2023

The Incongruence Principle of Evidence

Hillel J. Bavli
Southern Methodist University, Dedman School of Law

Author ORCID Identifier:

https://orcid.org/0000-0003-4888-2842

The Incongruence Principle of Evidence

HILLEL J. BAVLI*

Evidence law assumes that the meaning and value of information at trial is equal to the meaning and value of the same information in the real world. This premise underlies evidence policy, judicial applications of evidence law, and instructions to jurors for evaluating evidence. However, it is incorrect, and the law's failure to recognize this hinders its aims of accuracy and equality.

In this article, I draw on fields outside of law—including Bayesian inference and cognitive psychology—to develop a model of evidence that describes how jurors combine new evidence with prior beliefs (or "priors") to make inferences and judgments. I apply this model to derive a principle that I refer to as the incongruence principle of evidence. It states that the informational value of evidence at trial is not equal to the informational value of the same evidence in the real world. I show that, contrary to standard assumptions in evidence law, the trial setting degrades the value of evidence. It does this, on the one hand, by magnifying the influence of a juror's priors—including biases based on the race, gender, appearance, and other background characteristics of trial participants—and on the other hand, by employing a unique decisional framework that is susceptible to an inferential problem called "overfitting," which can lead to false-positive judgments when combined with biased and influential priors.

Finally, I show that the incongruence principle carries important implications for evidence law, including for the rule against character evidence, the hearsay rule, and impeachment evidence. I argue that recognizing the role of the incongruence principle in a juror's interpretation of evidence can help lawmakers to structure evidence law in a way that better achieves the goals of accuracy and equality.

---

* Associate Professor of Law, SMU Dedman School of Law. For helpful comments and exchanges regarding this and related work, I thank Edward Cheng, Bennett Capers, Michael Saks, Teneille Brown, Julia Simon-Kerr, Ronald Allen, Jonathan Koehler, Edward Imwinkelried, Eric Ruben, and the participants of the *2022 Evidence Summer Workshop* at Vanderbilt Law School. I also thank Ryan Demarest for his outstanding research assistance. I am grateful for generous financial support provided by the Law Faculty Excellence Fund.

I. INTRODUCTION

At the center of evidence law is an assumption that the meaning and value of information at trial is equal to the meaning and value of the same information in the real world.[1] This premise underlies federal and state evidentiary rules, judicial applications of evidence law, and instructions to jurors for evaluating evidence. However, it is incorrect. Information means different things in different contexts. It can give rise to one set of inferences for an observer in the real world and a completely distinct set of inferences for a juror at trial. Failure to account for this incongruence has arguably led to poor law and faulty judgments.

The law frequently relies on an oversimplified understanding of evidence as involving a fixed, abstract meaning or value. For example, law surrounding the admissibility of character evidence—evidence of a defendant's[2] prior acts or

---

1. The fundamental concept of evidentiary relevance is itself based on logical relevance and the tendency of the evidence, based on "acceptable generalization[s]" from the real world, to make a fact more or less likely. George F. James, *Relevancy, Probability and the Law*, 29 CALIF. L. REV. 689, 699 (1941); *see* GEORGE FISHER, EVIDENCE 22–25 (3d ed. 2013); *see also* FED. R. EVID. 401 advisory committee's note to 1975 proposed rule ("Whether the relationship [between the evidence and the case] exists depends upon principles evolved by experience or science, applied logically to the situation at hand."); CHRISTOPHER B. MUELLER, LAIRD C. KIRKPATRICK & LIESA L. RICHTER, EVIDENCE § 4.2, at 161 (6th ed. 2018); discussion *infra* Part II.

2. Neither the definition of character evidence nor the rule against character evidence is limited to defendants. For simplicity, however, throughout this Article, I often refer to a criminal defendant as the individual about whom evidence is offered at trial.

propensities offered to prove that the defendant acted in accordance with a certain character trait on the occasion in question[3]—relies on studies in the social sciences that examine the consistency of an individual's behavior across different circumstances.[4] However, these studies speak solely to the predictive value of character evidence in the real world: How well do prior acts actually predict future behavior?[5] What inferences and predictions can a real-world observer make based on an individual's past behavior? While these issues are certainly relevant to the probative value of character evidence at trial, the trial context is different from the real world, and these differences are critical for understanding the meaning and value of the evidence at trial. In other words, the law incorrectly assumes that the true predictive value of past behavior, or the inferential value for a real-world observer, determines the meaning and value of this information at trial. It incorrectly assumes that a person's past relevant behavior has the same meaning to a juror at trial as it would to a real-world observer.

Instead, evidence is observed and understood only in reference to its context. Information creates inferences that take shape when combined with an observer's broader perspective and prior beliefs.[6] This idea is encapsulated in the statistical principle known as Bayes' rule, which uses probabilities to describe how new evidence is combined with prior beliefs (or "priors") to arrive at an inference or judgment—i.e., a new, or "posterior," belief.[7]

This "Bayesian" framework for thinking about evidence reveals important problems. The trial setting itself tends to impact a juror's priors, thus changing the meaning of the trial's evidence in the juror's perception. It then augments the biasing effect of the juror's altered priors through its unique, confirmatory decisional framework. This leads to judgments that are heavily influenced by the juror's implicit biases, cognitive misperceptions, prior beliefs based on trial roles, and other prejudicial effects—and ultimately to incorrect verdicts.

Consider, for example, a scenario in which a law firm partner interviews a thirty-five-year-old woman for a prestigious position as an attorney at the law firm. During the interview, the partner learns various facts about the woman, including that she is well educated, has a number of years of relevant experience, and seems organized and hard-working. The partner also learns a few personal facts, including that the woman's husband is also an attorney, and that two of her young children tragically died of sudden infant death syndrome (SIDS), the sudden unexplained death of a baby less than one year old.

The partner may or may not ultimately hire the candidate. Either way, the hire is likely to be based on her qualifications and certainly not on the fact that two of her young children died. It is unlikely to occur to the interviewing partner that perhaps the woman murdered her children—even though the chances of two babies from

---

3. *See* FED. R. EVID. 404.

4. *See infra* note 35; *see also* FISHER, *supra* note 1, at 162–64, 218–33 (discussing studies regarding the predictability of behavior and providing legislative history for Rules 413–15 of the Federal Rules of Evidence).

5. *See infra* Part II.

6. *See infra* Section III.A.

7. *See* PETER D. HOFF, A FIRST COURSE IN BAYESIAN STATISTICAL METHODS 1 (2009).

the same family both dying of SIDS are quite low, and even assuming that the partner knows or has an idea about the statistical improbability of this occurrence.

However, in 1999, a UK jury indeed found a young woman named Sally Clark guilty of murdering her two young children based on little more than the fact that both of the children died allegedly (according to the defense) of SIDS.[8] Clark had features similar to those described in the above scenario—she was in her thirties, well-educated, and an attorney who practiced law with her husband at a UK law firm.[9] However, unlike the scenario above, the jury convicted Clark based on the improbability of two babies from the same family dying of SIDS.[10] Clark was ultimately exonerated and her conviction was overturned in a second appeal in 2003.[11]

To be sure, Clark's trial involved misleading statistical evidence that surely had an impact on the trial's outcome.[12] It is not my intention to suggest that the trial setting itself was the exclusive (or even the dominant) factor in bringing about Clark's conviction. However, the stark difference between the interviewing partner's interpretation and the jury's interpretation of the infants' deaths can likely be attributed to more than just the misleading statistical evidence. After all, at its core, the case's evidence involved little more than the rare occurrence of two unexplained infant deaths, and it is likely that both the jurors and the interviewing partner would be aware that such an occurrence is at least improbable.[13]

Rather, it is likely that the plain difference between the partner's interpretation and the jury's interpretation can be attributed—at least in part—to the different contexts and the respective priors of the interviewer and the jury. In the first scenario, the interviewing partner saw the SIDS deaths as a tragic set of incidents in a young attorney's background. In the second scenario, from the moment the jurors sat for trial, they had reason to suspect Sally Clark as a murderer. She had been presented to them as such by the prosecution, and although there is a legal presumption of innocence, the jurors cannot avoid viewing the defendant and the evidence in the context of the trial.

In the interview context, the question of whether the interviewee was a murderer is surely relevant in the sense that the partner would be unlikely to hire her had he suspected her of murdering her children. However, this possibility would be unlikely to even occur to him. Simple Bayesian reasoning (as well as a concept in

---

8. *See* R v. Clark [2003] EWCA (Crim) 1020, [1].

9. *Id.* at [8].

10. *See id.* at [96].

11. *See id.* at [181]. Tragically, after her release from prison in 2003, Sally Clark developed psychiatric illnesses and died of alcohol poisoning in 2007. *See* Sara Gaines & David Pallister, *Sally Clark's Death Accidental, Coroner Rules*, THE GUARDIAN (Nov. 7, 2007), https://www.theguardian.com/society/2007/nov/07/children.uknews [https://perma.cc/B6M2-2PG7].

12. *See* R v. Clark [2003] EWCA (Crim) 1020, at [94]–[110].

13. Presumably, the case also involved decisions by police, the government, and other case participants (e.g., the prosecutor's decision to charge Clark) that were unexposed to the faulty statistical evidence and yet still motivated by suspicion arising from the two infant deaths.

psychology known as the base-rate fallacy, described below)[14] provides a good explanation: Although the partner is sure to realize that it is unusual for two infants in the same family to die of SIDS, he is also sure to know that it is extremely unusual for a mother to murder her children. These pieces of information can be understood to reflect "new evidence" and a "prior," respectively, in a Bayesian framework. In this scenario, they counteract each other. Together, and considering the context of the job interview, the information regarding the deaths is unlikely to raise any suspicion. Standing alone, it is poor evidence of murder.

On the other hand, in Sally Clark's trial, the jurors were hyperfocused on the question, "Did Sally Clark murder her children?" They evaluated the evidence in light of this question, and their priors regarding Clark and the evidence had the opposite effect compared to the priors of the interviewer. The jurors' priors are likely to have contributed significantly to their finding that Clark committed murder notwithstanding the extraordinarily low likelihood in the real world that a mother would murder her children.

This is certainly not to say that trials necessarily lead to incorrect outcomes or that jurors should or should not view the evidence in the context of the broader trial. My intention is not to criticize the structure or analytical framework of US trials. Rather, it is to demonstrate and formalize the idea that the context in which evidence is presented is critical to the meaning and informational value of the evidence.

In this article, I build on my recent work[15] and draw on fields outside of law—including Bayesian inference and cognitive psychology—to develop a model of evidence that highlights the central role of an observer's prior beliefs in the formation of inferences from evidence. I apply this model—called the "reflective model"[16]—to derive a principle that I refer to as the *incongruence principle of evidence*. It states that the informational value of evidence at trial is not equal to the informational value of the same evidence in the real world; to the contrary, for most evidence, the trial setting degrades the value of evidence.

Specifically, the reflective model reveals two features of a trial that alter the meaning and value of information relative to its meaning and value in the real world. First, the trial setting magnifies the strength and influence of a juror's prior beliefs and prejudices regarding the background characteristics of the defendant and other trial participants.[17] A juror's perception of the evidence in a case is contextualized by trial roles and the juror's prior beliefs. I show how the trial process intensifies known cognitive biases—such as confirmation bias and the base-rate fallacy—that may hinder a juror's ability to make accurate judgments. I then show how these biases combine with a juror's preexisting beliefs, prejudices, and expectations regarding trial participants to create misleading and influential priors.

---

14. *See infra* Section III.D.1.

15. *See* Hillel J. Bavli, *Character Evidence as a Conduit for Implicit Bias*, 56 U.C. DAVIS L. REV. 1019 (2023) [hereinafter *Character Evidence as a Conduit for Implicit Bias*].

16. *See id.* at 1039–70.

17. As explained below, the priors of judges, witnesses, and other trial participants are also important. *See infra* Parts III–IV.

Second, unlike most real-world contexts, the trial setting involves an analytical framework that is, in many respects, analogous to a scientific hypothesis test—a method for testing whether a certain claim or hypothesis is supported by observed evidence.[18] As in scientific hypothesis testing, a trial involves a default, or "null," hypothesis (e.g., not guilty) that is rejected in favor of an alternative hypothesis (e.g., guilty) if there is sufficient evidence against it based on a predefined threshold, such as the reasonable-doubt standard in criminal cases.[19]

As in the sciences, however, this mode of analysis is susceptible to a problem called "overfitting," which involves misinterpreting random occurrences (or "noise") in the evidence as having true predictive or explanatory value for the matter in question, resulting in false-positive inferences.[20] Just as a scientific researcher observes a sample of data and models it to make inferences for future samples—essentially, developing a story or an accounting of the data to make predictions and conclusions—a juror observes evidence and develops a model or story around it to make inferences for the judgment in question. An overfit model—whether in law or science—overexplains the evidence, misinterpreting randomness for valuable information and rendering poor predictions and false judgments.[21]

Worse, as in the sciences, strong and influential priors tend to intensify this problem and bring its danger to fruition via known psychological biases that cause observers—whether researchers or jurors—to interpret evidence in a way that confirms their prior beliefs and expectations.[22] Thus, while the trial setting creates strong and biased priors, it also compounds the detrimental effects of these priors by employing an analytical framework that is particularly susceptible to their biasing influence. In short, the way that a juror processes information at trial is detrimentally different from the way that an observer processes information in the real world. This difference impacts a juror's perception of the evidence, and it systematically heightens the risk of certain cognitive biases and ultimately the biasing influence of prior beliefs.

Finally, after introducing the incongruence principle, I demonstrate its implications for a broad range of evidentiary contexts. First, I examine the sharp trend in the courts toward a more permissive approach to character evidence. I apply the incongruence principle to demonstrate that this trend is built on false assumptions. Specifically, I show that character evidence creates a perfect storm

---

18. *See generally* MORRIS H. DEGROOT & MARK J. SCHERVISH, PROBABILITY AND STATISTICS 530–623 (4th ed. 2012) (explaining hypothesis testing); DAVID COPE, FUNDAMENTALS OF STATISTICAL ANALYSIS 36–41 (2005) (explaining hypothesis testing).

19. *See* COPE, *supra* note 18, at 40 (discussing statistical significance).

20. *See infra* Section III.D.2; *see also* Tal Yarkoni & Jacob Westfall, *Choosing Prediction Over Explanation in Psychology: Lessons from Machine Learning*, 12 PERSPS. ON PSYCH. SCI. 1100, 1102 (2017) ("The tendency for statistical models to mistakenly fit sample-specific noise as if it were signal is commonly referred to as *overfitting*.").

21. *See infra* Section III.D.2.

22. *See* Hillel J. Bavli, *Credibility in Empirical Legal Analysis*, 87 BROOK. L. REV. 501, 509–14 (2022) [hereinafter *Credibility in Empirical Legal Analysis*]; discussion *infra* Section III.D.3.

for the devaluation of evidence: it can be extremely detrimental to the aims of evidence law even while being very probative in many real-world settings.

Second, I discuss related implications for impeachment evidence. I show that the false assumptions underlying the current trend toward a more permissive treatment of character evidence also entail important implications for witness-credibility evidence.

Third, I discuss implications for the perceptions of judges and witnesses. I emphasize that judges and witnesses are also observers at trial, such that the trial setting impacts their perception and interpretation of evidence just as it does for jurors. This carries important implications for a judge's evidentiary decisions and a witness's testimony, thus compounding the effects of devaluation via the influence of a juror's priors.

Fourth, I discuss implications for the rule against hearsay evidence, which limits the admissibility of out-of-court statements.[23] Specifically, I show that, because hearsay evidence involves inferences that derive first from a real-world observer of an event in question, followed by a trial-witness observer of an out-of-court statement, followed by a juror's observation of the witness's testimony, the incongruence principle carries important—although complex—implications for the rule against hearsay. In some hearsay contexts, the principle may even favor the admissibility of hearsay evidence, reflecting advantages of certain out-of-court statements regarding an underlying event over analogous statements at trial.

Lastly, I discuss critical implications for the aims of equal treatment and preventing the influence of biases on the basis of race, gender, appearance, economic status, education, and other background characteristics of litigants and witnesses. In particular, I highlight the significant role of implicit bias in the prior beliefs of jurors and other trial participants, and I show that the devaluing effect of the trial on evidence disproportionately disadvantages certain groups of litigants based on their background characteristics. I thus show how the incongruence principle connects concerns for accuracy and equality and how the trial setting interacts with a juror's prior beliefs and prejudices to devalue evidence and cause inaccurate judgments.

I proceed as follows: In Part II, I discuss the assumption in legislation, judicial opinions, and scholarship that the meaning and value of evidence at trial is equal to the meaning and value of the same evidence in the real world. In Part III, I describe the reflective model of evidence in simple, nontechnical terms, and I carefully define the concept of evidentiary value. I then apply this model to derive and examine the incongruence principle. In Part IV, I apply the incongruence principle in a range of evidentiary settings, and I discuss its important implications. In Part V, I conclude.

II. A FALSE PREMISE AT THE CENTER OF EVIDENCE LAW

Evidence law is premised on the idea that the meaning and value of evidence at trial and in the real world are equal. In this Part, I show that this basic—but

23. *See* FED. R. EVID. 801(a)–(c); discussion *infra* Section IV.D.

ultimately incorrect—assumption is a central pillar around which evidence law is built.

Even the most fundamental principles of evidence law seem to incorporate this assumption of equivalence. We can see this in the law's heavy reliance on real-world experience and science, combined with the absence of conditions or qualifications accounting for important differences between the trial setting and the real world.

Consider, for example, the notion of relevance—evidence law's gateway requirement for admissibility.[24] Under Federal Rule of Evidence 401, "[e]vidence is relevant if: (a) it has any tendency to make a fact more or less probable than it would be without the evidence; and (b) the fact is of consequence in determining the action."[25]

But what determines the relationship between evidence and a case? In other words, what distinguishes an appropriate inference from an inappropriate one? The answer lies in the everyday reasoning of individuals in the real world. Evidence is relevant if it has a tendency, based on "acceptable generalization[s]" from the real world, to make a fact more or less likely.[26] As the Advisory Committee notes to Rule 401 explain, "Whether the relationship [between the evidence and the case] exists depends upon principles evolved by experience or science, applied logically to the situation at hand."[27] Similarly, "[i]n ruling upon relevancy, the court must draw on its own experience, knowledge, and common sense in assessing whether a logical relationship exists between proffered evidence and the fact to be proven."[28] Scientific results may also assist a court in determining an association that may otherwise be difficult to discern.[29]

In short, in the eyes of the law, "[r]elevancy boils down to what our human experiences tell us, sometimes called common sense."[30] As one leading treatise explained:

> [H]ow can a judge know whether the evidence could reasonably affect an assessment of the probability of the fact to be inferred? In some instances, scientific research may show that the fact in issue is more likely to be true (or false) when such evidence is present than when it is not. *Ordinarily, however, the answer must lie in the judge's own experience, [the judge's] general knowledge, and [the judge's] understanding of human conduct and motivation.*[31]

---

24. *See* FED. R. EVID. 401.

25. *Id.*

26. James, *supra* note 1, at 699; FISHER, *supra* note 1, at 22–25.

27. FED. R. EVID. 401 advisory committee's note to 1975 proposed rule; *see also* MUELLER ET AL., *supra* note 1, § 4.2, at 161.

28. MUELLER ET AL., *supra* note 1, § 4.2 at 161.

29. *See id.*

30. United States v. Curtin, 489 F.3d 935, 948 (9th Cir. 2007); *see* MUELLER ET AL., *supra* note 1, §4.2, at 161.

31. *Curtin*, 489 F.3d at 948 (quoting 1 KENNETH S. BROUN, GEORGE E. DIX, EDWARD J. IMWINKELRIED, D. H. KAYE, ROBERT P. MOSTELLER, E. F. ROBERTS, JOHN W. STRONG & ELEANOR SWIFT, MCCORMICK ON EVIDENCE § 185, at 736 (6th ed. 2006) (modifications and

Moreover, a court's judgments about unfair prejudice—and its balancing of probative value and unfair prejudice under Rule 403[32] or an analogous state rule—also rely on the judge's real-world knowledge and experience, as well as inferences regarding a juror's real-world knowledge.[33]

Similarly, evidence law surrounding substantive categories of evidence, such as character evidence and hearsay, is grounded in common-sense inferences that one would draw from such evidence in the real world—generally without accounting for differences between the real world and the trial setting.

For example, policy surrounding the rule against character evidence[34] is, in part, grounded in behavioral studies that examine the consistency of an individual's behavior across different circumstances.[35] Moreover, policy surrounding the sexual assault exceptions to the rule against character evidence is grounded, in part, in the theory that individuals who commit sexual assault have a greater propensity than other criminal offenders to reoffend.[36] Similarly, California's statute establishing an exception to its rule against character evidence for "evidence of the defendant's commission of other domestic violence"[37] is based on the stated purpose "that in domestic violence cases, as in sex crimes, similar prior offenses are uniquely probative of guilt in a later accusation."[38]

---

italicization in original)).

32. FED. R. EVID. 403 ("The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence.").

33. *See, e.g.*, Wallace v. State, 932 S.W.2d 345, 347 (Ark. Ct. App. 1996) ("[W]e cannot say that the trial court was wrong in its apparent conclusion that the probative value of the evidence outweighed the danger of its *unfair* prejudice. After all, in considering the evidence, the jury is not required to set aside its common knowledge but has the right to consider all the evidence in the light of its own observations and experiences in the affairs of life."); Sweers v. Westfall, No. 09-0912, 2010 WL 786036, at *4 (Iowa Ct. App. March 10, 2010) ("First, jurors bring their common experiences to the courtroom. Certainly one of those common experiences is that motorists in the State of Iowa are required to and generally do carry insurance. Another of those common experiences is that when people are involved in a car accident, regardless of who is at fault, they frequently contact their insurance agent afterward. It is difficult to believe the admission of [appellant's] statement altered the collective wisdom that this jury already had before trial commenced.").

34. *See* FED. R. EVID. 404; *see also supra* note 3 and accompanying text.

35. *See, e.g.*, State v. Martinez, 195 P.3d 1232, 1236 (N.M. 2008) ("Modern scientific research now confirms what human beings have always observed in their own family and community relationships, that the average person is able to explain, and even predict, a subject's behavior with a significant degree of accuracy.") (citing studies); State v. Sullivan, 679 N.W.2d 19, 23–25 (Iowa 2004) (discussing studies regarding consistency in behavior relative to weight afforded to character evidence); State v. Dellay, 687 A.2d 435, 438 n.1 (R.I. 1996) (discussing research regarding the consistency of behavior in relation to character evidence).

36. *See* FED. R. EVID. 413–15; FISHER, *supra* note 1, at 218–26 (containing legislative history and other sources).

37. CAL. EVID. CODE § 1109 (West 2006).

38. People v. Johnson, 110 Cal. Rptr. 3d 515, 524 (Cal. Ct. App. 2010) (internal quotation marks omitted). The court continued by explaining that "proponents of the bill that became section 1109 argued for admissibility of such evidence because of the 'typically

In addition, policy surrounding the rule against hearsay and its exceptions is grounded in concern for the reliability of the hearsay statement. Hearsay is an out-of-court statement offered into evidence for the truth of the matter asserted in the statement.[39] For example, if David witnesses a car accident and later tells Wanda that the driver of a white minivan ran a red light, and then Wanda's testimony regarding this statement is offered at trial to prove the minivan driver's fault, the evidence is defined as hearsay and is generally inadmissible.[40] However, there are many exceptions to the rule against hearsay.[41] The rule itself is based on the concern that, *inter alia*, hearsay generally cannot adequately be tested through cross examination and is therefore unreliable.[42] Its exceptions are frequently based on assurances of reliability grounded in the circumstances surrounding the out-of-court statement.[43] For example, the excited utterance exception is based on the idea that people tend not to lie when making "[a] statement relating to a startling event or condition, made while . . . under the stress of excitement that it caused."[44]

However, the reliability concerns underlying the rule against hearsay and its exceptions relate predominantly, if not almost entirely, to real-world considerations. The hearsay exceptions, for example, are based on indicia of reliability for real-world inferences—for instance, that people are more reliable when responding to a startling event.[45] This is not to say that these indicia necessarily do not extend to inferences at a trial. Again, the point is only to emphasize that courts and legislatures assume a symmetry between real-world inferences and inferences at trial—between the meaning and value of the evidence (and the out-of-court statement in particular) in the real world and at trial. They do not consider, for example, whether indicia of reliability for an out-of-court statement, while holding true for an observer in the real world, might not hold true at trial. Such assurances of reliability may hold true as they pertain to the inferences of a real-world observer but may weaken or altogether dissipate when applied to an observer at trial.[46]

---

repetitive nature' of domestic violence. This pattern suggests a psychological dynamic not necessarily involved in other types of crimes." *Id.* (internal footnotes and citations omitted).

39. FED. R. EVID. 801.

40. *See id.* 801–02.

41. *See id.* 803–04.

42. *See* FISHER, *supra* note 1, at 378–79.

43. *Id.* at 407. The exceptions are also grounded in a rationale of necessity since the declarant is often unavailable at trial. *Id.*

44. FED. R. EVID. 803(2).

45. *See generally id.*

46. Note that there are instances in which the rules of evidence do contemplate the reliability of out-of-court statements regarding an event relative to in-court testimony regarding the same event. For example, for the special circumstances surrounding statements of identification, Rule 801(d)(1)(C) permits out-of-court statements that "identif[y] a person as someone the declarant perceived earlier" if "[t]he declarant testifies and is subject to cross-examination about [the] prior statement." *Id.* 801(d)(1)(C). This exclusion from hearsay is based on the relative reliability of out-of-court identifications. As one leading treatise explained, "Ordinarily, when a witness is asked to *identify* the assailant, or thief, or other person who is the subject of his testimony, the witness' act of pointing out the accused (or other person), then and there in the courtroom, is of little testimonial force. After all that has intervened, it would seldom happen that the witness would not have come to believe in

These and many other examples illustrate that evidence law is grounded in the assumption that the meaning and value of information observed in the real world—both in terms of its probative or predictive value and in terms of its unfair prejudice—extend to the same information observed at trial. This is certainly not to say that scientific studies or real-world human experience, logic, and intuition regarding an item of evidence are not important in determining the value of evidence at trial. I examine this relationship in the following Part. In the current Part, however, my aim is to emphasize that courts, legislatures, and scholars in fact rely on such studies and the probable inferences of real-world observers as the basis of principles and policies underlying evidence law.

Further, in addition to shaping policies surrounding the rules of evidence, courts and litigants rely on a juror's ability to draw on their common sense and real-world experience to make inferences and judgments. Consistent with the language of evidentiary relevance, courts generally instruct jurors to apply their common sense and experience to evaluate evidence and determine judgments. For example, a Texas model criminal jury charge provides:

> While you should consider only the evidence, you are permitted to draw reasonable inferences from the testimony and exhibits that are justified in the light of common experience. In other words, you may make deductions and reach conclusions that reason and common sense lead you to draw from the facts that have been established by the evidence.[47]

Indeed, it is standard to instruct jurors to follow their real-world common sense and experience, and it would be unusual for a court to account for differences

---

the person's identity." Commonwealth v. Weichell, 453 N.E.2d 1038, 1043–44 (Mass. 1983) (quoting 4 J. WIGMORE, EVIDENCE § 1130, at 277–79 (Chadbourn rev. ed. 1972)) (internal quotation marks omitted). In this unique context, evidence law implicitly reflects concern regarding the prior beliefs of jurors. In addition, in the broader hearsay context, because courts and legislatures are forced to confront the reliability of out-of-court (i.e., real-world) statements relative to testimony at trial, they have engaged in at least implicit comparisons. However, these comparisons are primarily concerned with issues of memory, perception, sincerity, the potential for threats, and other trustworthiness factors unrelated to the prior beliefs of jurors. *See generally* FISHER, *supra* note 1, at 510–74 (discussing rationales underlying exceptions under Rule 803). Although, in certain circumstances, such comparisons may also incidentally reflect such prior beliefs, the focus is generally on the trustworthiness of the evidence itself rather than on the impact of the trial setting on an observer's prior beliefs. *See generally* FED. R. EVID. 803 (exceptions that do not depend on the declarant's availability); FISHER, *supra* note 1, at 510–11 (discussing the Advisory Committee's suggestion that, unlike hearsay admitted under Rule 804, "hearsay admitted under Rule 803 sometimes is *better* than the declarant's live testimony"). Such comparisons are generally not based on an account of evidentiary value that involves a combination of new evidence and prior beliefs. *See infra* Part III.

47. STATE BAR OF TEX., TEXAS CRIMINAL PATTERN JURY CHARGES, CRIMES AGAINST PERSONS § C2.1, at 34–35 (2011). There is no doubt that common-sense inferences from evidence in the real world appropriately play an important role in fact-finding at trial. *See generally* FREDERICK SCHAUER, THE PROOF: USES OF EVIDENCE IN LAW, POLITICS, AND EVERYTHING ELSE (2022) (discussing the use of evidence to discover truth in the real world as well as in the courtroom).

between inferences in the real world and at trial. As one court has stated, "[t]he instructions in question were the standard instructions, given in nearly every jury trial, that jurors may draw on 'common sense,' 'experience,' and reasonable 'inferences' when evaluating the evidence."[48] The defendant in that case argued that "the [d]istrict [c]ourt's limiting instruction relating to [the proper and improper use of other-acts evidence offered for non-character purposes under Rule 404(b)] were effectively negated by contradictory or confusing instructions elsewhere in the charge."[49] However, the Third Circuit rejected the argument, finding "meritless" the suggestion that the court's invitation for the jury to rely on common sense and experience negated the court's limiting instruction and constituted "an invitation to consider propensity evidence."[50]

Similarly, courts expect lay witnesses—another category of trial participants and observers[51]—to rely on their real-world common sense and experience to make inferences in their testimony. For example, "[t]o prove that a defendant deviated from the standard of care, a plaintiff need not rely on expert testimony where the alleged negligent act is within the realm of common knowledge and everyday experience."[52] Indeed, "lay testimony results from a process of reasoning that is familiar in everyday life,"[53] and courts, again, generally do not account for differences between information in the real world and at trial.

In general, in neither the underlying policy nor the application of evidence law do legislatures or courts consider the effects of the trial setting itself on the observer of evidence when the observer applies her common sense and real-world experience to make inferences at trial. Instead, the meaning and value of information observed in the real world—both in terms of its probative or predictive value and in terms of its unfair prejudice—determine evidence policy (e.g., whether character evidence should generally be admissible or inadmissible), the courts' application of evidence law (e.g., whether a court finds a particular item of evidence to be relevant or admissible under the rules of evidence), and the courts' expectations and instructions for how jurors should evaluate evidence and determine verdicts.

This is not to say that courts do not evaluate evidentiary matters like probative value and unfair prejudice within the context of a trial. For example, in determining whether evidence should be excluded on the basis of its unfair prejudice under Rule

---

48. United States v. Homick-Van Barry, 240 F. App'x 966, 969 (3d Cir. 2007).

49. *Id.*

50. *Id.* at 969–70; *see also* United States v. Gainey, 111 F.3d 834, 836 (11th Cir. 1997) ("In evaluating the facts of a case, the law permits jurors to 'apply their common knowledge, observations and experiences in the affairs of life.' Such an instruction recognizes that in assessing credibility or the reasonableness of a position, people inherently apply conclusions about human behavior based on *common experiences of daily living*. For example, jurors may use 'common sense,' derived from the repetitive pattern of human behavior and experiences common to all of us, in discerning the reliability of a person who gives conflicting testimony.") (emphasis in original) (citation omitted).

51. *See infra* Sections IV.C–D.

52. Butera v. District of Columbia, 235 F.3d 637, 659 (D.C. Cir. 2001) (internal quotation marks and citations omitted).

53. United States v. Gyamfi, 805 F.3d 668, 674 (6th Cir. 2015).

403, a court may evaluate the likely effectiveness of a limiting instruction—an explicit instruction to the jury on the permissible and impermissible uses of the evidence.[54] This evaluation would surely consider the evidence and instruction within the context of the trial—to determine the effectiveness of the instruction for a juror in the trial setting in particular, rather than for an observer in the real world. Similarly, under Rule 403, courts evaluate the unfair prejudice associated with an item of evidence within the context of the trial.[55]

However, the primary focus is on the evidence itself, and although a court's analysis may well account for the jurors' perceptions, the evidence is generally evaluated with the basic assumption that jurors perceive the information as they would as an observer in the real world. Courts generally do not consider the effects of the trial setting on the jurors' general perception or their priors about the evidence, the case, or trial participants. If they do, it is only incidental and implicit, on a case-by-case basis, based on their evaluation of the probative value and unfair prejudice associated with an item of evidence within the context of a trial.

At the very least, therefore, the problem can be summarized as follows: Policy surrounding the rules of evidence is frequently developed on the basis of the perceived predictive value of evidence (based on, e.g., scientific studies) or the likely inferences of observers of the evidence in the real world. Similarly, judicial applications of evidence law are frequently made on the basis of inferences that an observer could or would make from evidence in the real world—notwithstanding a court's evaluation of the evidence in the context of a trial. In addition, courts frequently instruct jurors on how to evaluate evidence and arrive at a verdict on the basis of the expectation that jurors at trial evaluate evidence as they would in the real world. Finally, scholarship surrounding the rules of evidence frequently assumes the same—that the inferences that can be expected to result from evidence observed in the real world extend to the same evidence observed by jurors at trial.[56]

Legislatures, courts, and scholars generally ignore important differences between a real-world observer's prior beliefs and those of a juror at trial. When evidentiary rulings reflect such differences, it is usually only incidental to a court's analysis of the evidence in the trial setting. Consequently, in developing and

---

54. *See* FED. R. EVID. 105, 403.

55. *See* FED. R. EVID. 403 advisory committee's note to proposed rule ("'Unfair prejudice' within its context means an undue tendency to suggest decision on an improper basis, commonly, though not necessarily, an emotional one.").

56. *See, e.g.*, Teneille R. Brown, *The Content of Our Character*, 126 PENN. ST. L. REV. 1, 6 (2021) ("[T]his basic psychological process of forming immediate impressions of others lies beneath *all* of our social interactions. . . . All day long we constantly make predictions about others' behaviors and intentions based on their past actions and our assessments of their characters, which we incorrectly assume to be stable. Because this process is spontaneous and subconscious, we cannot stop doing it when we become jurors in trials.") (footnote omitted); *id.* at 10 ("We must revise the character evidence rules to reflect what psychologists already know—that people cannot *not* make character inferences.") (emphasis in original); Roger Park, *A Subject Matter Approach to Hearsay Reform*, 86 MICH. L. REV. 51, 54 (1987) ("The primary argument for admitting hearsay is simple and powerful: hearsay can be convincing evidence, and it is the sort of evidence on which we routinely rely in the most important affairs of home, state, and business.").

applying evidentiary rules, lawmakers neglect the substantial differences between the inferences that an observer would or should make from evidence in the real world and the inferences that a juror would make at trial.

However, as I argue in the following Part, notwithstanding the actual predictive value of evidence for the fact in question or the inferences that a real-world observer would make from the evidence, the value of evidence at trial is based on the inferences that a juror *at trial* will draw from the evidence, and these inferences are based on *both* the real-world value and meaning of the evidence on the one hand and the juror's priors on the other hand. In turn, the juror's priors at trial—and therefore the juror's inferences from the evidence—are substantially different than those of an observer in the real world, and they do not necessarily reflect the true predictive value of the evidence or the inferences that a real-world observer would make for the fact in question.

III. THE INCONGRUENCE PRINCIPLE

In this Part, I argue that the premise discussed in Part II—that the meaning and value of information at trial is equal to the meaning and value of the same information in the real world—is incorrect. I begin by introducing the reflective model of evidence, which describes a Bayesian process whereby case evidence is combined with a factfinder's prior beliefs to generate inferences and fact judgments. I then discuss the concept of evidentiary value. Specifically, I define a measure that I call the *ev ratio*, a quantification of evidentiary value that relates the probative value of an item of evidence to its unfair prejudice. Finally, I apply the reflective model to examine the impact of a trial on the value of evidence. In particular, I demonstrate the incongruence principle and show that, in many circumstances, the trial setting causes a significant devaluation of evidence with respect to the inferences that an observer draws from it at trial.

A. The Reflective Model

The reflective model is premised on the idea that jurors process evidence logically, in line with the principle in statistics known as Bayes' rule.[57] To reiterate, this means that jurors combine their prior beliefs with new evidence to arrive at a judgment—a posterior belief that accounts for the new evidence.[58] For example, a

---

57. *Character Evidence as a Conduit for Implicit Bias*, *supra* note 15 at 1026–27.

58. *Id.*; *see also* Andrew Gelman, *Bayesian Statistics: What's It All About*, STAT. MODELING, CAUSAL INFERENCE & SOC. SCI. (Dec 13, 2016), https://statmodeling.stat.columbia.edu/2016/12/13/bayesian-statistics-whats/ [https://perma.cc/Z94W-MRF7] ("Bayesian statistics uses the mathematical rules of probability to combine data with prior information to yield inferences which (if the model being used is correct) are more precise than would be obtained by either source of information alone"); HOFF, *supra* note 7, at 1 ("[I]t can be shown that probabilities can numerically represent a set of rational beliefs, that there is a relationship between probability and information, and that Bayes' rule provides a rational method for updating beliefs in light of new information. The process of inductive learning via Bayes' rule is referred to as *Bayesian inference*.") (emphasis in original).

person who lives in a desert may have strong priors regarding the likelihood that it will not rain. But if she wakes up and sees dark clouds in the sky, she will update her beliefs to reflect this new evidence. Whether she brings an umbrella depends on how heavily she weights the new evidence against her priors.

In a previous article, titled *Character Evidence as a Conduit for Implicit Bias*, I developed the reflective model to describe how jurors incorporate character evidence into a verdict, and particularly, to explain the influence of a juror's prior beliefs—including implicit biases regarding race, appearance, and other background characteristics of a defendant—in determining a verdict based on character evidence.[59] Similar to the desert example above, a juror updates her prior beliefs in light of new evidence, and the influence of her priors depends on how heavily she weights the new evidence in updating her beliefs. To answer this question, the reflective model distills a basic principle from Bayesian statistics that I call the "shrinkage principle," which is based on a statistical idea called "shrinkage estimation."[60] The shrinkage principle states that when there are two sources of information—prior beliefs and new evidence—a juror will combine these sources of information by weighting each in proportion to how precisely it speaks to the matter in question relative to the other source of information.[61] Precision here refers to the inverse of variability or uncertainty.[62] It is distinct from accuracy, which measures the proximity of a measurement to its true value.[63] Precision is a measure of informational sharpness, and for purposes herein, it involves two components: the uniformity and the strength of the information.[64] Uniformity reflects the tightness or similarity of the information with respect to itself and the event in question while strength reflects both the quantity and certainty surrounding the information.[65] So, priors will be relied on most heavily when they speak precisely to the matter in question—in terms of both uniformity and strength—and when the new evidence is imprecise.[66] Conversely, they will be relied on less heavily when they speak only imprecisely to the matter in question or the new evidence is relatively precise.[67]

For example, assume that two young brothers get into a physical altercation. Each blames the other, and their mother is interested in determining which of the brothers started the fight and which simply defended himself. She has prior beliefs regarding the matter: based on previous experiences, the mother believes that one of the brothers—the older one—is most likely to have been the aggressor. However, she also collects "new evidence." For example, she asks each brother to explain what happened, and she observes their physical injuries. Therefore, to arrive at a judgment, she combines her priors with the new evidence. Pursuant to

---

59. *Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 1039–69.
60. *Id.* at 1053–55.
61. *Id.* at 1041–54.
62. *Id.* at 1042.
63. *Id.* at 1041.
64. *Id.* at 1051–53.
65. *Id.*
66. *Id.* at 1054.
67. *Id.*

the shrinkage principle, she is likely to weight each in proportion to its relative precision.

For instance, if the older brother was clearly and consistently the aggressor in many previous similar incidents, she will rely more heavily on her priors than if his role in the prior incidents was less clear or less frequent (reflecting less evidential strength) or was less consistent or less similar to the current incident (reflecting less evidential uniformity). At the same time, her reliance on her priors depends not only on their precision with respect to the matter in question, but on their precision *relative to* that of the new evidence. For example, she would rely on her priors more heavily if the new evidence involved simply different stories from each brother without any further evidence (providing only a weak informational signal), and she would rely on them less heavily if a clearer informational signal emerges from the evidence—for instance, if the brothers' stories both clearly point to one brother as the aggressor, and the observed physical injuries are consistent with their stories (all reflecting the uniformity and strength of the new evidence).[68]

Similarly, when prior-bad-acts character evidence is offered at a criminal trial, according to the reflective model, the evidence can be understood as informing a juror's prior beliefs regarding the defendant's character.[69] At trial, other-acts character evidence plays a role similar to the prior incidents in the fighting-brothers scenario: although not *directly* relevant to the act in question, it provides information about the defendant's propensity to act in a certain way and thereby about the likelihood that the defendant committed the act in question.[70] Just as the prior fighting incidents informed the mother's prior beliefs, character evidence provides context for the matter in question—in a sense, a prior to be combined with "new evidence" (such as physical evidence or eyewitness testimony) *directly* regarding the matter in question to determine a judgment.

Character evidence is in fact a special case (relative to other types of evidence) that involves *two* levels of priors regarding the defendant's character. Specifically, a juror brings to trial a set of preexisting beliefs and prejudices regarding the background characteristics of the defendant.[71] These may arise from a wide range of factors in the juror's life experience that influence the juror's beliefs and perspectives. The juror combines these *priors* with character *evidence* offered at trial to form an assessment of the defendant's character. This character judgment then serves as a new prior with respect to the "new," non-character, evidence offered in the case.[72] This prior (i.e., the juror's judgment regarding the defendant's

---

68. The prior can serve to confirm the new evidence, or it can pull in the opposite direction of the new evidence. For example, if there is weak new evidence that one brother started the fight while there is a strong prior in the opposite direction, the prior may dominate the judgment even if it is inconsistent with the new evidence.

69. Hillel J. Bavli, *An Aggregation Theory of Character Evidence*, 51 J. LEGAL STUD. 39, 49–53 (2022) [hereinafter *An Aggregation Theory of Character Evidence*].

70. As discussed in Part IV, the reflective model describes character evidence with some additional complexity to reflect the fact that there are two levels of prior beliefs regarding a defendant's character—one set of prior beliefs that a juror brings to trial and one set of prior beliefs that emerges from the prior-bad-acts *evidence* at trial.

71. *Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 1055–65.

72. *Id.*

character) and the non-character event evidence in the case are ultimately combined to form a verdict.[73]

Thus, *Character Evidence as a Conduit for Implicit Bias* develops the reflective model to explain the role of a juror's prior beliefs and prejudices when a court admits character evidence. In the following Section, I generalize the reflective model by extending the same Bayesian reasoning to the broader evidence landscape.

B. Generalizing the Reflective Model

Without getting into the technical properties of shrinkage,[74] the idea of the generalized reflective model is this: Jurors come to a trial with a range of preexisting beliefs, prejudices, and stereotypes—that is, prior beliefs. When they hear evidence, they combine these priors with the evidence to form fact judgments and ultimately a verdict. Of course, they do not simply attach *all* of their prior beliefs to each and every piece of evidence. Pursuant to the shrinkage principle, they rely on relevant priors more or less heavily based on how precisely the priors speak to the matter in question and also based on how imprecisely the case evidence speaks to the matter in question.[75]

For example, I have argued that a juror's priors regarding a defendant's background characteristics—and thus regarding the defendant's character—will be more influential when the court admits character evidence and explicitly sanctions a character judgment and character inferences than when a court forbids character evidence and emphasizes to the jury the importance of focusing on the concrete evidence in the case.[76] The character evidence in effect *activates* the juror's priors regarding the defendant's character.[77]

This activation is based on the shrinkage principle. The juror's priors will play a greater role when there is relatively weak event evidence in a case, leaving large evidentiary gaps, and when the prosecutor relies on character evidence to fill these gaps.[78] In this case, the issue of the defendant's character is salient, and a juror's priors regarding the defendant's character speak precisely to the issue. Contrast this with a case in which the focus is on concrete event evidence regarding the matter in question—such as fingerprint evidence, eyewitness testimony, and the defendant's statements to the police—and in which the court disallows character evidence and instructs the jury to determine a verdict based on the evidence.[79] In this case, the issue of the defendant's character—although undoubtedly still an issue in the jurors' minds—is less salient. The jurors' priors regarding the defendant's character are therefore less precise with respect to the immediate matter in question—now the act at issue in the case and not the defendant's character—and

---

73. *Id.*

74. For an in-depth explanation of how shrinkage estimation applies to character evidence and other legal contexts, see *id.* at 1045–55.

75. *Id.* at 1053–65.

76. *Id.* at 1055–65.

77. *Id.* at 1060.

78. *Id.* at 1055–66.

79. *Id.*

so they are less influential.[80] This is especially so in light of the relatively precise event evidence in the case.

The reflective model is easily extended to other evidentiary contexts to describe the influence of prior beliefs on a fact judgment or a verdict. First, there are various evidentiary contexts that involve character evidence in one form or another. For example, the witness-credibility context, in which federal and state evidentiary rules loosely permit evidence regarding a witness's character for truthfulness or untruthfulness, involves the same issues as in the discussion above.[81] In this setting, courts regularly admit evidence of a witness's prior dishonest acts—and even their prior bad acts more generally—to demonstrate the witness's character for dishonesty.[82] As above, a juror will have prior beliefs regarding the witness's character for truthfulness based on the witness's background characteristics. According to the shrinkage principle, these priors are combined with new evidence regarding the witness's character for truthfulness to form a judgment regarding the witness's credibility and, ultimately, the veracity of the witness's testimony.

Moreover, the reflective model applies to prior beliefs other than those regarding character. To illustrate the reflective model in a more general context, consider witness testimony that a defendant in an attempted murder trial acted in self-defense when he shot at the alleged victim in the case. The witness testifies that the alleged victim was a driver who exchanged words with the defendant, a pedestrian who, according to the witness, drew his firearm only after the alleged victim emerged from his car with a knife to attack the defendant. Two jurors may assess this evidence very differently depending on their respective priors.

Both jurors may have strong priors regarding the use of firearms. One may have prior beliefs that cause her to interpret this evidence very favorably for the defendant while the other may have prior beliefs that cause her to interpret the defendant's use of a gun—and perhaps even his possession of a gun—as signaling aggression. The reflective model formalizes and makes explicit the idea that the inferences that a juror draws from the evidence may be as dependent on the juror's own priors as on the evidence itself.

As above, pursuant to the shrinkage principle, the influence of a juror's prior beliefs will depend on the precision of the priors relative to the precision of the evidence. For instance, in the self-defense example, if the evidence is overwhelming in one direction or the other, two jurors with opposing priors are nevertheless likely to determine the same judgment. Moreover, the precision of the prior is important: A prior about the use of guns is likely to be less influential (although perhaps still somewhat influential) if the defendant's act involved a knife rather than a gun. Similarly, a weaker prior—e.g., one that involves mixed beliefs about the use of guns—is likely to be less influential.

The idea that priors about firearms are activated by and therefore influential in a case that involves the alleged use of a gun for self-defense is not surprising. However, there may also be many other priors at play in the case. For example, a juror may have strong priors relevant to a background characteristic of the alleged

---

80. *Id.*

81. *See* FED. R. EVID. 608–09.

82. *See id.*

victim or the defendant—for example, strong priors about the violent or peaceful character of people of a certain race or sex, about the benevolence of doctors, or about the conduct of people with kids or of people who drive a minivan. How heavily these priors influence a juror's judgment and what evidence activates these priors—i.e., makes them more or less influential—depends on their relative precision with respect to the matter in question.

The reflective model provides a framework for considering these questions, and it provides a coherent theory based on well-accepted principles of Bayesian inference and cognitive psychology to understand the implications for different categories of evidence. Importantly, the reflective model does not simply state that prior beliefs will have influence on a verdict. Instead, the shrinkage principle at the center of the reflective model implies that certain types of evidence will trigger certain priors such that they become active and influential with the evidence but would remain subdued or uninfluential without it. This idea has significant implications for evidence policy.

For example, in considering whether to admit character evidence under Rule 413 in a sexual assault case, a court is likely to undertake a balancing of the probative value and unfair prejudice associated with the evidence under Rule 403.[83] In considering the prejudicial effect of the evidence, it is likely to consider any particularly inflammatory aspects of the evidence as a factor against admission.[84] However, it is *unlikely* to consider the unfair prejudice associated with the evidence's activation of prior beliefs based on race, appearance, and other background characteristics of the defendant.[85] Similarly, legislatures have generally not considered these effects in deciding policy surrounding the admissibility of character evidence, impeachment evidence, hearsay evidence, or other categories of evidence.[86] The reflective model highlights the importance of considering such effects in determining evidence policy and in making admissibility decisions.

Finally, it is important to emphasize that the reflective model is well-supported by findings in cognitive psychology and juror behavior.[87] To summarize, the model is grounded in the assumption that jurors are accuracy-maximizers in that they incorporate new evidence in their beliefs in a way that optimizes the various sources of information available to them.[88] Jurors are in fact far from perfect. It is well-accepted that they, like people in general, are subject to many biases and cognitive quirks.[89] However, the primary purpose of the jury is to determine accurate judgments and correct verdicts,[90] and there is substantial evidence

---

83. *See id.* 403, 413.

84. *See* FED. R. EVID. 403 advisory committee's note to proposed rule.

85. *See Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 1025, 1029–32, 1073–78.

86. *See id.*

87. In this Section, I only provide a brief summary of the psychological findings in support of the reflective model. For a more detailed discussion, see *id.* at 1065–69.

88. *Id.* at 1065.

89. *See* STEVEN PINKER, RATIONALITY: WHAT IT IS, WHY IT SEEMS SCARCE, WHY IT MATTERS (2021); MICHAEL J. SAKS & BARBARA A. SPELLMAN, THE PSYCHOLOGICAL FOUNDATIONS OF EVIDENCE LAW (2016).

90. *See* FED. R. EVID. 102 ("These rules should be construed so as to administer every

indicating that jurors take their responsibility seriously and perform reasonably well.[91]

There is also empirical support for the argument that jurors are Bayesian thinkers and that they behave at least somewhat in line with the shrinkage principle.[92] Moreover, there is a body of literature in cognitive psychology that supports the reflective model's idea that the influence of a juror's priors depends on the priors' activation and relative precision with respect to the matter in question.[93]

In summary, for any category of evidence, the reflective model implies that prior beliefs and context play a substantial role in the inferences that are generated by the evidence. Specifically, according to the model, fact judgments are based on neither priors nor new evidence in isolation but rather depend on a combination of them as determined by the shrinkage principle. As I explain below, this conclusion leads directly to the incongruence principle.

C. Evidentiary Value

Let us define error and accuracy in terms of the difference and proximity between an actual judgment and a correct judgment.[94] For example, a jury may find that a defendant acted with intent when in fact she did not, a jury may find a civil

---

proceeding fairly, eliminate unjustifiable expense and delay, and promote the development of evidence law, to the end of ascertaining the truth and securing a just determination."); *see also* Estes v. Texas, 381 U.S. 532, 540 (1965) ("Court proceedings are held for the solemn purpose of endeavoring to ascertain the truth . . . ."); *Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 1065–66.

91. Brian H. Bornstein & Edie Greene, *Jury Decision Making: Implications for and from Psychology*, 20 CURRENT DIRECTIONS PSYCH. SCI. 63, 65–66 (2011) ("there is plenty of evidence that jurors also use careful, systematic processing strategies"; "jurors perform their duties reasonably well . . . .").

92. *See Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 1065–69 (summarizing studies); Tamara Shengelia & David Lagnado, *Are Jurors Intuitive Statisticians? Bayesian Causal Reasoning in Legal Contexts*, 11 FRONTIERS PSYCH., Feb. 5, 2021, at 1, 3–4, 9 ("Overall the results indicate that people's qualitative reasoning is mostly accurate and follows qualitative predictions of Bayesian models in predictive, diagnostic and explaining away inferences. These findings reinforce results from previous studies where Bayesian probabilistic reasoning was observed."); Hillel J. Bavli & Reagan Mozer, *The Effects of Comparable-Case Guidance on Awards for Pain and Suffering and Punitive Damages: Evidence from a Randomized Controlled Trial*, 37 YALE L. & POL'Y REV. 405, 432–50 (2019) (reporting empirical results consistent with shrinkage estimation).

93. *See Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 1065–69 (discussing empirical results in the "category activation" literature); *see also* C. Neil Macrae & Galen V. Bodenhausen, *Social Cognition: Thinking Categorically About Others*, 51 ANN. REV. PSYCH. 93, 96–100 (2000) (explaining that "[a]ccording to recent thinking on the topic, mere exposure to a stereotyped target may be insufficient to trigger category activation" and discussing the role of "perceivers' temporary processing goals and their general attitudes (i.e., prejudice level) toward the members of the category in question") (citations omitted). *See generally* Jonathan J. Koehler, *The Base Rate Fallacy Reconsidered: Descriptive, Normative, and Methodological Challenges*, 19 BEHAV. & BRAIN SCIS. 1, 5–11 (1996).

94. *See* Hillel J. Bavli, *The Logic of Comparable-Case Guidance in the Determination of Awards for Pain and Suffering and Punitive Damages*, 85 U. CIN. L. REV. 1, 12–13 (2017).

defendant not liable when in truth he should be found liable based on the law and the true facts in the case, and a jury may find that a plaintiff worked ten hours per week overtime when in fact she worked eight hours per week overtime. These examples reflect error.

Error is, in part, a function of the evidence in a case.[95] It is also a function of the prior beliefs that a juror combines with the evidence to arrive at a judgment.[96] Error, as defined above, can be deconstructed into bias and variance.[97] If we imagine repeating a judgment many times, each time with different jurors, judges, and presentations of the evidence, bias is the difference between the expected (or average) judgment and the correct judgment.[98] It reflects systematic error. On the other hand, variance is the dispersion of the judgments across different adjudications.[99] It reflects randomness, or unsystematic error. It is the inverse of precision.

In simple terms, a prior belief or an item of evidence can involve both a biasing effect (thus harming accuracy) and a precision-enhancing effect (thus improving accuracy).[100] The meaning and value of evidence, information, or prior beliefs—in both the real world and at trial—can be stated in terms of their effects on both bias and variance (i.e., the elements error).[101]

In light of this, it is convenient to define a measure of value that relates the precision-enhancing effect of evidence to its bias-introducing effect. Moreover, although the evidence-law concepts of probative value and unfair prejudice are not equivalent to the statistical ideas of a precision-enhancing effect and a bias-introducing effect (respectively),[102] these ideas are closely correlated. In particular, the probative value of evidence—often described as the tendency to make a material fact in a case more or less likely[103]—can be understood as involving a precision-enhancing effect on the judgment at hand, while unfair prejudice can be understood as involving a bias-introducing effect.[104]

---

95. *An Aggregation Theory of Character Evidence*, *supra* note 69, at 46–49.

96. *Id.* at 50–53.

97. *Id.* at 46–49.

98. *Id.*

99. *Id.*

100. *Id.* at 50–53.

101. Underlying the combination of priors and case evidence, and the shrinkage principle in particular, is an idea in statistics known as a bias-variance tradeoff. See *Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 1041–53. In Bayesian statistics, applying a prior can improve accuracy by trading a relatively small degree of bias for the purpose of gaining a relatively large degree of precision. *Id.*

102. For example, unfair prejudice can involve the introduction of confusion and therefore variability (i.e., a reduction in precision).

103. *See* FED. R. EVID. 401; FISHER, *supra* note 1, at 23–24.

104. Unfair prejudice reflects a broader range of effects than just bias as an element of error. For example, unfair prejudice can relate to the introduction of uncertainty through confusion, and some unfair prejudice arises purely out of policy concerns. *See, e.g.*, FED. R. EVID. 407–411. However, bias—in the sense of systematic error—comprises an important component of unfair prejudice. For example, a court may exclude unduly provocative evidence because it would be unfairly prejudicial in the sense of harming accuracy through the introduction of bias. Moreover, even policy considerations can arguably be incorporated

Therefore, for convenience, and in light of the close relationship between the evidence concepts of probative value and unfair prejudice on the one hand and the statistical concepts of precision-enhancement and bias-introduction on the other, I define evidentiary value as follows:

> The *evidentiary value ratio* ("ev ratio") is a measure of evidentiary value that is equal to the probative value of an item of evidence (or information) divided by the unfair prejudice associated with the evidence (or information).

For example, saying that an item of evidence has an ev ratio (sometimes referred to herein just as a "value," an "evidentiary value," or an "informational value") of two implies that the probative value of the evidence is twice as great as its unfair prejudice. An ev ratio of one-half implies that the probative value of the evidence is only half as great as its unfair prejudice.

Note two features of this definition: First, the ratio views probative value and unfair prejudice as implicitly involving expected (or mean) values—averaging over, e.g., different interpretations of the evidence by different jurors. Second, this definition involves some abuse of terminology in that I apply it in the real-world context as well as the trial context (whereas terms like "evidentiary," "probative value," and "unfair prejudice" live in the domain of trial evidence). I do this for obvious reasons—to compare the value of information in the real world to the value of the same information at trial, and specifically in terms recognized by evidence law.

Ultimately, I define the ev ratio for clarity and convenience, and my analysis herein is not sensitive to the two features discussed above or the precise relationship between evidence law and accuracy. For example, my arguments do not depend on whether the ev ratio is defined in terms of variance and bias or in terms of probative value and unfair prejudice, and they do not depend on whether I define the ev ratio in terms of expected values or another reasonable interpretation of probative value and unfair prejudice.

D. The Reflective Model's Incongruence Implication

The reflective model highlights and formalizes the central role of an observer's prior beliefs with respect to the meaning and value of evidence. This role implies that the *context* in which an observer receives information is central to the observer's interpretation of the evidence and to evidentiary value. Context therefore represents an important layer in the accuracy of a judgment.

For our purposes herein, let us divide the world into two categories of context: the trial and the real world. In reality, each of these categories involves a wide variety of contexts, but, for simplicity, I focus on substantial differences between the trial evidentiary context as a whole and the real-world evidentiary context as a whole, and I highlight individual *sub*-contexts when important.

---

in the meaning of accuracy. *See* Hillel J. Bavli, *An Objective-Chance Exception to the Rule against Character Evidence*, 74 ALA. L. REV. 121, 150–51 (2022) (discussing policy-disfavored bias).

The trial is similar to a laboratory experiment in certain respects. It is, in a sense, a sterile environment that we carefully construct to test a hypothesis (e.g., guilty or not guilty, or liable or not liable) through carefully crafted evidence and instructions that focus the observer's attention on the task at hand and that emphasize the importance of deciding the matter in question by evaluating the evidence presented. It also involves a range of implicit but important informational factors that can shape the observer's interpretation of the evidence and the observer's judgment. The real-world evidentiary context is very different. It involves a mishmash of information and a wide variety of informational and decisional tasks. It is not a sterilized environment in the sense that, even in the context of specific decisional or informational tasks—such as hiring an employee, deciding which television to buy, or forming an impression of a new neighbor—there is a fairly uncensored flow of information. In addition, in the real world, although there are exceptions, the information is not generally processed in a formal and highly-regulated analytical environment.

There are of course many differences between the trial setting and the real world. However, in the current Section, I will focus on two differences that are particularly central to an observer's interpretation of evidence. Specifically, my aim is to highlight two features of the trial setting that degrade the value of evidence relative to its value in the real world—the trial's effect on a juror's prior beliefs and the trial's analytical framework.

1. The Trial's Effect on Priors

The trial setting tends to create strong, biased, and influential priors. This occurs for a number of reasons. First, it occurs as a result of the roles of trial participants. Consider, for example, a criminal trial in which a defendant is charged with armed robbery. Although there is a theoretical presumption of innocence, the defendant begins trial with anything but a clean slate in terms of the jurors' priors. From the moment the jurors set eyes on the defendant, they view the defendant as someone who has been charged with a serious violent crime and as someone whom the police chose to arrest and whom the prosecutor decided to prosecute after evaluating the evidence.

These prior beliefs are not necessarily illogical (even if undesirable). Whether or not the defendant in fact committed the crime in question, there is generally a far higher probability that the defendant—having been arrested and charged with the crime—in fact committed the crime than a random individual. Nevertheless, even ignoring non-accuracy policy reasons to view these priors as a source of evidence devaluation, as discussed in the following Sections, they could have a damaging effect on accuracy in the context of a trial. In particular, when combined with a trial's analytical framework, they introduce bias by encouraging jurors to search for an explanation for the evidence that is consistent with the juror's priors and the

charge against the defendant.[105] In turn, they increase the potential for false-positive findings and verdicts.[106]

Second, these priors do not exist in isolation. Rather, as discussed earlier, a juror brings to trial a broad range of preexisting beliefs, stereotypes, and prejudices.[107] While an observer may attach these priors to anyone observed to have the background characteristics to which the priors pertain, there is good reason to believe that a juror's preexisting beliefs and prejudices will attach particularly readily and sharply to influence a juror's interpretation of evidence in the trial context.

Specifically, trial roles—such as the role of an individual as a defendant in a criminal trial—are likely to confirm and strengthen prior beliefs. This is based on a well-researched and well-settled phenomenon in psychology known as confirmation bias—the tendency to interpret new evidence in a way that confirms prior beliefs.[108] Thus, a juror arrives at a trial with priors based on their preexisting beliefs, stereotypes, and prejudices.[109] In a Bayesian updating process, these priors are combined with "new evidence" that is observed upon perceiving the defendant's background characteristics, hearing the nature of the trial, and being exposed to the trial context and the roles of trial participants. The juror updates her priors, attaching relevant priors to the defendant based on the defendant's background characteristics and interpreting this new information and the role of the defendant so as to confirm and strengthen her preexisting beliefs.[110] Thus, from an early stage of the trial—even before any court-sanctioned evidence has been presented—the juror already has a strong set of priors that are based on her preexisting stereotypes and prejudices and that are biased against the defendant.[111] Then, the juror hears trial evidence and combines her priors with it—interpreting it in light of the juror's priors and using it to confirm these priors—to arrive at a judgment.[112]

Moreover, the juror is in a laboratory setting in which she knows that she will be tasked, for example, with the serious matter of judging the defendant. Although there is strong evidence that people begin to form impressions about others from the moment they begin to observe them, this effect is likely particularly acute in the trial context, in which a juror is present to judge the defendant based on limited

---

105. This involves a form of the psychological phenomenon known as confirmation bias. *See infra* note 108 and accompanying text; *see also infra* Section III.D.2.

106. *See infra* Section III.D.2.

107. *See supra* notes 74–75.

108. *See* Raymond S. Nickerson, *Confirmation Bias: A Ubiquitous Phenomenon in Many Guises*, 2 REV. GEN. PSYCH. 175, 175 (1998); *see also* DAVID A. LAGNADO, EXPLAINING THE EVIDENCE: HOW THE MIND INVESTIGATES THE WORLD 172–85 (2022) (discussing confirmation bias and causal modeling).

109. *Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 1056.

110. *Id.*

111. *Id.* This assumes that the juror's preexisting beliefs disfavor the defendant; however, depending on the juror's priors and the defendant's particular background characteristics, they can also *favor* the defendant.

112. *Id.*

information and then to decide the defendant's fate based on the juror's judgment.[113]

In addition, as explained in Sections III.A–B, the introduction of certain evidence can itself activate priors that would not otherwise have a substantial influence on a verdict.[114] For example, character evidence may activate a juror's priors regarding the character of a defendant—including priors regarding the background characteristics of the defendant—even if those priors would have remained relatively subdued in the absence of such evidence.[115] However, although this effect may occur when an observer considers character evidence in the real world (e.g., information regarding a contractor's prior work to make inferences for the contractor's work on a future project, or information regarding a neighbor's prior acts to develop an impression), the effect is likely to be particularly acute when the evidence is presented at trial.

This is based on the reflective model: At trial, the jury is presented with specific evidence that is to be used (together with the jurors' common sense) for a specific inferential purpose and to fulfill a specific fact-finding task. Jurors are often instructed as such explicitly by a litigant (such as a prosecutor) and sometimes even by the court. At the same time, jurors are frequently provided with insufficient information to make such inferences without relying heavily on their priors. Thus, they may rely on their priors to fill evidentiary gaps and make inferences that are necessary to fulfill the assigned task. This is different than most real-world decisional contexts, in which tasks are often nonbinary and less well-defined or in which there are good alternatives to relying on priors—such as relying on other sources of information, declining to make inferences based on insufficient information, or incorporating greater uncertainty into an ultimate decision.[116]

Third, in evaluating evidence, the trial setting causes jurors to commit a well-settled psychological error known as the "base-rate fallacy."[117] A base rate represents the general prevalence of an event in a population.[118] This value is important for estimating the likelihood of the event in light of new evidence. For example, in medicine, if a patient tests positive for a disease, a physician would evaluate and act on this new evidence only in light of the general prevalence of the disease in the relevant population. A low general prevalence may indicate a low likelihood of the patient having the disease—notwithstanding a positive test result. The base-rate fallacy describes a cognitive bias that causes observers of evidence to undervalue the general prevalence of an event and overvalue the new evidence or

---

113. *See* Brown, *supra* note 56, at 7–14 (discussing the formation of character impressions).

114. *See supra* note 57–83 and accompanying text.

115. *See Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 1060.

116. For an example of this effect, see discussion of character evidence in Section IV.A. *See also infra* Section III.D.4.

117. *See generally* Daniel Kahneman & Amos Tversky, *On the Psychology of Prediction*, 80 PSYCH. REV. 237, 237–47 (1973) (discussing the base-rate fallacy).

118. Jonathan J. Koehler, *When Do Courts Think Base Rate Statistics Are Relevant?*, 42 JURIMETRICS 373, 374 (2002) ("The social science literature defines a base rate as a proportion—the relative frequency with which an event occurs or an attribute is present in some reference population.").

individual result.[119] A variation of this involves replacing a prior reflecting the general prevalence of an event in a population with an individuating prior reflecting the observer's beliefs regarding the event or the new evidence.[120]

The trial context tends to cause this error in judgment by creating an environment in which jurors are hyperfocused on the event at issue.[121] This is supported by research in psychology indicating that individuals commit the base-rate fallacy on the basis of the perceived relevance of the information.[122] Therefore, it is easy for a juror to lose sight of realistic prevalence estimates and instead replace them with priors that are heavily influenced by the trial context and a juror's prior beliefs and prejudices as described above. In other words, in evaluating the trial evidence, the juror relies heavily on the individuating priors that she has attached to the evidence and the defendant (or other trial participants).

Indeed, this error in judgment may have caused the false conviction in the Sally Clark case discussed in the Introduction.[123] In that case, the prosecutor, expert, and jurors, all hyperfocused on the defendant and the matter in question, may have replaced their general-prevalence estimates for the event in question—the general prevalence of a mother murdering her two young children—with priors that were likely heavily influenced by the trial setting and the role of Sally Clark as the defendant in a murder trial. Whereas the general prevalence of this extremely rare event would likely have counterbalanced the evidence indicating that Clark's two kids died by unexplained causes—also a rare event—the jurors' replacement priors (those influenced by the trial) were unable to offset the improbability of the evidence, thus resulting in Clark's conviction.[124]

We can therefore summarize this Section as follows: Due to trial roles and other features of the trial process, the strengthening of preexisting beliefs, stereotypes, and prejudices during the trial process, and the effects of various cognitive biases—including confirmation bias and the base-rate fallacy—the trial setting causes particularly strong, biased, and influential priors. These priors have the potential to cause substantial unfair prejudice by shifting the basis of a juror's judgment away from the evidence and toward the juror's subjective and biased prior beliefs.

2. The Trial's Analytical Framework

The second feature of a trial that degrades the value of evidence is the trial's peculiar analytical framework. In particular, while the trial setting unduly strengthens and biases a juror's priors, it simultaneously exacerbates these effects

---

119. *See* Maya Bar-Hillel, *The Base-Rate Fallacy in Probability Judgments*, 44 ACTA PSYCHOLOGICA 211, 211 (1980).

120. This can be viewed as equivalent to the standard base-rate fallacy, just with an added iteration of Bayesian updating.

121. *See generally* Bar-Hillel, *supra* note 119, at 215–32 (examining causes underlying the base-rate fallacy).

122. *See id.* at 216–17 (discussing causes underlying the base-rate fallacy and arguing that people neglect base rates when they view the information as irrelevant to the judgment in question).

123. *See supra* notes 8–24 and accompanying text.

124. *See id.*

by employing an analytical framework that is particularly susceptible to their biasing influence.

In many respects, the U.S. trial mirrors a scientific hypothesis test. However, it lacks important features that exist in evidence evaluation in the sciences to avoid false-positive errors that are otherwise likely to result from a researcher's prior beliefs, hopes, and expectations.

A hypothesis test is a scientific method for testing whether a certain claim or hypothesis is supported by observed evidence.[125] As in scientific hypothesis testing, in a trial, there is a "null hypothesis" (e.g., not guilty) that is rejected in favor of an "alternative hypothesis" (e.g., guilty) if there is sufficient evidence against it, which depends on a predefined threshold of proof (analogous to a level of "statistical significance" in the sciences), such as the reasonable-doubt standard in criminal cases.[126]

However, hypothesis testing entails significant risks and limitations—so much so that the American Statistical Association, "the world's largest community of statisticians,"[127] recently published an unprecedented statement criticizing the current use of hypothesis testing and establishing principles for improvement.[128] Indeed, "[t]he statistical community has been deeply concerned about issues of *reproducibility* and *replicability* of scientific conclusions,"[129] and hypothesis testing sits at the center of the "reproducibility crisis," the severe situation in the sciences and social sciences ("the sciences") wherein many empirical studies result in false-positive findings—that is positive findings that are not in fact true—such that these findings cannot be reproduced in a new study.[130] This situation, also called the "credibility crisis," has led the research community and the public to doubt empirical findings in the sciences.[131] This doubt, in turn, has led to forceful measures, such as bans on certain forms of hypothesis testing and a wide variety of requirements aimed at preventing false-positive findings and improving reproducibility.[132]

---

125. *See generally* DEGROOT & SCHERVISH, *supra* note 18, at 530–623; COPE, *supra* note 18, at 36–41; *see also Credibility in Empirical Legal Analysis*, *supra* note 22, at 507–09.

126. *See* COPE, *supra* note 18, at 40 (discussing statistical significance).

127. *About ASA*, AM. STAT. ASS'N, https://www.amstat.org/about-asa [https://perma.cc/X6FD-JN3R].

128. *See* Ronald L. Wasserstein & Nicole A. Lazar, *The ASA's Statement on p-Values: Context, Process, and Purpose*, 70 AM. STATISTICIAN 129 (2016); *see also* NAT'L ACADS. OF SCIS., ENG'G, & MED., REPRODUCIBILITY AND REPLICABILITY IN SCIENCE (2019).

129. Wasserstein & Lazar, *supra* note 128, at 129 (emphasis in original).

130. *Id.* (quoting Roger Peng, *The Reproducibility Crisis in Science: A Statistical Counterattack*, 12 SIGNIFICANCE 30 (2015)); *see also Credibility in Empirical Legal Analysis*, *supra* note 22, at 501–06 (describing the reproducibility crisis).

131. Wasserstein & Lazar, *supra* note 128, at 129.

132. *Id.* (describing a psychology journal's ban on "null hypothesis significance testing"); *see also* David Harrington et al., *New Guidelines for Statistical Reporting in the Journal*, 381 NEW ENG. J. MED. 285 (2019) (discussing new guidelines and requirements); *see, e.g.*, *New Manuscripts, Statistical Reporting Guidelines*, NEW ENG. J. MED., https://www.nejm.org/author-center/new-manuscripts [https://perma.cc/M2FQ-4N8Z] (discussing requirements for reporting statistical results).

The risks associated with hypothesis testing—and the risk of false-positive findings in particular—extend in substantial respects to the trial setting. However, whereas the scientific community has begun to implement a wide range of measures to prevent false-positive findings from hypothesis testing, courts have not implemented any such precautions to protect against the dangers of hypothesis testing in particular.[133]

The problem at the center of the credibility crisis in the sciences is an improper research practice called data fishing.[134] This practice (also called data dredging or *p*-hacking) involves searching for and selectively reporting "significant" results or distinct patterns in the data that are favorable in one respect or another to the researcher.[135] In summary, data fishing can often allow a researcher to reach any result she seeks by searching different methodologies and manipulations of the data until she obtains the desired result.[136] As Ronald Coase stated, "[I]f you torture the data enough, nature will always confess."[137]

However, the "discovered" result is invalid. To understand why this is, define a *population* as a group of objects of interest, and define a *sample* as a subgroup of those objects.[138] For example, if I am interested in knowing the average weight of a certain lizard found in the Amazon—the population of interest—I may collect a sample from that population and infer the average population weight from the average weight found in the sample. Similarly, I may use sampling to infer a causal effect of a certain drug in a population of interest.

Thus, statistical inference enables a researcher to learn about a population by studying patterns in a sample from the population.[139] Data fishing, then, leads to invalid results because the probability-based hypothesis test relies on the idea that an observation in a sample is significant if, assuming the null hypothesis is true, the observation is sufficiently extreme, thus providing support for the rejection of the null hypothesis in favor of the alternative hypothesis.[140] That is, trying out many testing models and methodologies enables the researcher to find extreme results by probability alone (by definition) rather than as a true signal regarding the population. In other words, the researcher interprets a nuance of the sample as a

---

133. *See generally Credibility in Empirical Legal Analysis*, *supra* note 22, at 549–53.

134. *See id.* at 502 (explaining data fishing).

135. *See* Joseph P. Simmons, Leif D. Nelson & Uri Simonsohn, *False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant*, 22 PSYCH. SCI. 1359, 1359 (2011); *see also Credibility in Empirical Legal Analysis*, *supra* note 22, at 502.

136. *Credibility in Empirical Legal Analysis*, *supra* note 22, at 509–14 (demonstrating how a researcher may search for and obtain a desired result).

137. R.H. COASE, HOW SHOULD ECONOMISTS CHOOSE? 16 (1982), *reprinted in* R.H. COASE, ESSAYS ON ECONOMICS AND ECONOMISTS 15, 27 (1994).

138. *See generally* DEGROOT & SCHERVISH, *supra* note 18, at 376–94; COPE, *supra* note 18, at 27–31.

139. *See generally* DEGROOT & SCHERVISH, *supra* note 18, at 376–94; COPE, *supra* note 18, at 27–31.

140. *See Credibility in Empirical Legal Analysis*, *supra* note 22, at 509–14. For example, I may reject a null hypothesis that lizards of a certain type weigh an average of half a pound if the average in a relatively large sample is three pounds.

feature of the population; and the extreme result would not arise in a new study that involved either a new sample of data or a new methodology.[141]

Now, in the trial setting, the way in which parties manipulate and present evidence is similar to data fishing in certain respects and distinct in other respects. First, a key component of data fishing involves not only searching for a desired result but also selectively reporting it—that is, reporting the methodology that led to the desired result without reporting the numerous other methodologies tested.[142] In the trial setting, the selective-reporting component of data fishing does not generally apply. This is because parties do not present themselves as neutral researchers; rather, it is known that they search the data—i.e., the available evidence—for patterns that support their argument.

Instead, the search for evidentiary support in litigation is more similar to the data fishing analog—called "overfitting"—in the fields of machine learning and statistical modeling.[143] Overfitting can be described as follows:

> [O]ur standard goal in statistical modeling is to develop a model that can capably generalize to new observations similar, but not identical, to the ones we have sampled. We generally do not care very much about how well we can predict scores for the observations in our existing sample, since we already *know* what those scores are. In this sense, the prediction error that we compute when we fit a model on a particular dataset is only a proxy for the quantity we truly care about, which is the error term that we would obtain if we were to apply our trained model to an entirely new set of observations sampled from the same population.[144]

Thus, overfitting can be described as an error that occurs when a model overexplains the evidence in a sample in the sense that it interprets nuances or patterns in the sample as true features of the population.[145] As a result, when the model is applied to future samples (e.g., when a drug that was found effective in an initial sample of individuals is administered to a "new" individual), it will not perform well to predict its patterns. This is to say that a model that best explains a sample of evidence—for example, past behavior, past temperatures, or past test results—is not necessarily one that will best predict a future sample.[146]

For example, in Figure 1, the curved line perfectly fits the sample data. However, this model of the population, based on the sample, likely overfits the data—it interprets randomness (or "noise") in the sample data as a feature of the population. Consequently, if applied to a different sample from the population, it is

---

141. *Id.*

142. *Id.*

143. *See* Yarkoni & Westfall, *supra* note 20, at 1102 ("The tendency for statistical models to mistakenly fit sample-specific noise as if it were signal is commonly referred to as *overfitting*.") (emphasis in original); *see also Credibility in Empirical Legal Analysis*, *supra* note 22, at 516–19 (discussing overfitting).

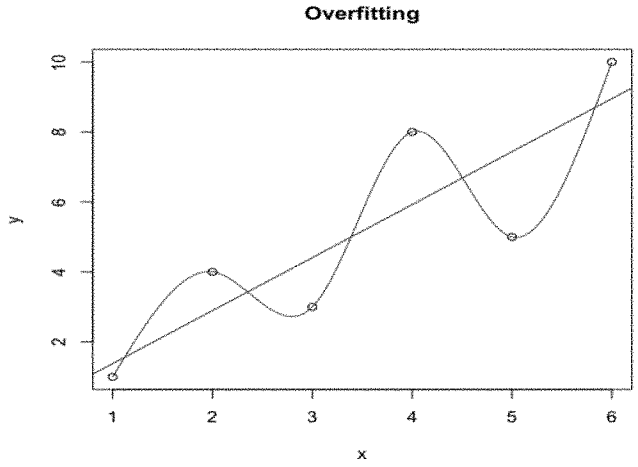
144. Yarkoni & Westfall, *supra* note 20, at 1102 (emphasis in original).

145. *See id.*

146. *Id.* at 1100–01.

likely to perform worse as a predictor of the data than a simpler model, such as the straight line in the figure.

Figure 1:
Comparison between a good model (straight line) and an overfit model (curved line).



Proof at trial is similarly susceptible to error resulting from overfitting. Evidence can be understood as a sample from which we (via a trial) want to infer the truth. In other words, we can view a state of perfect information about the facts of a case as a population of values from which we can simply compute the truth about a case and the correct judgment. However, we do not have perfect information; rather, we have limited evidence from which we may impute facts and infer a judgment regarding the true facts.

Like hypothesis testing in the sciences, we begin with a null hypothesis of, for example, "not guilty." We then observe the evidence. As in other hypothesis testing contexts, we measure the extremeness of the evidence against some predetermined threshold, and we reject the null hypothesis if the evidence is sufficiently extreme.[147]

For example, if from a certain sample of evidence, the jury imputes the fact that the defendant was in the vicinity of a crime at the time that it occurred, the evidence that led to this conclusion could be due to randomness—that is, he was randomly in the vicinity even though he did not in fact commit the act in question—or to the incorrectness of the null hypothesis of "not guilty." This particular evidence is likely not sufficiently extreme to satisfy a reasonable-doubt standard. However, if the jury also finds that the defendant had a motive to commit the crime and had recently purchased a weapon that matches the weapon used in the crime, it may conclude that the evidence is too extreme to be random under the

147. *See generally* DEGROOT & SCHERVISH, *supra* note 18, at 530–623 (discussing hypothesis testing).

null hypothesis. It may therefore reject the null of "not guilty" in favor of an alternative hypothesis of "guilty."

In the trial setting, overfitting can be said to occur when a verdict depends on an evaluation of the evidence that interprets randomness associated with the sample of evidence as a feature of the population of correct facts and inferences—that is, as a true inferential signal for the population of true facts, or more specifically, for the judgment at hand. Let us consider this more carefully. Overfitting describes a modeling problem: The model (or story) accounting for the evidence misinterprets sampling noise as a feature of the population. But at its core, it is an inferential problem. The overfit model is problematic because it is not predictive of other samples from the population—it does not allow accurate inferences for predicting future samples. At trial, overfitting describes a similar problem. It involves overexplaining the evidence such that the observer derives meaning from mere noise and therefore makes inferences that are in fact not predictive of the judgment in question.

In other words, in the trial setting, a model's predictive value can be understood as its value for predicting, or drawing inferences about, the outcome or other judgment at hand. A researcher observes sample data and models it for inference for future samples—in a sense, develops a story or accounting of it with respect to the broader population. Similarly, a juror in a trial observes case evidence and models it—draws inferences to develop a story around it—to infer the judgment in question. Just as an overfit model in science has poor predictive value for future samples, an overfit model at trial—that is, an accounting of the evidence that misinterprets randomness as true inferential signal—has poor predictive value for the judgment.

Now, here's the tricky part: Whether at trial or in a scientific study, overfitting connects the dots (i.e., the data or evidence) too closely.[148] But this does *not* mean that the dots have no probative value (i.e., predictive or inferential value) for the judgment at hand. Overfitting incorrectly interprets *noise* as having inferential value—even if any given dot indeed has such value. Consider the overfit (curved) model in Figure 1. Each data point on the curve has inferential value. It is the way in which they are all connected, and the inferences that are drawn from this accounting, that is problematic.

This represents a difficult task for both jurors and scientists. It takes seeing the forest through the trees: the problem is with the story, the methodology, and the modeling rather than with any individual data point. This is part of the reason that data fishing is so difficult to address in the sciences: the error, although potentially severe, is far from obvious, and many researchers do not realize that their research methodology renders the results invalid. At trial, the problem is arguably even more difficult, especially when certain evidence seems compelling. However, the problem of overfitting is crucial because it ultimately causes false-positive findings, or in the trial setting, false convictions and other false-positive judgments.

Finally, it is important to highlight that although various analytical contexts are subject to the dangers of data fishing and overfitting, the analytical framework of the trial—essentially a hypothesis test—is particularly vulnerable to these dangers.

---

148. *See supra* Figure 1.

This is due to the unique features of a hypothesis test, which involves an explicit or implicit evaluation of the extremeness of the evidence, a comparison between the extremeness of the evidence and a predetermined proof threshold, and a binary decision to reject or not reject a null hypothesis in favor of an alternative hypothesis. An analysis of real-world reasoning and decision-making is beyond the scope of the current article. Suffice it to say that it generally does not involve hypothesis testing or its unique features.[149] To be sure, in limited circumstances, a particular real-world analytical context may similarly mirror a hypothesis test in important respects. In these exceptional circumstances, the conclusions herein may apply.

In short, in the absence of special precautions, the features of a trial's analytical framework facilitate overfitting and ultimately false-positive judgments by permitting and encouraging an observer to search for and overemphasize extreme features, to see patterns in randomness, and to reject or not reject a hypothesis based on whatever result the observer is inclined to find. As I explain in the following Section, these effects, in turn, promote judgments that are heavily influenced by a juror's priors.

3. Strong Priors and a Vulnerable Analytical Framework: A Recipe for Error

The previous Sections highlight two features of the trial setting that give rise to a high risk that the value of evidence at trial will be substantially degraded relative to its value in the real world. In this Section, I emphasize the compounding effects of these two features—strong and influential priors and a vulnerable analytical framework—in degrading the value of evidence and causing false-positive findings.

While a trial's analytical framework is vulnerable to overfitting, strong and influential priors bring this danger to fruition. As I mentioned above, it is common for data fishing to occur without a researcher's awareness that their methodology invalidates the study's results. In the sciences, a researcher may frequently have an expectation—that is, a prior belief—that will lead her to overfit a model to confirm a result in line with her expectation (or desired result).[150] After all, most research is motivated by some expected result, such as an expected causal effect of a drug or procedure.

In this sense, prior beliefs are at the root of data fishing and the credibility crisis. Researchers are not objective observers. Whether grounded in profit, politics, or getting ahead in one's profession, motivational biases play a substantial and detrimental role in empirical research, and their vehicle for causing error is data fishing.[151]

---

149. *See generally* LAGNADO, *supra* note 108 (discussing how people analyze evidence in the real world).

150. *See supra* Sections III.A–B, D. Similarly, a researcher may overtly search for a desired result that derives from or reflects her prior beliefs.

151. *See Credibility in Empirical Legal Analysis*, *supra* note 22, at 503 (discussing motivational bias and data fishing); *see also* Gilberto Montibeller & Detlof von Winterfeldt, *Cognitive and Motivational Biases in Decision and Risk Analysis*, 35 RISK ANALYSIS 1230, 1230 (2015) (explaining motivational biases as "conscious or subconscious distortions of

As in the sciences, an observer's priors in the courtroom can have a substantial impact on the outcome of a trial. Jurors are generally not motivated by profit, politics, or getting ahead in their profession or performance as a factfinder.[152] However, as explained in Section III.D.1, by the time jurors begin to hear formal evidence in a case, they are likely to already have strong priors regarding the litigation parties, the case, and the evidence.

In turn, strong priors cause overfitting via confirmation bias—the tendency to seek out and interpret evidence in a way that confirms the observer's prior beliefs—and, more generally, by combining with new evidence pursuant to the shrinkage principle, such that a juror's subjective priors may dominate her judgment in the analytical framework of a trial.[153]

As psychologist David Lagnado has emphasized, "it's not surprising that our hypothesis testing is . . . oriented towards testing for positive effects. We construct a model of what we believe happened, and look for positive effects of this model. We find it less natural to search for . . . effects that our model does not predict."[154] Moreover, "[a]s well as preferring questions that are more likely to yield a positive answer, people also prefer tests with more extreme probabilities of giving a positive answer."[155] Lagnado further highlights the dangerous combination of a "systematic bias in evaluation" and "an asymmetric testing strategy":

> Thus, if one adopts a positive test strategy, and overestimates the diagnostic value of positive test outcomes, then a distinct bias towards the target hypothesis will follow. This seems a common error in crime investigations . . . . It also occurs in studies of hypothesis-testing, such as tasks requiring people to identify unknown objects based on shared features; in such tasks people mainly conduct positive tests and often overestimate the value of positive outcomes.[156]

In his recent book, Lagnado emphasizes that people are relatively good at using data to devise hypotheses, stories, and explanations for evidence but are relatively poor at evaluating models and hypotheses with respect to the evidence.[157] In other words, while "[w]e have a gift for creating novel hypotheses," we have great

---

judgments and decisions because of self-interest, social pressures, or organizational context"). Motivational biases also play a substantial role in our reasoning in the real world. *See* SCHAUER, *supra* note 47, at 226–38 ("Understanding motivated reasoning is crucial for understanding when and how evidence matters in the world, and when and how it does not.").

152. *See supra* notes 88–91 and accompanying text.

153. *See supra* Sections III.A–B.

154. LAGNADO, *supra* note 108, at 178 (discussing confirmation bias and causal modeling); *see also id.* at 176–77 (citing studies).

155. *Id.* at 178.

156. *Id.* at 180 (citations omitted).

157. *Id.* at 11–12 ("[W]hile we are adept at generating models to explain evidence, we are less capable at evaluating these models against the evidence. We can reason *from* evidence, but not so readily *about* the evidence or about the relation between our model and the evidence.") (emphasis in original).

difficulty "evaluating how well [evidence] supports our hypotheses as compared to other alternatives, and assessing the quality and reliability of the evidence itself."[158]

The analytical framework of a trial, rooted in a methodology of classical hypothesis testing, is precisely the kind of analysis that people find so difficult, and many of the deficiencies in reasoning that make this type of analysis so difficult are closely related to the influence of an observer's prior beliefs.[159] "Many shortcomings [in a person's ability to evaluate evidence] have been catalogued: people focus on a single story and neglect alternatives; they search for evidence to support their hypotheses and downplay contrary evidence; they fail to scrutinize their sources of evidence; they misinterpret statistical argument."[160]

In summary, the trial setting[161] at once promotes biased and influential priors and an analytical framework that is particularly vulnerable to the biasing effect of these priors. Jurors are poor at evaluating hypotheses against the evidence. They create stories that confirm their priors, modeling the evidence such that confirming inferences are overemphasized and disconfirming inferences are explained away or underemphasized. This is supported by empirical findings in the fields of cognitive psychology and juror behavior, and it has a strong theoretical grounding in the field of Bayesian inference.

The inferential value of evidence at trial is based not only on the evidence itself but on the priors of the observer. The trial setting promotes strong, biased, and influential priors—priors based on race, sex, appearance, education, and a broad range of other background characteristics, and of course, priors based on the trial context itself. By producing such priors and applying an analytical context that encourages a heavy reliance on them, the trial setting can have a substantial degrading effect on the value of evidence relative to its value in the real world. This results in biased judgments, unequal treatment, and false convictions and other false-positive judgments.

In the following Section, I apply the analysis above to discern principles for determining the extent to which evidence will be susceptible to a trial's devaluing effect.

4. The Influence of Prior Beliefs

The analysis above implies that the trial setting degrades the value of evidence relative to its value in the real world. This devaluing effect is grounded in the unique features of a trial, which alter a juror's priors and cause them to have undue influence on a case's outcome. In other words, the influence of a juror's prior beliefs is central to the devaluation of evidence at a trial.

However, the influence of a juror's prior beliefs is premised on their relevance—and more specifically, pursuant to the shrinkage principle, how precisely they speak to the matter in question.[162] As such, while a juror's biases can impact a trial in many ways, the harmful effects discussed above are most

158. *Id.* at 12.
159. *See id.* at 172–83.
160. *Id.* at 12–13.
161. *See supra* notes 17–22 and accompanying text; *see also supra* Section III.D.1.
162. *See supra* Sections III.A–B.

concerning (i.e., they are most harmful to the aims of accurate factfinding) when, pursuant to Bayesian updating and the shrinkage principle, the juror's priors have a significant role in the juror's judgment relative to the role of case evidence (or "new evidence") regarding the matter in question.[163]

According to the shrinkage principle, when a juror combines two sources of evidence—priors and new evidence—she will weight her priors in proportion to their relative precision.[164] In particular, a prior will be most influential when it speaks precisely to the matter in question and when the new evidence is relatively imprecise with respect to the matter in question. Therefore, the shrinkage principle and the analysis above imply the following three "evidence devaluation principles" regarding the trial setting's tendency to degrade the value of evidence via the influence of a juror's priors:

> (1) Imprecision of evidence. The devaluation of evidence will be most severe when the reasoning that connects the evidence to the matter in question leaves substantial gaps in its chain of inferences, reflecting the evidence's imprecision with respect to the matter in question. This imprecision and the ensuing inferential gaps cause jurors to rely more heavily on their priors in deciding the matter in question. Conversely, when evidence provides a precise informational signal with respect to the matter in question, jurors will rely less heavily on their priors.
>
> (2) Precision of priors. The devaluation of evidence will be most severe when the reasoning that connects the evidence to the matter in question involves inferences for which a juror's priors are particularly precise, thus creating an intermediate judgment that invokes a juror's heavy reliance on her priors. Conversely, when the reasoning associated with evidence does not involve inferences for which the juror's priors are precise, the juror will rely less heavily on her priors.
>
> (3) Inferential flexibility. The devaluation of evidence will be most severe when the reasoning that connects the evidence to the matter in question involves a particularly complex chain of inferences and a high degree of inferential flexibility, thus giving rise to a higher risk of overfitting (or over-explaining the evidence in a way that misinterprets randomness as a predictive informational signal) based on prior beliefs.

These devaluation principles determine the extent of the trial setting's devaluing effect on evidence relative to its value in the real world. As discussed earlier, for purposes of this paper, I define evidentiary value as a ratio of probative value and unfair prejudice.[165] Therefore, the devaluation of evidence can take the form of a decrease in probative value, an increase in unfair prejudice, or a combination of the two. The influence of a prior in the Bayesian framework frequently comes with an

---

163. The discussion in Section III.D.1 suggests that the trial setting may generally magnify the strength and influence of priors relative to those in real-world contexts. This Section, on the other hand, addresses the influence of priors in different trial contexts.

164. *See supra* Sections III.A–B.

165. *See supra* Section III.C.

effect on both bias and variance, reflected in both components of the ratio. While the influence of priors in the real world frequently leads to an improvement in accuracy and an increase in evidentiary value, this is often not the case at trial. Due to the two features discussed above—the trial's effect on an observer's priors and the trial's unique analytical framework—the inferences arising from certain categories of evidence are unduly influenced by the observer's priors.

Importantly, this effect does not in any way override the significant role of the shrinkage principle in determining the influence of prior beliefs. Rather, the above features act through shrinkage to cause the devaluation of evidence. For example, the trial's strengthening of a juror's priors impacts the juror's judgment by strengthening their potential precision and thereby increasing their potential influence via the shrinkage principle. But pursuant to shrinkage, the influence of the priors—however strong—still require a judgment, or a "temporary processing goal[]," to which they are relevant.[166] Moreover, although a prior's strength impacts its potential precision, its precision depends substantially on the uniformity of the evidence with respect to the judgment in question—that is, how closely it speaks to the matter at issue.

Thus, in light of the devaluation principles, the trial setting can be expected to degrade the value of different categories of evidence differently. In the following Part, I apply these principles to examine their implications for various evidentiary contexts.

IV. IMPLICATIONS

In this Part, I analyze the implications of the incongruence principle for a number of evidentiary contexts—and particularly, for character evidence (IV.A), witness-credibility evidence (IV.B), the perceptions of judges and witnesses (IV.C), and hearsay evidence (IV.D). Then, in Section IV.E, I discuss implications for the law's aim of equal treatment for individuals of different races, genders, appearances, economic and educational circumstances, and other background characteristics.

A. Character Evidence

Character evidence causes a perfect storm for the devaluation of evidence in the trial setting. This is due to two unique features of character evidence: it involves a character judgment, and it provides only vague information to enable such a judgment.

In *Character Evidence as a Conduit for Implicit Bias*, I apply the shrinkage principle to argue that the admission of character evidence causes jurors to rely heavily on their prior beliefs regarding the race, sex, appearance, and other background characteristics of a defendant in determining a verdict.[167] In particular, I argue that a juror's priors regarding the background characteristics of the defendant generally remain relatively subdued when the judgment in question is

166. Macrae & Bodenhausen, *supra* note 93, at 99.

167. *See Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 1055–65.

not character related. This is because the priors reflect broad and imprecise generalizations relative to the event evidence (or lack thereof) in the case—for example, eyewitness testimony, the defendant's prior statements to the police, or fingerprint evidence.[168]

However, the immediate factual matter in question changes when a court admits character evidence—as does the relevance of a juror's priors. Character evidence involves a judgment regarding the defendant's character and then inferences from this character judgment for the ultimate fact judgment in question.[169] A juror's priors regarding the background characteristics of a defendant—priors that pertain specifically to the defendant's character—become very precise when a court admits character evidence and explicitly or implicitly sanctions a character judgment and inferences from it for the event in question.[170] Moreover, the character evidence itself usually provides relatively imprecise information regarding the defendant's character—for example, with respect to the consistency of the defendant's character, or how the defendant's acts in other contexts would extend to the circumstances at issue.[171] These evidentiary gaps are arguably filled with the defendant's priors, which speak precisely to the immediate matter in question—the defendant's character.[172]

The analysis in Part III carries important implications for these effects. Specifically, it implies that the trial setting itself plays an important role in the effects of character evidence discussed immediately above. This, in turn, explains why character evidence can be so detrimental to the aims of evidence law even while being so probative in the real world.

To understand why this is, consider a juror's priors at trial relative to an observer's priors in the real world. A juror's priors at trial are substantially more precise with respect to character reasoning than are an observer's priors for character reasoning in the real world. This is due to a number of factors: First, as discussed above, due to trial roles and other features of the trial process, a juror's priors are substantially strengthened by the trial setting itself.[173] Therefore, they are stronger—and more precise—at trial than in the real world. Second, unlike the real world or in trials in which no character evidence is introduced, when a court admits character evidence, the immediate matter in question becomes the defendant's character rather than an underlying noncharacter fact question. Therefore, while priors regarding character are likely always to have some effect, pursuant to the shrinkage principle, they are likely to be particularly uniform—and therefore particularly precise—with respect to the matter in question when character

168. *See id.*

169. *See generally* FED. R. EVID. 404(a)(1), (b)(1). For example, character evidence showing that the defendant has committed two prior assaults of a similar nature as the assault in question is probative through the following chain of inferences: the defendant has committed assaults in the past; therefore, the defendant has a character or propensity for violent acts of this type; therefore, he is more likely to have committed the act in question.

170. *Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 1058–62.

171. *See id.* at 1062–63.

172. *See id.*

173. *See supra* Section III.D.

evidence is admitted at trial.[174] Third (and relatedly), a juror's priors regarding character are particularly salient when a prosecutor or court explicitly invites the juror to engage in character reasoning.[175] And fourth, a juror's priors fulfill a substantial gap-filling role when at trial.[176] Specifically, contrary to the real world, a juror at a trial in which the court admits character evidence is specifically asked to make character inferences to arrive at a judgment. At the same time, however, the juror is provided insufficient information to make such inferences—usually just a small and biased sample of prior acts. Therefore, she is likely to rely heavily on her priors in arriving at a character assessment and in determining the underlying fact judgment.[177]

Moreover, in addition to increasing the influence of priors, the trial setting's unique analytical framework exacerbates this effect by augmenting the trial's vulnerability to their biasing effects—and to false-positive judgments in particular—through overfitting. As discussed, the trial is distinct from most real-world analytical settings. It involves a hypothesis test of sorts—one in which a juror considers the extremeness of the evidence and then rejects or does not reject the null hypothesis of, for example, not guilty in favor of the alternative hypothesis of guilty if the extremeness of the evidence exceeds the prescribed threshold.[178]

Overfitting is particularly concerning in this context because, due to confirmation bias, jurors may search for and attach undue importance to evidence that supports their priors. In other words, by seeking out patterns and extremes in a setting in which jurors receive information regarding *other* events that are not relevant to the matter in question except through vague yet influential character inferences, there is a substantial risk of interpreting noise or random patterns as signaling extreme evidence under the null hypothesis of not guilty.[179] In turn, in light of the binary decision before the jury, this interpretation can prompt the jurors' rejection of the null hypothesis of not guilty in favor of a guilty verdict.

In this context, an important factor determining the susceptibility of the analysis to overfitting is the degree of inferential flexibility (and complexity) associated with the evidence.[180] As discussed above, strong priors promote the role of confirmation bias in factfinding.[181] In turn, inferential flexibility facilitates the effects of a juror's priors and the role of confirmation bias by leaving more of the factfinding to the juror's intuition or "common sense." For example, compare character evidence to eyewitness testimony. Assuming that an eyewitness is credible, the witness's testimony that she saw the defendant approach the alleged victim at the time and place of the alleged crime leaves relatively little to common

---

174. *See supra* Sections III.A–B.

175. *See supra* notes 92–93 and accompanying text; *see also* Macrae & Bodenhausen, *supra* note 93, at 96–100 (discussing "category activation").

176. *See supra* Sections III.A–B.

177. *See supra* Sections III.A–B.

178. *See supra* Section III.D.

179. *See supra* Sections III.D.2–3.

180. Inferential flexibility is similar to methodological degrees of flexibility in research, or "researcher degrees of freedom"—a primary driver of overfitting in the sciences. Simmons et al., *supra* note 135, at 1359.

181. *See supra* Section III.D.

sense—it leaves little to inferential flexibility. On the other hand, even assuming that other-acts character evidence (e.g., two prior assaults of a similar nature to the act in question) is perfectly credible in the sense that the juror can be certain that the defendant committed the alleged prior bad acts, there is still a very substantial inferential gap—that is, flexibility that is left for the juror to sort out through "common sense." One juror may interpret the other-acts character evidence as leading strongly to the inference that the defendant committed the act in question while another juror may place little weight on the evidence, believing instead that people change significantly over time and across distinct situations.[182]

Tied to the inferential flexibility associated with character evidence is the complexity of the inferences connecting the evidence to the case. It's not like other forms of evidence: it requires a key probabilistic assumption that an individual's prior acts and current acts are all bound together by a certain force that permits an inference from those acts for the matter in question.[183] This inferential chain first requires a set of inferences connecting the alleged prior acts to the central force—the defendant's character—and then a second set of inferences connecting the first set of inferences—and the juror's character judgment in particular—to the act in question. This complexity gives rise to substantial inferential flexibility, and the complexity of the ultimate inferential chain signals a mode of reasoning similar to that illustrated in Figure 1—a mode of reasoning susceptible to overfitting.

In summary, character evidence involves a perfect storm for false-positive judgments at trial even while providing valuable information in the real world. At trial in particular, it invokes strong, undesirable priors that can be expected to have a substantial influence on a juror's interpretation of the evidence. Moreover, it requires complex reasoning and involves a particularly high degree of inferential flexibility—both of which, in turn, make the trial's already vulnerable analytical framework highly susceptible to overfitting and false-positive judgments.

In light of this, courts and legislatures should reassess the many exceptions to the rule against character evidence, exceptions that are generally premised, implicitly or explicitly, on the incorrect assumption that the value of the evidence at trial is similar or equivalent to the value of the same evidence in the real world.[184] Moreover, they should examine the current trend in the courts toward a more permissive treatment of character evidence, and more broadly, they should create law and policy that reflect the trial setting's dramatic devaluing effect on character evidence.

B. Impeachment Evidence and Witness Credibility

Character evidence appears at trial in various forms. One common and important form is credibility evidence. Under Federal Rules of Evidence 608 and 609, a party may offer evidence of a *witness's* prior acts to impeach the witness by proving that the witness has a character for untruthfulness and therefore should not

182. *Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 1039.

183. *See generally* FISHER, *supra* note 1, at 153–60 (discussing the "propensity box").

184. *See supra* Part II.

be trusted as a witness.[185] Rule 608 pertains to prior acts generally while Rule 609 pertains to prior convictions—including convictions for crimes unrelated to truthfulness (on the theory that prior serious crimes create an inference of untrustworthiness and untruthfulness).[186]

The implications for character evidence discussed in Section IV.A above generally apply to character-based credibility evidence also. The consequences of such effects are in some instances less severe since, in this context, the juror's priors affect the juror's perception of the witness's credibility rather than, for example, whether a defendant committed the act in question directly. However, the effects can nevertheless be as severe as those described in Section IV.A. For example, a witness's testimony may be critical to the outcome of a case, and a juror's judgment of the witness's credibility may be central to the juror's interpretation of the evidence. Moreover, in some instances, a defendant may also be a witness, in which case, credibility evidence may affect the case's outcome via both credibility inferences and character inferences about the defendant more broadly.

C. The Perception of Judges and Witnesses

Before discussing implications for hearsay, it is important to emphasize that trial participants other than jurors also have priors that may impact the presentation and interpretation of evidence and ultimately the factfinder's judgment. As with a juror's priors, the trial setting can have a substantial impact on a verdict via these priors.[187]

For example, a witness's priors may impact how a witness testifies; a judge's priors may impact how a judge rules on evidentiary and procedural issues; and an attorney's priors may impact how the attorney presents a client's argument. For instance, just as a trial may degrade the value of evidence by affecting a juror's priors regarding the evidence, it may have a similar devaluing effect by impacting a witness's priors and thereby influencing the way in which the witness testifies. This, in turn, can often compound the trial's impact on a *juror's* priors.

D. Hearsay Evidence

Hearsay is an out-of-court statement offered for the truth of the matter asserted in the statement.[188] For example, assume that to prove that a defendant committed the act in question, a prosecutor seeks to introduce testimony by Andrea that her friend, Joe, told her that he saw the defendant commit the act in question. Andrea's testimony constitutes hearsay because it is about Joe's out-of-court statement, and this statement is offered to prove the truth of the matter asserted in the statement—that the defendant committed the act in question. The rules of evidence generally prohibit hearsay, but with many exceptions.[189]

---

185. FED. R. EVID. 608–09.
186. *Id.*
187. *See generally supra* Sections III.A–B, D.
188. FED. R. EVID. 801(a)–(c).
189. *See id.* 802–04.

The implications of the incongruence principle for hearsay evidence are more complex than its implications for character evidence. In some circumstances they may strongly disfavor the admissibility of hearsay evidence, while in other circumstances, the implications for hearsay are arguably less severe than for character evidence. In still other circumstances, they may even *favor* the admissibility of hearsay evidence.

First, hearsay evidence generally does not invoke a juror's priors as strongly as does character evidence. This is because hearsay generally relates directly to the factual matter in question rather than the defendant's character in particular, and it generally does not involve the features above that cause unusually precise priors for the matter in question.[190]

Second, however, the implications for hearsay are more complex because hearsay evidence uniquely involves at least two effects that work in opposing directions—reflecting both the juror's priors and the trial witness's priors. Unlike other categories of evidence, hearsay involves testimony regarding an out-of-court statement offered for the substance of that statement.[191] In the previous Section, I highlighted the implications of the incongruence principle for trial participants other than jurors. One important category of implications pertains to witnesses. In particular, when a witness participates in a trial, any statement that they make at trial regarding the underlying event at issue is impacted by the trial setting. After all, the witness's perception is influenced by their priors just as a juror's perception is. Of course, some testimony may be impervious to the influence of priors, but other testimony may not be. The witness's perception of the underlying events may well be subject to the shrinkage principle and the influence of the witness's priors.

However, hearsay evidence alleviates this concern to some degree. This is because the testimony presents a statement that was made out of court—*prior to the devaluing impact of the trial*. That is, whereas most testimonial contexts involve a trial witness and an event witness as one in the same, the hearsay context involves an event witness—called a "declarant"[192]—and a trial witness, who testifies regarding the declarant's statement.[193] To be sure, a trial witness's perception of a declarant's statement may be influenced by the witness's priors. But there is arguably less inferential flexibility in perceiving and recounting a statement. On the other hand, the actual statement of the individual who observed the event in question—the declarant—is unmarred by the influence of the trial.

This is not to say that hearsay evidence is necessarily reliable. Certainly not. The general rule against hearsay has sound foundations. However, just as many categories of statements involve indicia of reliability and have therefore given rise to exceptions to the general rule,[194] the incongruence principle supplies a factor favoring out-of-court statements.

Consider for example, the exclusion from the rule against hearsay under Rule 801(d)(1)(C), which permits testimony by a witness regarding her own prior out-of-court statements when the statements "identif[y] a person as someone the declarant

190. *See generally id.* 801.
191. *See id.*
192. *See id.* 801(b).
193. *See id.* 801(c).
194. *See generally id.* 803–04.

[and witness] perceived earlier."[195] A primary rationale for this rule is that the out-of-court identification may be more reliable than an in-court identification due to the unique nature of identification evidence, which may be more susceptible to unreliability arising from the passage of time before trial, potential threats against the witness, and other influences.[196]

The incongruence principle carries similar implications for hearsay evidence more broadly. The trial itself would have a devaluing effect on a statement regarding an underlying event. Therefore, an out-of-court statement untouched by the trial setting has certain advantages over a statement made at trial.

Of course, we must also consider effects of the trial setting that work to devalue hearsay evidence. First, the trial has a devaluing effect on hearsay evidence based on the juror's priors. For example, like character evidence, hearsay involves certain gaps that are vulnerable to the influence of a juror's priors. For instance, the juror often lacks information regarding the credibility of the declarant—the individual who made the out-of-court statement—or other circumstances surrounding the out-of-court statement.[197] These gaps may be filled by the juror's priors regarding the defendant, the witness, or the evidence.

Second, we must consider the effect of the trial on the trial witness. The trial witness observes the declarant's statement unaffected by the trial but is then exposed to the trial setting before and during her presentation of the evidence to the jury. Testimony that recounts an earlier statement may be less susceptible to the influence of the witness's priors than testimony about an underlying event; nevertheless, the trial setting may alter the witness's memory and narration of the earlier statement.

Finally, and closely related to the points above, the trial setting creates a unique danger for hearsay evidence in that the prior beliefs of a real-world observer interact indirectly—through a trial witness—with the prior beliefs of jurors, observers at trial. For example, at trial, a juror may observe a statement of a declarant through the testimony of a witness. Pursuant to the shrinkage principle, the juror combines her prior beliefs with the evidence to make inferences and judgments.[198] It is possible, for instance, that the juror will attribute excessive precision and therefore weight to the out-of-court statement (rather than, e.g., discounting the value of the out-of-court statement to reflect its relative unreliability)—more than would be intended by the declarant. This may occur as a consequence of the relatively hollow and unidimensional nature of testimony regarding statements (as opposed to events), or perhaps even to some form of metacognition regarding the unimpaired nature of the statement.

As indicated above, the effect of the trial setting on the value of hearsay evidence is complex. An in-depth analysis of the normative implications of the incongruence principle for hearsay evidence is beyond the scope of this article. Suffice it to say that it has significant implications for the reliability and meaning

---

195. *Id.* 801(d)(1)(C).

196. *See supra* note 46; *see also* Commonwealth v. Weichell, 453 N.E.2d 1038, 1043–44 (Mass. 1983); *see generally* FISHER, *supra* note 1, at 510–74 (discussing rationales underlying exceptions under Rule 803).

197. *See* FISHER, *supra* note 1, at 376–82 (discussing reliability concerns).

198. *See supra* Sections III.A–B.

of the evidence in light of the trial setting's impact on the prior beliefs of jurors and witnesses—an important topic for future analysis.

E. Implications for Equality

Finally, it is important to emphasize that, aside from the incongruence principle's implications for accuracy and policies of procedural fairness and preventing false convictions, it has significant implications for the aims of equal treatment and preventing the influence of implicit and explicit prejudices based on the race, sex, religion, appearance, economic status, education, accent, and other background characteristics of litigation parties and witnesses.

The centerpiece of the incongruence principle is the heightened influence of an observer's prior beliefs when the observer hears evidence at trial rather than in the real world. As explained in Part III, these priors can be heavily influenced by the observer's stereotypes and prejudices regarding the background characteristics of a defendant or other trial participant.

Therefore, the incongruence principle implies that the devaluing effect of the trial on evidence will disproportionately disadvantage certain groups of trial participants based on their race, sex, appearance, economic status, and other background characteristics.[199] For example, I have argued that character evidence invites jurors to rely on their prior beliefs, stereotypes, and prejudices in determining a verdict. This reliance will of course have different implications for individuals of different races, genders, professions, etc.[200] The incongruence principle highlights the important role of the trial setting in causing such reliance and its implications for inequality.

V. CONCLUSION

Evidence law regularly draws on the real-world predictive value of information or the inferences that a real-world observer would or should make from evidence. This much is clear from legislative history, Advisory Committee notes, judicial opinions, and scholarship. Simply stated, the real-world setting is central to judicial and legislative understandings of how jurors use evidence at trial, and it is central to the rules of evidence, their underlying policies, and their application in the courts. Indeed, this reliance on the real-world evidentiary context is well-founded. After all, jurors (and other trial observers) draw on their everyday experience, knowledge, and reasoning to make inferences from evidence at trial.

However, as the incongruence principle implies, it is a mistake to rely on real-world settings without accounting for important differences between the real-world and trial evidentiary contexts. In particular, the incongruence principle brings to light the trial setting's devaluing effect on evidence.

The analysis above applies Bayesian reasoning to demonstrate this devaluing effect. It emphasizes the significant role of prior beliefs in the inferences that jurors and other trial participants draw from evidence, and it shows how the trial setting is

---

199. *See Character Evidence as a Conduit for Implicit Bias*, *supra* note 15, at 36–44.
200. *See id.* at 26–33.

particularly susceptible to an inferential problem called overfitting, which can lead to false-positive judgments via the influence of strong priors.

I apply the incongruence principle and its accompanying evidence devaluation principles in a range of evidentiary contexts in order to demonstrate a few of its important implications. My analysis suggests that lawmakers should account for this devaluation when creating law and policy surrounding different categories of evidence, and courts should account for it when analyzing evidence in a particular case. Ultimately, my analysis is not intended to criticize the U.S. trial's analytical framework. Rather, it is to highlight the importance of recognizing the incongruence between information in the real world and information at trial, and particularly, the devaluing effect of the trial setting on evidence through both its influence on the prior beliefs of trial participants and the vulnerabilities in its analytical framework. Accounting for this effect would enable legislatures and courts to better achieve the accuracy, fairness, and equality aims of evidence law.