

ESSAY

HINDSIGHT EVIDENCE

*Maggie Wittlin**

Judges and juries are frequently called upon to evaluate a party's actions in retrospect—with the benefit of hindsight. Traditionally, courts and scholars have been understandably wary about how hindsight bias influences verdicts, focusing on how to keep outcome information away from jurors and how to minimize its influence on adjudication. But outcome information can be probative evidence: Bad outcomes can be indicative of bad decisionmaking. In this Essay, I aim to rehabilitate the use of outcome information by conceptualizing it as a new category of evidence: hindsight evidence. First, I develop a framework for deciding how much weight to afford hindsight evidence and whether it should be admitted to a jury under Rule 403, which requires judges to weigh probative value against prejudicial impact. As for relevance and probative weight, I show that hindsight evidence is probative to the extent that facts supporting one party's theory of the case have a greater tendency to generate that outcome than facts supporting the other party's theory. As for prejudice, I review the research on hindsight bias and the factors that mitigate its impact. Finally, I apply this framework to four paradigmatic examples from diverse areas of the law—civil rights, contracts, special education law, and civil procedure—where courts have disagreed about whether to consider hindsight evidence. Ultimately, I conclude that a deeper theoretical understanding of how hindsight evidence operates will allow courts to embrace its value more readily.

INTRODUCTION	1324
I. HINDSIGHT EVIDENCE DEFINED.....	1329
II. RELEVANCE	1332
A. The Affirmative Case for Relevance.....	1333

* Associate in Law, Columbia Law School. For helpful conversations and comments, I thank Ian Ayres, Ilya Beylin, Kiel Brennan-Marquez, Richard Briffault, Rick Brooks, Ed Cheng, Jeff Fagan, Dennis Fan, Heather Gerken, Victor Goldberg, David Goodwin, Zohar Goshen, Logan Gowdey, Bernard Harcourt, Claudia Haupt, Bert Huang, Rob Jackson, Dan Kahan, Avery Katz, Richard Lempert, Jim Liebman, Erin McCormack-Herbert, Alex Nunn, Michael Pardo, Dave Pozen, Jaya Ramji-Nogales, Alex Raskolnikov, Gustavo Ribeiro, Dan Richman, Bob Scott, Dan Simon, Christina Parajon Skinner, David Sklansky, Alex Stein, Kate Stith, Allison Tait, Dane Thorley, Alan Trammell, Ryan Williams, everyone who discussed the paper with me during job interviews, and participants in the Associates and Fellows Workshop at Columbia Law School.

1. Bayesian Reasoning	1334
2. Inference to the Best Explanation	1341
B. Must Bear on an Issue of Fact	1344
III. PROBATIVE VALUE	1351
A. Inevitable Hindsight Evidence and Prior Odds	1352
B. Nonbinary Hindsight Evidence	1355
C. Res Ipsa Loquitur	1357
IV. PREJUDICE, BALANCING, AND ADMISSION	1359
A. Hindsight and Outcome Biases	1359
B. Debiasing	1362
1. Consider the Opposite	1363
2. Consequence Severity	1367
3. Bifurcation	1368
4. Rules that Counteract Hindsight Bias	1369
5. Summary	1370
C. Admitting Hindsight Evidence	1370
V. APPLICATION TO AREAS OF DOCTRINAL CONFUSION	1372
A. Excessive Force and Self-Defense	1372
B. Liquidated Damages	1378
C. Individuals with Disabilities Education Act	1384
D. Judgment as a Matter of Law	1390
CONCLUSION	1393

INTRODUCTION

A police officer approaches a car occupied by two men. The passenger, known to the officer as a suspect in several prior crimes, fits the description of a man who robbed a nearby store minutes earlier. The officer and his partner move toward the vehicle with their weapons raised. The officer orders the men inside to raise their hands—he must do so three times before they comply, and when their hands are raised the occupants exchange a look. As the officer approaches the car, the driver reaches into his coat. The officer will later describe this as a “quick movement,” like “he was going to reach for a weapon,” although other witnesses will dispute that characterization. At that moment, the officer fires into the car, killing the driver instantly. A search of the driver’s body fails to find a weapon.

If the driver’s family files a section 1983 lawsuit against the officer, claiming that the officer used excessive force in violation of the Fourth Amendment, can the plaintiff introduce evidence that the driver was unarmed? Does the absence of any weapon make it less likely that the officer’s actions were “objectively reasonable” in light of the facts and

circumstances confronting [him]” “on the scene”?¹ In this actual case, *Sherrod v. Berry*, a majority of the Seventh Circuit, sitting in banc, answered these questions in the negative.² The fact that the deceased was unarmed is not at all relevant to the jury’s determination, the majority wrote, because the jury must not “possess more information than the officer possessed when he made the crucial decision.”³ Three dissenting judges (and one concurring judge) disagreed. According to these dissenters, knowledge that the driver was unarmed could help the jury evaluate both the “type of movement” the deceased made and the credibility of the officer-witness.⁴

Which side was correct, the majority or the dissent? The implications of that answer extend far beyond the domain of excessive force cases. Judging decisionmaking in retrospect—with the benefit of hindsight—is a central function of the courts. In a negligence case, a plaintiff who has suffered some harm⁵ tries to prove that the defendant breached the duty to behave reasonably.⁶ Similarly, in a securities fraud case, a plaintiff who has experienced a loss aims to demonstrate that the defendant acted with scienter, an “intention ‘to deceive, manipulate, or defraud.’”⁷ Parents who sue under the Individuals with Disabilities Education Act (IDEA)⁸ must show that the school board’s plan for their child’s education was unreasonable,⁹ and they often come armed with test results showing their child did not, in fact, progress under the plan.¹⁰ A party to a contract

1. *Graham v. Connor*, 490 U.S. 386, 396–97 (1989).

2. 856 F.2d 802, 805, 807 (7th Cir. 1988) (en banc).

3. *Id.* at 805.

4. *Id.* at 810–11 (Cummings, J., dissenting); see also *id.* at 813 (Flaum, J., dissenting) (“[T]he evidence presented at trial on the issue of whether the deceased was armed was relevant . . .” (emphasis omitted)). In subsequent excessive force cases around the country, where an officer claims the person he shot appeared to be reaching for a weapon but no weapon was found, most, but not all, courts have followed the Seventh Circuit. See *infra* section V.A (discussing hindsight evidence in excessive force cases).

5. See Restatement (Second) of Torts § 281 (Am. Law Inst. 1965) (requiring actor’s conduct cause an “invasion of interest” in order to sustain a negligence claim).

6. Arthur Best & David W. Barnes, *Basic Tort Law* 89–91 (2d ed. 2007).

7. *Tellabs, Inc. v. Makor Issues & Rights, Ltd.*, 551 U.S. 308, 313 (2007) (quoting *Ernst & Ernst v. Hochfelder*, 425 U.S. 185, 193 n.12 (1976)); see also 15 U.S.C. § 78j(b) (2012) (prohibiting use of “any manipulative or deceptive device” in contravention of SEC’s rules).

8. 20 U.S.C. §§ 1400–1482 (2012).

9. See *Bd. of Educ. v. Rowley*, 458 U.S. 176, 207 (1982) (stating plan must be “reasonably calculated to enable the child to receive educational benefits”); see also *Fuhrmann ex rel. Fuhrmann v. E. Hanover Bd. of Educ.*, 993 F.2d 1031, 1040 (3d Cir. 1993) (applying “*Rowley*’s threshold determination of a ‘reasonable calculation’ of educational benefit”); *Roland M. v. Concord Sch. Comm.*, 910 F.2d 983, 992 (1st Cir. 1990) (“[T]he issue is . . . whether [the plan] was ‘reasonably calculated’ to provide an ‘appropriate’ education . . .” (quoting *Rowley*, 458 U.S. at 207)).

10. See Dennis Fan, Note, No IDEA What the Future Holds: The Retrospective Evidence Dilemma, 114 *Colum. L. Rev.* 1503, 1523–24 (2014) (discussing “actual-progress evidence”).

seeking to avoid a liquidated damages clause must show the clause was an ex ante unreasonable estimate of damages, and it must do so after one party has *actually* suffered damages.¹¹ Judges must frequently decide whether the outcome of a decision is probative of that decision's quality and if so, whether the fact-finder should consider the outcome in evaluating the actor's conduct.

Yet courts have failed to systematically consider information about the outcome in individual cases as a potential source of probative value—as a category of evidence bearing on ex ante probability or the quality of an actor's decisionmaking. Instead, courts determine what to do with outcome information in a scattershot fashion.¹² They occasionally acknowledge its relevance¹³ but just as often label outcome information “irrelevant” and exclude it, as in *Sherrod*, or attempt to mitigate its influence on the decisionmaker.¹⁴ They almost never provide a considered explanation for *why* the information is or is not relevant evidence. Instead, courts either assert relevance as obvious¹⁵ or more typically, decline to consider the information because to do so would be to judge the actor's decisions “with the 20/20 vision of hindsight.”¹⁶

The haphazard nature of these decisions suggests that a robust theory of when outcome information is relevant and probative could improve rational decisionmaking in the courts. Although a few legal scholars have noted that outcome information may, under certain circumstances, be probative,¹⁷ none has set forth a theory of when this

11. Restatement (Second) of Contracts § 356 (Am. Law Inst. 1981).

12. See *infra* Part V (discussing four areas of law where courts have diverged regarding relevance of hindsight evidence).

13. See *infra* sections V.A–C (noting that whether victims were armed is deemed relevant in self-defense cases, actual damages are sometimes treated as relevant in liquidated damages cases, and educational outcomes are sometimes deemed relevant in IDEA cases).

14. See *infra* sections V.A–C (noting that whether victims were armed is often deemed irrelevant in excessive force cases, actual damages are sometimes treated as irrelevant in liquidated damages cases, and educational outcomes are sometimes deemed irrelevant in IDEA cases).

15. See, e.g., *Cypress-Fairbanks Indep. Sch. Dist. v. Michael F.*, 118 F.3d 245, 253 (5th Cir. 1997) (endorsing district court's test for whether education plans are reasonably calculated to provide educational benefits under IDEA, where test includes consideration of whether “positive academic and non-academic benefits are demonstrated”).

16. *Graham v. Connor*, 490 U.S. 386, 396 (1989); see, e.g., *Sherrod v. Berry*, 856 F.2d 802, 805 (7th Cir. 1988) (en banc) (“Knowledge of facts and circumstances gained after the fact . . . has no place in the . . . proper post-hoc analysis of the reasonableness of the actor's judgment. Were the rule otherwise . . . the jury would possess more information than the officer possessed when he made the crucial decision.”).

17. See Chris William Sanchirico, *Finding Error*, 2003 Mich. St. L. Rev. 1189, 1194 (arguing it is rational to use outcome information to evaluate whether another person knew or should have known outcome was likely to occur, if that person “may have been in a position to know” about the outcome's likelihood but may now not “fully reveal what information that position provided”); see also Mitu Gulati, Jeffrey J. Rachlinski & Donald C. Langevoort, *Fraud by Hindsight*, 98 Nw. U. L. Rev. 773, 786 (2004) (“If poor predict-

information should be considered as evidence in the courtroom. In this Essay, I take up that challenge and aim to rehabilitate the use of outcome information by conceptualizing it as a new category of evidence: hindsight evidence. Hindsight evidence is relevant to the extent that one party's story of what happened has a greater tendency to generate that hindsight evidence—to produce that result—than the other party's story. This is so for the simple reason that bad decisions tend to produce worse outcomes than good decisions, so a bad outcome is more indicative of a bad decision than a good decision. Under this framework, hindsight evidence is indeed frequently relevant.

This is not to say courts' concerns about prejudice are unfounded. Hindsight bias—the tendency to view an event as more likely or predictable after it happened than it actually was before it occurred¹⁸—and its cousin, outcome bias—the tendency to judge decisionmaking in light of outcome, independent of how likely or predictable the outcome was¹⁹—are real and troubling. Social scientists have been studying hindsight bias and outcome bias for decades, and their research overwhelmingly demonstrates that after learning about an outcome, people view that outcome as more predictable than it actually was²⁰ and people judge decisionmakers inappropriately based on the outcomes of their decisions.²¹ Research examining how these biases manifest themselves in

tions and misconduct are more likely to produce a failed forecast than good predictions and honest conduct, then the fact of the bad outcome, by itself, supports an inference that the conduct was unreasonable.”); Fan, *supra* note 10, at 1547 (“This Note suggests first that a nonexclusionary approach that permits relevant retrospective evidence should prevail.”).

18. Dustin P. Calvillo & Abraham M. Rutchick, Domain Knowledge and Hindsight Bias Among Poker Players, 27 *J. Behav. Decision Making* 259, 259 (2014).

19. Jonathan Baron & John C. Hershey, Outcome Bias in Decision Evaluation, 54 *J. Personality & Soc. Psychol.* 569, 570 (1988).

20. See Jeffrey J. Rachlinski, A Positive Psychological Theory of Judging in Hindsight, 65 *U. Chi. L. Rev.* 571, 580 (1998) [hereinafter Rachlinski, *Judging in Hindsight*] (“Virtually every study on judging in hindsight has concluded that events seem more predictable than they actually are.”); see also Jay J.J. Christensen-Szalanski & Cynthia Fobian Willham, The Hindsight Bias: A Meta-Analysis, 48 *Organizational Behav. & Hum. Decision Processes* 147, 162–64 (1991) (analyzing 122 studies of hindsight bias and concluding it occurs under some conditions); Rebecca L. Guilbault et al., A Meta-Analysis of Research on Hindsight Bias, 26 *Basic & Applied Soc. Psychol.* 103, 111–15 (2004) (analyzing ninety-five studies of hindsight bias and determining what factors exacerbate or mitigate bias); Neal J. Roese & Kathleen D. Vohs, Hindsight Bias, 7 *Persp. on Psychol. Sci.* 411, 411 (2012) (noting over 800 academic articles discuss hindsight bias).

21. See, e.g., Philip J. Mazzocco, Mark D. Alicke & Teresa L. Davis, On the Robustness of Outcome Bias: No Constraint by Prior Culpability, 26 *Basic & Applied Soc. Psychol.* 131, 144–45 (2004) (demonstrating outcome bias in American undergraduates); Francesca Gino et al., No Harm, No Foul: The Outcome Bias in Ethical Judgments 7 (*Harvard Bus. Sch., Working Paper No. 08-080*, 2009), <http://ssrn.com/abstract=1099464> (on file with the *Columbia Law Review*) (demonstrating “outcome information impacts an observer’s evaluation of the ethicality of a target’s behavior”).

legal contexts suggests that both judges and jurors are, at times, prone to fall victim to these biases.²²

The problem of overvaluing outcome information does not, however, negate its probative value where fact-finders are justified in reassessing the propriety of the actor's decision in light of that evidence. In bench trials and judicial hearings, where judges generally do not exclude evidence for unfair prejudice or confusion,²³ relevant evidence should be considered and weighed in accordance with its probative value. In jury trials, judges deciding whether to admit the evidence should weigh its probative value against the danger of biased assessment.²⁴

This Essay offers a theoretical framework for evaluating when courts should consider hindsight evidence. Part I defines "hindsight evidence." Part II provides a framework that courts can use to decide when hindsight evidence is *relevant* under Federal Rule of Evidence 401. To be relevant, hindsight evidence must bear on a question of fact, defined as a question on which the parties may use evidence to educate the jury. Part III then addresses how courts should think about the *probative weight* of hindsight evidence in several specific circumstances: when a cause of action requires that evidence of an outcome be presented (for example, the existence of harm in a negligence case), when outcomes vary in magnitude, and when outcomes vary in their individual characteristics. This theory justifies—and partly renders unnecessary—the doctrine of *res ipsa loquitur*.

22. See, e.g., Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, The "Hidden Judiciary": An Empirical Examination of Executive Branch Justice, 58 Duke L.J. 1477, 1512–16 (2009) (observing administrative law judges were more likely to say a store had improperly detained an African American teenager based on his race when he was not found with stolen merchandise than when he was); Kim A. Kamin & Jeffery J. Rachlinski, Ex Post ≠ Ex Ante: Determining Liability in Hindsight, 19 Law & Hum. Behav. 89, 98 (1995) (finding a sizable majority of mock committee members decided, in foresight, that town should not take certain precautions to protect property owner from flood damage; however, when damage was already done, majority of mock jurors decided, in hindsight, that failure to take precaution was negligent); Gregory N. Mandel, Patently Non-Obvious: Empirical Demonstration that the Hindsight Bias Renders Patent Decisions Irrational, 67 Ohio St. L.J. 1391, 1406–11 (2006) (finding mock jurors in patent case were more likely to find an invention obvious when they knew the solution to the problem posed versus when they received only the problem).

23. See *United States v. Preston*, 706 F.3d 1106, 1117 (9th Cir. 2013) ("Rule 403 is inapplicable to bench trials."), *rev'd on other grounds*, 751 F.3d 1008 (9th Cir. 2014) (en banc); *Schultz v. Butcher*, 24 F.3d 626, 632 (4th Cir. 1994) ("[W]e hold that in the context of a bench trial, evidence should not be excluded under 403 on the ground that it is unfairly prejudicial."); *Gulf States Utils. Co. v. Ecodyne Corp.*, 635 F.2d 517, 519 (5th Cir. Unit A Jan. 1981) ("The exclusion of this evidence under Rule 403's weighing of probative value against prejudice was improper. This portion of Rule 403 has no logical application to bench trials."). But see Christopher B. Mueller & Laird C. Kirkpatrick, *Evidence* § 4.9, at 178 (5th ed. 2012) (suggesting judges exclude unfairly prejudicial evidence in bench trials "to ensure the integrity of the factfinding process and to make clear for the parties and the appellate court the basis for the trial court's decision").

24. See Fed. R. Evid. 403.

Part IV discusses the problem of hindsight bias and addresses when a judge should admit hindsight evidence in a jury trial. It suggests judges should pay attention to the magnitude of hindsight bias observed in psychological studies—which is significant but far from universally dispositive—and should attend to differences between the laboratory and courtrooms that may mitigate the bias’s effects. Part IV then discusses several debiasing techniques and suggests how and when they can be most useful.

Part V applies the framework by reviewing a selection of paradigmatic examples from diverse areas of law where courts encounter hindsight evidence and have disagreed about whether to use it. These include: excessive force cases like *Sherrod v. Berry*, where a police officer shoots a suspect and claims he thought the suspect had a gun, but there is evidence that the suspect was unarmed; contract cases addressing whether a liquidated damages clause was a reasonable forecast of damages, and there is evidence of actual damages; cases under the IDEA where a disabled student’s parents claim the school board did not create a plan “reasonably calculated” to achieve educational benefits, and there is evidence of the student’s actual progress or lack thereof; and motions for judgment as a matter of law, when information about an actual jury’s decision could be used to determine what a “reasonable jury” could decide. These examples demonstrate that courts’ failure to understand outcomes as a category of evidence has led to confused thinking about the appropriate use of outcome information and a failure to admit or employ this information where it could be probative and helpful. Each example includes a discussion of how courts should use the framework developed here to think about hindsight evidence. Ultimately, these examples suggest that courts should admit and consider hindsight evidence more frequently than they currently do.

I. HINDSIGHT EVIDENCE DEFINED

Whenever a fact-finder is asked to consider evidence of an outcome in order to judge an earlier action in retrospect, that fact-finder has encountered “hindsight evidence.” To formalize this definition, the term “hindsight evidence”²⁵ as used here refers to outcome information in-

25. Several courts and scholars have employed the term “hindsight evidence,” often with a negative connotation, to refer to the use of outcomes to prove earlier states of affairs. See, e.g., *S. Bancorporation v. Comm’r.*, 847 F.2d 131, 137 (4th Cir. 1988) (discussing the prohibition on using “hindsight evidence” to evaluate reasonableness of a taxpayer’s depreciation deduction); *McGee v. Bowen*, 647 F. Supp. 1238, 1248–49 (N.D. Ill. 1986) (criticizing reliance on “hindsight evidence,” showing applicant was not disabled in 1966, to determine that she was not disabled for long in 1964); *Gulati et al.*, supra note 17, at 809–11 (discussing whether “hindsight evidence” may be used in securities fraud cases). They have also occasionally used the term “retrospective evidence,” see, e.g., *Fan*, supra note 10, at 1505 (defining “retrospective evidence,” used frequently in cases under the IDEA), or simply “outcome information,” see, e.g., *Sanchirico*, supra note 17, at 1194

roduced to prove either (1) the objective ex ante likelihood that the outcome would materialize (such as whether a toaster was likely to explode)²⁶ or (2) the decisionmaking quality of an actor whose decision relied on an estimate of the outcome's likelihood. The second category, which focuses on actors' decisionmaking, encompasses both actors who rely on a prediction of a *future* outcome, such as how likely it is that a natural disaster will occur in the next year,²⁷ and actors who rely on an appraisal of an *existing* state of affairs, such as how likely it is a criminal suspect has contraband in his car.²⁸ In the first instance, a plaintiff might attempt to introduce evidence that a natural disaster occurred to show that the defendant's failure to take precautions was negligent. In the second, the police might attempt to introduce the presence of cocaine to show that their search of a plaintiff's car was reasonable. In both cases, the court should evaluate the admissibility of the evidence under the framework introduced here.

This definition of "hindsight evidence" is broad by design and is meant to encompass most situations in which a court is likely to experience hindsight-bias anxiety—the worry that a fact-finder will overestimate ex ante likelihood or improperly evaluate decisionmaking based on later-received information.²⁹ The umbrella of hindsight evidence unites seemingly distinct types of information into a single analytical framework based on this shared characteristic.

Hindsight evidence is, however, limited to *individual* outcomes. It is distinct from cases where a party presents a large, representative sample of outcomes and an expert testifies that proper statistical analysis of that data yields reliable conclusions about the cause of the outcomes. Courts have more frequently recognized the probative value of that kind of

(discussing "rational cross-person use of outcome information"). This Essay introduces the term "hindsight evidence" as used to describe a certain category of potentially useful evidence.

26. See Oren Bar-Gill & Elizabeth Warren, Making Credit Safer, 157 U. Pa. L. Rev. 1, 7 (2008) (referring to "exploding toasters" as "unreasonably dangerous products").

27. See Kamin & Rachlinski, *supra* note 22, at 93 (describing natural-disaster-likelihood experiment).

28. See Jonathan D. Casper, Kennette Benedict & Jo L. Perry, Juror Decision Making, Attitudes, and the Hindsight Bias, 13 Law & Hum. Behav. 291, 294–96 (1989) (describing use of search of suspect's home as stimulus for mock jurors in hindsight experiment).

29. It does not, however, include subsequent remedial measures evidence, which is excluded by Rule 407. Although evidence of subsequent remedial measures may pose a danger of hindsight bias, see Dan M. Kahan, The Economics—Conventional, Behavioral, and Political—of "Subsequent Remedial Measures" Evidence, 110 Colum. L. Rev. 1616, 1622–24 (2010), a subsequent remedial measure is not itself an outcome. Instead, it follows an outcome, and it may be probative either of prior knowledge that the precaution was necessary to avert an accident, see *id.* at 1623, 1633, or of a postaccident, revised probability of injury without the precaution.

statistical proof.³⁰ In toxic torts cases, for example, plaintiffs commonly introduce epidemiological studies that compare the incidence of disease in a population that has been exposed to the toxin at issue to the incidence of disease in an unexposed population, thereby deriving the relative risk of the exposed group.³¹ The relative likelihood of developing the disease is then used to demonstrate causation.³² The Supreme Court has also relied on statistical evidence demonstrating race or sex discrimination in a number of contexts, including jury pool selection³³ and employment.³⁴ Courts have been less willing to use statistical evidence to make inferences about decisionmaking in an individual case.³⁵

The relevance of statistical evidence is easier to understand (and stomach) than the relevance of hindsight evidence. With a large sample, a party's expert can demonstrate that his or her results are "statistically significant"—that under generally accepted academic norms,³⁶ they are

30. See Fed. Judicial Ctr., Reference Manual on Scientific Evidence 214 (3d ed. 2011) ("Statistical studies suitably designed to address a material issue generally will be admissible under the Federal Rules of Evidence.").

31. See Bert Black & David E. Liliendorf, Epidemiologic Proof in Toxic Tort Litigation, 52 Fordham L. Rev. 732, 755–61 (1984) (discussing epidemiological methodologies); Steve Gold, Note, Causation in Toxic Torts: Burdens of Proof, Standards of Persuasion, and Statistical Evidence, 96 Yale L.J. 376, 379–80 (1986) (discussing "relative risk").

32. See Fed. Judicial Ctr., *supra* note 30, at 612 ("Some courts have reasoned that when epidemiologic studies find that exposure to the agent causes an incidence in the exposed group that is more than twice the incidence in the unexposed group . . . the probability that exposure to the agent caused a similarly situated individual's disease is greater than 50%."); 2 Margie Tyler Searcy, A Guide to Toxic Torts § 10.02 (Cumulative Supp. 1996) (discussing how epidemiological evidence may be used to show general or, less frequently, specific causation).

33. See *Turner v. Fouche*, 396 U.S. 346, 360 (1970) (holding petitioners "made out a prima facie case of jury discrimination" by showing substantial disparity between percentage of African Americans in population and jury pool and showing disparity originated at point of subjective judgment).

34. See *Bazemore v. Friday*, 478 U.S. 385, 398–401 (1986) (suggesting petitioners' "multiple regression analyses designed to demonstrate that blacks were paid less than similarly situated whites" were probative).

35. See, for example, *McCleskey v. Kemp*, in which a condemned black man, convicted of killing a white police officer, proffered the now-iconic Baldus study, which used over 2,000 Georgia murder cases to demonstrate that defendants who killed white victims—particularly black defendants who killed white victims—were more likely to receive the death penalty than defendants who killed black victims. 481 U.S. 279, 286–87 (1987). The petitioner argued this evidence demonstrated systemic racial bias, Brief for Petitioner at 26, *McCleskey*, 481 U.S. 279 (No. 84-6811), 1986 WL 727359, but the Court held the study "clearly insufficient to support an inference that any of the decisionmakers in *McCleskey's* case acted with discriminatory purpose." *McCleskey*, 481 U.S. at 297.

36. Recently, elements of classical hypothesis testing have come under attack. See Regina Nuzzo, Statistical Errors, 506 Nature 150, 150–52 (2014) (discussing criticism of use of p-values in hypothesis testing); David Trafimow & Michael Marks, Editorial, 37 Basic & Applied Soc. Psychol. 1, 1–2 (2015) (banning null-hypothesis significance testing in the journal).

“so unlikely to have occurred by chance alone that it is reasonable to attribute the result to something else.”³⁷ By eliminating other possibilities through statistical analysis, experts can make a case for what that “something else” is: They can assure a fact-finder that their side’s story is the only legitimate explanation for the data presented. Hindsight evidence offers no such assurances. A single data point cannot yield firm conclusions in the same way that a large, representative data sample can. But “a brick is not a wall”;³⁸ evidence need not be conclusive to be relevant. And as demonstrated in the next Part, hindsight evidence is often relevant.

II. RELEVANCE

Three consecutive Federal Rules of Evidence—Rules 401, 402, and 403—establish the foundational framework for admissibility of evidence in federal court. Rule 401 defines “relevant evidence” as evidence that “has any tendency to make a” “fact . . . of consequence in determining the action” “more or less probable than it would be without the evidence.”³⁹ Rule 402 then provides that relevant evidence is generally admissible.⁴⁰ Rule 403, however, gives judges broad discretion to exclude relevant evidence when its “probative value is substantially outweighed by” the risk of “unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence.”⁴¹ When no other rule or law governs the admissibility of a challenged piece of evidence, the court weighs probative value against prejudice to determine whether that evidence should go before a jury.

These next three Parts trace the considerations set out in these Rules—relevance, probative value, and weighing probative value against the risk of prejudice—thereby establishing a framework for determining when hindsight evidence should be admitted and when it should be excluded. Many of these considerations apply to bench trials as well as to jury trials. Most courts agree that Rule 403 exclusion does not apply to

37. D.H. Kaye, *Is Proof of Statistical Significance Relevant?*, 61 Wash. L. Rev. 1333, 1333 (1986); see, e.g., *Segar v. Smith*, 738 F.2d 1249, 1261–63 (D.C. Cir. 1984) (“[T]he odds are far less than one in a thousand that the observed disparities for any year could have occurred by chance. A study is generally considered to be statistically significant when the odds that the result occurred by chance are at best one in 20.”).

38. Charles T. McCormick, *Handbook of the Law of Evidence* § 152, at 317 (1954).

39. Fed. R. Evid. 401.

40. Fed. R. Evid. 402.

41. Fed. R. Evid. 403. Rule 403 departs from Jeremy Bentham’s idea of “free proof,” under which evidence would be excluded *only* for costliness, waste of time, or delay. See 1 Jeremy Bentham, *Rationale of Judicial Evidence* 1 (John S. Mill ed., 1827) (“[M]erely with a view to rectitude of decision . . . no species of evidence whatsoever . . . ought to be excluded: . . . [in cases where evidence should not be admitted] the reason for the exclusion rests on other grounds; viz. avoidance of vexation, expense, and delay.”); Frederick Schauer, *On the Supposed Jury-Dependence of Evidence Law*, 155 U. Pa. L. Rev. 165, 169 (2006) (discussing Bentham’s “free proof” approach).

bench trials, as there is no way to filter evidence before it reaches the judge's eyes and ears.⁴² But judges acting as fact-finders still endeavor to afford evidence its appropriate probative value and avoid prejudice.⁴³ The following discussion of the probative value and pitfalls of hindsight evidence therefore applies even where the evidence will not be excluded and a judge simply wants to put it to its best possible use. The discussion begins with relevance.

This Part addresses when hindsight evidence is and is not relevant. It first presents a general, affirmative case for relevance, demonstrating under two central theories of relevance discussed in the evidence literature that whenever outcomes are correlated with decisionmaking quality—whenever better decisions tend to produce better outcomes and vice versa—outcome information should influence a rational fact-finder's beliefs about decisionmaking quality. It then discusses when hindsight evidence is *not* relevant: namely, where it does not bear on a fact of consequence to the action.

A. *The Affirmative Case for Relevance*

Under Rule 401, relevant evidence is that which has “any tendency to make a fact more or less probable than it would be without the evidence.”⁴⁴

Evidence theorists disagree about how best to understand the mechanism by which evidence makes a fact more or less plausible than it would otherwise be. Two camps currently dominate the discussion. First, proponents of Bayesian reasoning rely on probability theory to model how rational jurors should update their beliefs about consequential facts given new evidence.⁴⁵ Evidence is relevant under this model to the extent that it is more consistent with one factual proposition than it is with the negation of that proposition.⁴⁶ Second, proponents of reasoning by inference to the best explanation contend that jurors find facts by

42. See *supra* note 23 (citing judicial statements of Rule 403's inapplicability to bench trials).

43. See, e.g., *Schultz v. Butcher*, 24 F.3d 626, 631–32 (4th Cir. 1994) (“For a bench trial, we are confident that the district court can hear relevant evidence, weigh its probative value and reject any improper inferences.”).

44. Fed. R. Evid. 401.

45. See Reid Hastie, Introduction [hereinafter Hastie, Introduction] to *Inside the Juror: The Psychology of Juror Decision Making* 3, 11 (Reid Hastie ed., 1993) [hereinafter *Inside the Juror*] (“[M]any legal scholars [have] . . . propose[d] the application of [Bayes's] theorem as a prescription or description for factfinders' reasoning in legal cases.”); see also, e.g., Ian Ayres & Barry Nalebuff, *The Rule of Probabilities: A Practical Approach for Applying Bayes' Rule to the Analysis of DNA Evidence*, 67 *Stan. L. Rev.* 1447, 1457–66 (2015) (applying Bayesian updating to DNA evidence); Kahan, *supra* note 29, at 1631–32 (discussing role of Bayes's Theorem in updating beliefs given evidence).

46. Kahan, *supra* note 29, at 1631.

choosing the best explanation⁴⁷ for the evidence presented at trial. Evidence is relevant to the extent that it “is explained by, and hence justifies” the narrative offered by the party introducing the evidence.⁴⁸

This Essay does not choose between these theories but instead demonstrates that under *both* theories, hindsight evidence is often relevant. And it is relevant for essentially the same reason under each theory: When one occurrence tends to lead to another, evidence of the second occurrence is suggestive of the first. Under Bayesian reasoning, this is so because evidence of the second occurrence is more likely to surface if the first occurrence happened than if it did not. Under inference to the best explanation, it is true because the first occurrence explains the second. In the case of hindsight evidence, where bad decisions lead to bad outcomes more than good decisions lead to bad outcomes, a bad outcome is probative of a bad decision. That is the central insight into relevance. What follows is detail.

1. *Bayesian Reasoning.* — Bayesian theory provides a model for how a rational juror, or any rational fact-finder, should evaluate new evidence as he or she receives it.⁴⁹ For decades, legal scholars have used this theory for its normative value: Bayesian reasoning is idealized, rational fact-finding, and to the extent that actual legal reasoning deviates from Bayesian reasoning, these scholars have suggested, the deviation may suggest a need for improvement.⁵⁰

The Bayesian model presumes that before receiving the evidence, a fact-finder—tasked with determining whether a factual proposition is true or false—has some belief about how probable it is that the fact is true.⁵¹ For instance, before receiving certain testimony, a juror in a

47. See Ronald J. Allen & Alex Stein, Evidence, Probability, and the Burden of Proof, 55 Ariz. L. Rev. 557, 567–71 (2013) (arguing adjudicative fact-finding is inference to best explanation reasoning); Michael S. Pardo & Ronald J. Allen, Juridical Proof and the Best Explanation, 27 Law & Phil. 223, 223–25 (2008) [hereinafter Pardo & Allen, Juridical Proof] (claiming juridical-proof processes involve “abduction” or “inference to the best explanation”).

48. Pardo & Allen, Juridical Proof, *supra* note 47, at 242 (defining probative value by extending their definition of relevance).

49. See Richard O. Lempert, Modeling Relevance, 75 Mich. L. Rev. 1021, 1023 (1977) [hereinafter Lempert, Modeling Relevance] (“Since the law assumes that a factfinder *should* be rational, this is a normative model . . .”).

50. See Hastie, Introduction, *supra* note 45, at 12 (“A majority of the twenty-two entries in [a 1986 issue of *Boston University Law Review*] promote[d] the application of Bayes’ theorem to elucidate or improve legal reasoning from evidence.”); Lempert, Modeling Relevance, *supra* note 49, at 1056 (“The Bayesian model is normative—it specifies the way in which jurors are expected to evaluate evidence . . .”); Sandy Zabell, A Mathematician Comments on Models of Juror Decision Making, *in* Inside the Juror, *supra* note 45, at 263, 263–64 (“[A]s a theory of consistency, it helps us to police our beliefs.”).

51. This pre-evidence probability estimate is called the “prior odds.” Richard D. Friedman, A Very Brief Primer on Bayesian Methods in Evidence, AALS Evidence Sec. Newsl. (Ass’n of Am. Law Sch., Wash., D.C.), 2002, at 3–4 [hereinafter Friedman, Bayesian Methods].

murder case may believe the odds are 25% that the defendant killed the victim. The fact-finder then uses the new evidence to update this subjective probability estimate. Bayes's Rule, an equation derived from basic formulas in probability theory,⁵² dictates how the fact-finder should perform this updating to achieve maximum accuracy. It instructs the fact-finder to derive a new probability estimate by multiplying her old probability estimate by a mathematical representation of the strength of the new evidence known as the "likelihood ratio."⁵³

The likelihood ratio—the key to Bayesian updating—represents how much the new evidence supports one hypothesis over another.⁵⁴ Formally, it is the ratio of the probability (or frequency) of observing the evidence if the hypothesis is true, divided by the probability (or

52. For the derivation of Bayes's Rule, see Michael O. Finkelstein & William B. Fairley, *A Bayesian Approach to Identification Evidence*, 83 Harv. L. Rev. 489, 498–99 (1970), and Lempert, *Modeling Relevance*, *supra* note 49, at 1022–24 & n.12.

53. In equation form, where "O" is odds, "P" is probability, "E" is evidence, and "G" is the proposition in question ("guilt"), Bayes's Rule dictates:

$$O(G|E) = \frac{P(E|G)}{P(E|\sim G)} O(G)$$

Lempert, *Modeling Relevance*, *supra* note 49, at 1023.

In other words, the odds of guilt, given the new evidence, are the prior odds of guilt multiplied by the probability of observing the evidence if the defendant is guilty, divided by the probability of observing the evidence if the defendant is not guilty. *Id.* at 1023–24.

54. See I.J. Good, *Weight of Evidence: A Brief Survey*, in *Bayesian Statistics 2*, at 249, 250–52 (J.M. Bernardo et al. eds., 1985) ("[W]eight of evidence is equal to the logarithm of the Bayes factor [or likelihood ratio]."). The formulation of Bayes's Rule featured in note 53 addresses binary factual propositions: Did the defendant do it or not? There is also a more general form of Bayes's Theorem that applies when the fact-finder attempts to determine the value of a continuous variable, such as the speed of a vehicle, the amount of damages, or as relevant here, the *ex ante* probability of a negative outcome. The basic idea is the same, however. The evidence shifts the rational fact-finder's estimated relative probability that the variable assumed a given value based on the relative likelihood of observing the evidence if the variable took that value. A witness's testimony that a car "whizzed by in a flash" might be more consistent with the car going eighty miles per hour than with a car going fifty miles per hour, and it might be more consistent with the car going fifty miles per hour than with the car going twenty miles per hour. The evidence would influence the fact-finder's assessment of the car's speed accordingly.

The continuous form of Bayes's Theorem is best expressed as:

$$P(Q|E) = \frac{P(E|Q)}{P(E)} P(Q)$$

$P(Q)$ is the prior probability density function—the relative probability of the variable taking each value before the evidence is observed. $P(E|Q)$ is the "likelihood function," corresponding to the distribution of evidence produced as a function of the variable at issue. Tony Lancaster, *An Introduction to Modern Bayesian Econometrics* 10–11 (2004). $P(E)$ is a normalizing factor, interpretable as the overall probability of observing the evidence. And $P(Q|E)$ is the posterior probability density function, conditional on the evidence.

frequency) of observing the evidence if the hypothesis is false.⁵⁵ The importance of this ratio is intuitive: The more that the evidence is correlated with the truth of a factual proposition, the more the presence of the evidence indicates that the proposition is true.⁵⁶

For example, say that in the murder case mentioned above, a witness testifies that he or she observed the defendant kill the victim. The jurors would each consider everything they know about that testimony—the witness’s demeanor on the stand, any incentives the witness had to lie, the lighting conditions at the time of the crime, and so on—and evaluate how likely it is that they would hear that testimony if the defendant *did* kill the victim, compared to how likely it is that they would hear that testimony if the defendant *did not* kill the victim. To the extent that this evidence is more likely to surface if the defendant is the killer than if he is not, the evidence is relevant to guilt. In the language of Rule 401, the more this likelihood ratio differs from one, the greater the tendency of the evidence to make the fact more or less probable than it would otherwise be. If the jurors started out believing there was a 25% chance that the victim was guilty—1:3 odds—and they concluded that the testimony was six times as likely to be presented if the defendant were guilty than if he or she were innocent, they would each multiply those two numbers— $(6:1) \times (1:3)$ —and conclude the odds are 6:3, or a 2/3 probability, that the defendant is guilty.

Hindsight evidence is relevant from a Bayesian perspective when its likelihood ratio differs from one—when the evidence is more likely to appear under one party’s story than the other’s. Some hindsight evidence is introduced to prove only the *objective* ex ante probability of the outcome materializing. This use of hindsight evidence is not particularly controversial.⁵⁷ For example, in products liability cases, the question may be whether the product was “unreasonably dangerous to the user” as judged by the expectations of an ordinary consumer.⁵⁸ If a toaster

55. Louis Kaplow, Likelihood Ratio Tests and Legal Decision Rules, 16 Am. L. & Econ. Rev. 1, 2–3 (2014).

56. See Jonathan J. Koehler, The Psychology of Numbers in the Courtroom: How to Make DNA-Match Statistics Seem Impressive or Insufficient, 74 S. Cal. L. Rev. 1275, 1300 n.69 (2001) (“Many scholars agree that the probative value of statistical evidence can be represented in terms of a likelihood ratio.”).

57. See Reid Hastie et al., Juror Judgments in Civil Cases: Hindsight Effects on Judgments of Liability for Punitive Damages, 23 Law & Hum. Behav. 597, 609–10 & n.3 (1999) [hereinafter Hastie et al., Juror Judgments] (“[U]nder many conditions this form of learning from experience should not be called an error or a bias at all.”); Sanchirico, supra note 17, at 1198 (“No one could doubt that the fact finder ought to use its knowledge of the outcome in assessing the ex ante objective probability that the event would occur.”).

58. Restatement (Second) of Torts § 402A(1) & cmt. i (Am. Law Inst. 1965). Courts have described this test as a “purely objective determination.” See, e.g., Vaughn v. Nissan Motor Corp. in U.S.A., 77 F.3d 736, 739 (4th Cir. 1996) (“[A] plaintiff’s extraordinary quality . . . has no place in the purely objective determination of whether the product itself is unreasonably dangerous to an ordinary consumer.”). Under the Third Restatement,

explodes,⁵⁹ does the explosion itself tend to prove that the toaster had a relatively high probability of exploding? It does, per the Bayesian likelihood ratio. If an “unreasonably dangerous” toaster is taken to mean one that is highly likely to explode, we are more likely to observe an explosion if the toaster was unreasonably dangerous than if it was not.⁶⁰ As Professor Chris Sanchirico notes, “This is as natural as increasing one’s assessment that a horse is fast after hearing that it won a race.”⁶¹

Now take the more controversial case, where the fact-finder uses hindsight evidence to judge an actor’s decision in retrospect. The probative value of this evidence still corresponds to the likelihood ratio: the probability of the outcome materializing given that the actor made a *bad* decision, relative to the probability of the outcome materializing given that the actor made an *acceptable* decision. In other words, to determine whether the outcome is probative, determine whether there is a correlation between the decisionmaking quality and the occurrence of the outcome.

Where the actor’s decision relied on a probability assessment, hindsight evidence may be useful for either of two purposes. First, if the issue is whether the defendant acted reasonably in light of a risk *known* to the actor, hindsight evidence can tend to prove what that risk *was*, as explained above. For example, if the question is whether it was reasonable to sell a certain toaster, and the defendant knew how dangerous the toaster was but the fact-finder does not, hindsight evidence can tend to prove the level of objective dangerousness, so the fact-finder can then decide whether the risk was unreasonably high. Second, if the issue is whether the actor’s *assessment* of the risk level was reasonable, the probative value of hindsight evidence depends on two

product defect is defined with reference to “*foreseeable* risks of harm posed by the product.” Restatement (Third) of Torts: Prods. Liab. § 2(b) (Am. Law Inst. 1998) (emphasis added).

59. See Bar-Gill & Warren, *supra* note 26, at 7 (introducing the exploding toaster as an example of a dangerous product).

60. A more precise fact-finder would use the continuous form of Bayes’s Theorem here. The fact-finder would start out with some idea of how likely each probability of injury is, expressed as a prior probability distribution. This will probably be weighted heavily toward lower probabilities, since it is unlikely a company would allow a product with a high risk of injury onto the market. The likelihood function here is a line with slope 1, since the probability of observing the hindsight evidence is the probability that the toaster would explode. When these two functions—the prior distribution and the likelihood—are multiplied (and normalized), they yield a posterior probability distribution shifted right relative to the prior distribution: The fact that the injury materialized makes high-risk values more likely than was previously thought. If “unreasonably dangerous” is defined to be at least a 1/10,000 chance of explosion, an ideal fact-finder would integrate the posterior distribution from 1/10,000 to 1 to find the total posterior probability that the product was unreasonably dangerous.

61. Sanchirico, *supra* note 17, at 1198; see also Keith N. Hylton, *The Theory of Tort Doctrine and the Restatement (Third) of Torts*, 54 *Vand. L. Rev.* 1413, 1430 (2001) (“Bayesians will update their predictions of the probability an event will happen after it happens.”).

analytical steps:⁶² The outcome must tend to prove that the actor's probability assessment was *objectively* wrong, and objective incorrectness must tend to prove that the actor's probability assessment was *unreasonable* (or failed to meet the pertinent decisionmaking standard). If both of these are true, the outcome evidence is probative.

For essentially every legal question where outcome information is available in court, that information may indeed be probative in this way. For example, we hold people liable for the results of their negligence because we believe negligence is correlated with injury. Similarly, we require probable cause to conduct a search in part because we believe probable cause is correlated with finding evidence of a crime.⁶³ In these and many other instances, then, outcome information is probative of ex ante decisionmaking.⁶⁴

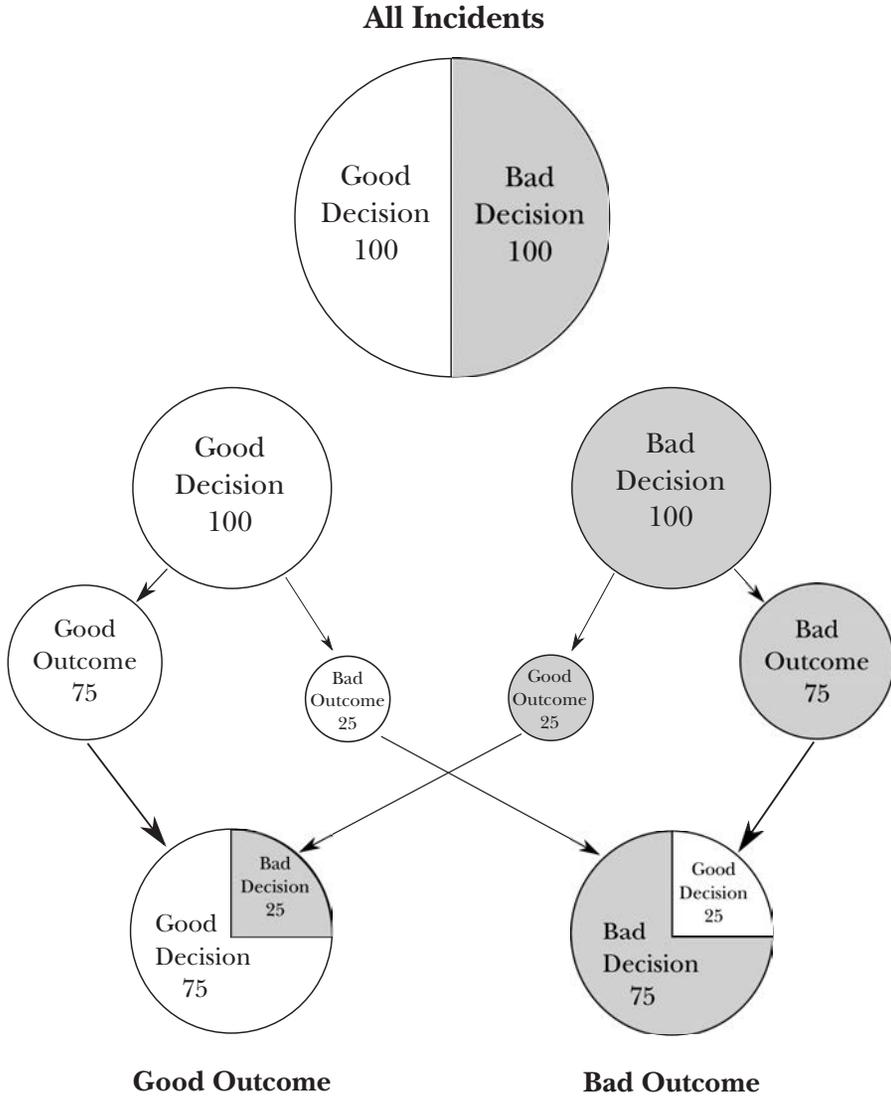
Figure 1 represents schematically how updating in light of hindsight evidence works—and why it makes sense—under a hypothetical scenario. The figure depicts 200 decisions in a world where half of the decisions are good and half are bad. A rational fact-finder initially believes there is a 50% chance the decision at issue was bad. In this hypothetical world, 75% of good decisions result in a good outcome, so of the 100 good decisions, 75 had a good outcome and 25 had a bad outcome. Similarly, 75% of bad decisions result in a bad outcome, so of the 100 bad decisions, 25 had a good outcome, and 75 had a bad outcome. Of the cases with a good outcome, then, 75 are from the “good decision” set, and 25 are from the “bad decision” set. A good outcome changes the initial 50% probability that the decision was bad to a 25% probability that the decision was bad. Same with the bad outcome cases: 75 of the 100 cases with a bad outcome involved a bad decision. So after observing a bad outcome, a rational fact-finder with perfect information about this world would believe there was a 75% probability the decision was bad.

62. See Sanchirico, *supra* note 17, at 1199 (noting occurrence of accident bears on whether or not defendant believed accident would occur if there is a correlation “between the defendant’s assessment of the chance of an accident and the actual chance of an accident”).

63. If a legal rule or standard is *not* designed to improve outcomes—if it advances some other interest, such as autonomy—outcome may not be indicative of compliance. In that case, hindsight evidence is irrelevant.

64. See *infra* Part V (presenting other examples in which outcome information is probative). Apply this thinking to a hypothetical in Mark Kelman, David E. Fallas & Hilary Folger, *Decomposing Hindsight Bias*, 16 *J. Risk & Uncertainty* 251, 259–60 (1998), which describes a weather forecaster who says there is an 80% probability it will rain on Saturday. It does not rain. “If a ‘factfinder’ were trying to ascertain whether the weatherman was ‘reasonable’, or made good probability estimates,” the authors write, “the fact that it [did not rain] should be of no moment.” *Id.* at 260. If the fact-finder knows, as they assume he or she does, that the weather pattern acted in a way that exceeds all contemporary meteorological understanding, this is true—as a reasonable weatherman would not be more likely to be correct than an unreasonable weatherman. But in the more likely event that the fact-finder does not know what caused the sunny day, the outcome may indeed be of some moment.

FIGURE 1: BAYESIAN UPDATING IN LIGHT OF HINDSIGHT EVIDENCE



Of course, actual jurors will not go through any of these precise mathematical calculations. Rarely, if ever, will jurors have subjective probability assessments that match precisely the proportion of decisions that are bad or the probability of a bad outcome given a good or bad decision. And these precise proportions are only illustrative, in any event:

The circumstances surrounding any individual litigation are unique.⁶⁵ The “proportion” of instances in which the hindsight evidence would arise is the product of an imagined counterfactual, not objective reality.⁶⁶

In addition, actual jurors are far from perfect Bayesians: They do not perform difficult mathematical updating, they judge using heuristics and biases that subject them to cognitive traps, and they use evidence to construct narratives rather than evaluate the evidence piecemeal.⁶⁷ But thinking about evidentiary issues in terms of Bayesian reasoning has two

65. See Alex Stein, *Bayesian Skepticism Justified*, 1 *Int'l J. Evidence & Proof* 339, 341 (1997) (“[O]bjective frequentist data are normally not available in litigation, which usually arises out of unique events.”).

66. The uniqueness of each case is the source of the “reference class problem,” a critique of mathematical models of evidence. See Ronald J. Allen & Michael S. Pardo, *The Problematic Value of Mathematical Models of Evidence*, 36 *J. Legal Stud.* 107, 109–16 (2007) (discussing reference-class problem); Edward K. Cheng, *A Practical Solution to the Reference Class Problem*, 109 *Colum. L. Rev.* 2081, 2083 (2009) (setting up problem before proposing a solution). Critics note that determining a prior probability of a fact and a likelihood ratio requires placing the case at hand in a class of cases. But each case has many characteristics. For example, if fact-finders want evidence on the prior odds of a broker giving fraudulent advice, do they consider all brokers, all American brokers, those in big cities, those in the northeast, or just those in New York City? Do they consider both men and women? Brokers of all ages? Those obtained by any means, by responding to advertisements, or just by responding to subway advertisements? And so forth. The only objectively reasonable reference class is a class of one: the case at issue. It is impossible to get that data. But when no numerical data is available anyway, there is little reason to believe jurors would do better with a perfectly specified reference class than with just the facts at hand. In either event, they will rely on their own life experiences and understanding of the world to evaluate—likely in a heuristic, non-numeric way—the base rate of the relevant behavior and the probative value of the evidence.

67. See Kenworthy Bilz, *We Don't Want to Hear It: Psychology, Literature and the Narrative Model of Judging*, 2010 *U. Ill. L. Rev.* 429, 435 & n.24 (citing Daniel Kahneman and Amos Tversky's research on heuristics and biases); Richard O. Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 *B.U. L. Rev.* 439, 448 (1986) [hereinafter Lempert, *New Evidence*] (noting research shows people are “at best inconsistent” Bayesians); Zabell, *supra* note 50, at 263 (discussing how human behavior diverges from Bayesian model). Professors Nancy Pennington and Reid Hastie have demonstrated that instead of incorporating evidence piecemeal, jurors construct coherent narratives to account for the evidence presented. See Nancy Pennington & Reid Hastie, *The Story Model for Juror Decision Making*, in *Inside the Juror*, *supra* note 45, at 192, 192–99 (summarizing ten years of research into story model). Professor Dan Simon's work on coherence-based reasoning indicates that as jurors develop conclusions from evidence, those conclusions feed back into their understanding of the evidence and reshape their interpretation of it. See Dan Simon, *In Doubt: The Psychology of the Criminal Justice Process* 174–76 (2012) (discussing coherence-based reasoning).

In addition, numerically precise Bayesian reasoning runs up against several logical paradoxes, most famously the “conjunction paradox”: Proof of each element of a claim by a preponderance of the evidence does not entail proof of the entire claim by a preponderance of the evidence when the probabilities of the independent elements are multiplied together. See Edward K. Cheng, *Reconceptualizing the Burden of Proof*, 122 *Yale L.J.* 1254, 1256–57 (2013) [hereinafter Cheng, *Burden of Proof*] (discussing conjunction paradox).

main virtues. First, it has normative value.⁶⁸ Bayesian updating is, in an important sense, rational decisionmaking.⁶⁹ Denying fact-finders access to evidence that would allow them to update rationally, without a good reason to do so, is troublesome and contravenes the rule and norm that relevant evidence is generally admissible. Second, it operates as a heuristic guide to decisionmaking for lawyers and judges as well as jurors.⁷⁰ Judges, in particular, can think in Bayesian terms when gauging the probative value of evidence in relation to prejudice. And even if jurors are imprecise with their estimates and vulnerable to cognitive traps, they may reasonably learn of an outcome and in some coarser manner, revise upward their assessment that the outcome was likely or the defendant was negligent.⁷¹

But even the most ardent “Bayesian-skeptic” should welcome hindsight evidence, as it is also relevant under the rival theory of reasoning from courtroom evidence: inference to the best explanation.

2. *Inference to the Best Explanation.* — Whereas Bayesian updating is an idealized, mathematical model of decisionmaking, inference to the best explanation is a nonprobabilistic, psychologically realistic description of human reasoning.⁷² It aligns with the empirically supported “story telling model” of juror decisionmaking, whereby jurors construct narratives to account for the evidence presented at trial.⁷³ And it mirrors how trials proceed in practice: The two sides vie for the jury’s verdict by presenting

68. See Lempert, *Modeling Relevance*, supra note 49, at 1056 (“The Bayesian model is normative—it specifies the way in which jurors are expected to evaluate evidence . . .”); Zabell, supra note 50 at 264 (stating Bayesian reasoning is useful as “a model of normative consistency”).

69. Lempert, *Modeling Relevance*, supra note 49, at 1023 (“This [Bayesian] formula describes the way knowledge of a new item of evidence (E) would influence a completely rational decision maker’s evaluation of the odds that a defendant is guilty (G).”).

70. See Ronald J. Allen, *A Reconceptualization of Civil Trials*, 66 B.U. L. Rev. 401, 402 (1986) (“It is becoming increasingly obvious, for example, that Bayesian approaches can best be used heuristically as guides to rational thought and not as specific blueprints for forensic decisionmaking.”); Lempert, *Modeling Relevance*, supra note 49, at 1056 (“Although I do not think that jurors in actual cases should be urged to use Bayesian calculations in evaluating evidence, nevertheless, I believe the approach taken here is of more than strictly academic interest.”); Richard A. Posner, *An Economic Approach to the Law of Evidence*, 51 Stan. L. Rev. 1477, 1479 (1999) [hereinafter Posner, *Economic Approach*] (“The value of Bayes’ theorem in the law of evidence is heuristic.”).

71. See Friedman, *Bayesian Methods*, supra note 51, at 4 (“Bayesian thinking does not require numerification. We all think in terms of magnitude without necessarily assigning numbers; ‘unlikely,’ for example, is a non-numerical expression of probability.”).

72. See Pardo & Allen, *Juridical Proof*, supra note 47, at 225 (contrasting inference to the best explanation with probabilistic approaches).

73. See generally Pennington & Hastie, supra note 67, at 192 (explaining and discussing empirical support for story model).

competing accounts of what happened, each insisting that the evidence supports its own story better than the other side's.⁷⁴

In this schema, lawyers generate potential explanations of the evidence and jurors select the best one based on criteria of consistency, simplicity, explanatory power, coherence with the fact-finder's background beliefs, internal coherence, and any other epistemic criteria that reasonable people use to judge the plausibility of an explanation.⁷⁵ This process of selecting explanations by comparing relative plausibility replicates much of "commonsense logic."⁷⁶ Under this framework, evidence is relevant "if it is explained by the particular explanation offered by the party offering the evidence, which in turn justifies that explanation as correct, assuming the explanation concerns a fact that matters to the substantive law."⁷⁷

This understanding of relevance is similar to the Bayesian version, but there are some differences. For one, Bayesian reasoning conceptualizes the likelihood that a proposition is true quantitatively—in terms of numerical probabilities—whereas inference to the best explanation understands relative plausibility to be a purely qualitative assessment.⁷⁸ Also, Bayesian updating incorporates evidence piecemeal: Each piece of evidence has a likelihood ratio that changes the fact-finder's probability assessment; by contrast, people engaging in inference to the best explanation evaluate evidence holistically, asking, "Which story better explains this collection of evidence?"⁷⁹ The relevance of a piece of evidence can be understood as the change in that holistic assessment after the evidence comes in.

Hindsight evidence is probative under this view, then, to the extent that one party's account is a better explanation of the outcome than the other party's. Again, when the fact in question is the objective *ex ante*

74. See Allen & Stein, *supra* note 47, at 568–69 (noting fact-finders compare competing stories at trial); Pardo & Allen, *Juridical Proof*, *supra* note 47, at 234 (discussing structure of proof in trials).

75. Pardo & Allen, *Juridical Proof*, *supra* note 47, at 229–30; see also Allen & Stein, *supra* note 47, at 568 ("[T]he variables that inform judgments of plausibility are all the things that convince people that some story may be true . . ."); Larry Laudan, *Strange Bedfellows: Inference to the Best Explanation and the Criminal Standard of Proof*, 11 *Int'l J. Evidence & Proof* 292, 296 n.13 (2007) ("Other familiar proposals for the explanatory virtues include plausibility, economy, loveliness (sic), probability, epistemic responsibility (that is, the earnest search for alternatives), high posterior probability, intuitiveness, and the refutation of alternatives."). Jurors may also come up with their own explanations for the evidence.

76. John R. Josephson, *On the Proof Dynamics of Inference to the Best Explanation*, 22 *Cardozo L. Rev.* 1621, 1623 (2001).

77. Pardo & Allen, *Juridical Proof*, *supra* note 47, at 241–42.

78. See Allen & Stein, *supra* note 47, at 572 (describing plausibility criteria as "qualitative rather than quantitative").

79. See *id.* at 576 (stating natural reasoning allows for "a global (or holistic) assessment of evidence").

probability of the outcome materializing, evidence that the outcome did in fact materialize is generally relevant. A high probability of the outcome, stemming from some strong *ex ante* disposition,⁸⁰ is a better explanation of the outcome than a low probability stemming from a weak disposition. If one party submits that a toaster was dangerous and the other submits it was not dangerous, calling an explosion a freak occurrence,⁸¹ the “explosive-disposition” explanation better explains the explosion than the “rare-but-possible-event” explanation. The power of the explanation may be weakened in light of other evidence, of course. But without any other information, jurors can reasonably incorporate the outcome into their inferential process.

When hindsight evidence is used to judge an actor’s decision in retrospect, it will be probative to the extent that a narrative depicting an unreasonable decision is a better explanation for the outcome than a narrative depicting a reasonable decision.⁸² Most fact-finders understand from experience that bad decisions tend to have bad consequences to a greater degree than good decisions have bad consequences. The world looks something like Figure 1. And this tendency is not a mere correlation,⁸³ a bad decision *creates* and so *explains* an increased likelihood of a bad outcome. A doctor who makes a bad medical decision—by ignoring literature that contraindicates her proposed treatment, going against established medical opinion, or examining a patient with insufficient care—creates a higher risk of a bad outcome. The story advanced by a plaintiff in a medical malpractice case, then, explains the outcome. The outcome, in turn, tends to prove the plaintiff’s story.

This explanatory structure becomes more complicated when the actor’s decision does not *cause* the outcome to materialize. Such is the case when the “outcome” is the revelation of a state of affairs that existed at the time the actor made the decision, such as the presence or absence of contraband in a suspect’s car when a police officer decided to search it or an occurrence that the actor predicted but did not cause, such as a natural disaster. In those cases, the quality of the actor’s decision explains

80. The notion of a “disposition” is highly contested in philosophical literature, see generally Sungho Choi & Michael Fara, Dispositions, *Stan. Encyclopedia Phil.* (2012), <http://plato.stanford.edu/entries/dispositions> (on file with the *Columbia Law Review*) (summarizing scholarly debate), but a dictionary definition should suffice here: “the tendency of something to act in a particular way under specific circumstances.” Disposition, Merriam-Webster (3d new international unabridged ed. 2015), <http://unabridged.merriam-webster.com/unabridged/disposition> [<http://perma.cc/56YX-HDJR>].

81. See *supra* notes 58–60 and accompanying text (discussing exploding toaster example).

82. If the standard for the cause of action is something else, such as intent or fraud, substitute that standard for “reasonableness.”

83. See Gilbert H. Harman, *The Inference to the Best Explanation*, 74 *Phil. Rev.* 88, 93–94 (1965) (distinguishing inference to the best explanation from inductive reasoning relying on mere correlation independent of explanatory relationship between correlated occurrences).

the *co-occurrence* of the decision and the result.⁸⁴ In the case of the search, a bad decisionmaking process—no probable cause⁸⁵—accounts for a decision to search the vehicle, accompanied by the absence of contraband. Similarly, a bad decisionmaking process explains a decision not to take an antidisaster precaution, accompanied by a natural disaster.

The probative value of hindsight evidence is similar under the Bayesian framework and the inference to the best explanation framework. This should not be surprising: “[T]here is no reason to see explanatory and Bayesian approaches as necessarily incompatible.”⁸⁶ The likelihood ratio is the relative probability of observing the evidence under two competing hypotheses. When the probability of observing the evidence under one hypothesis relates directly to how well that hypothesis explains the evidence, Bayesian probative value and relevance under inference to the best explanation are closely linked.⁸⁷ Such is the case with hindsight evidence: Bad decisions tend to produce bad outcomes. That is why the likelihood of observing a bad outcome given a bad decision is higher than the likelihood of observing a bad outcome given a good decision.

B. *Must Bear on an Issue of Fact*

After building the relevance case for hindsight evidence, it is time to break it down: to separate relevant hindsight evidence from irrelevant hindsight evidence. The previous section relied on the tendency of “good decisions” to produce good outcomes and “bad decisions” to produce bad outcomes. But these terms oversimplify the fact-finder’s inquiry. This section divides evaluation of a decision into two components: (1) determining the factual circumstances surrounding the decision based on evidence presented and (2) judging the decision given those factual circumstances. Only hindsight evidence bearing on the first component, the facts, is relevant under Rule 401. The section explores the line between fact and judgment and explains why certain determina-

84. See Fed. R. Evid. 104 (addressing “conditional relevance”).

85. See *United States v. Ross*, 456 U.S. 798, 809 (1982) (noting that for class of “searches of vehicles that are supported by probable cause . . . a search is not unreasonable if based on facts that would justify the issuance of a warrant, even though a warrant has not actually been obtained”).

86. Pardo & Allen, *Juridical Proof*, *supra* note 47, at 253.

87. See *id.* (“[E]xplanatory considerations drive inferences as to the likelihood of various potential explanations.”); see also Cheng, *Burden of Proof*, *supra* note 67, at 1258–72 (arguing Bayesian reasoning and inference to the best explanation are compatible because Bayes’s Theorem is well suited for comparative hypothesis testing, so it can be used to compare competing explanations for evidence). But see Allen & Stein, *supra* note 47, at 594–602 (criticizing Cheng’s reconciliation of these two understandings and claiming: “Mathematical probability is a system of reasoning that one must either use in its entirety or not use at all. There is no room for picking and choosing.”).

tions, including what constitutes reasonable professional judgment, concern factual circumstances that hindsight evidence can prove.⁸⁸

Under Rule 401, hindsight evidence—like any evidence—must make a “fact” more or less probable than it would otherwise be.⁸⁹ In many cases, there is little question that hindsight evidence meets this standard because the evidence bears on the concrete, factual circumstances that existed at the time of the events in question.

A jury evaluating an actor’s decision must determine what information was available to the actor at the time in order to decide whether the actor appropriately assessed and acted on that information.⁹⁰ If one version of the facts—the information available to the decision-maker at the time of the decision—was more likely to produce the outcome than another version, the outcome is probative of those facts. To take a simple example: A case turns on whether a police officer had probable cause to search a vehicle, and the officer claims he was justified because when he approached the vehicle, it smelled to him like burnt marijuana. The presence of burnt marijuana located during the search would tend to support the officer’s contention, whereas the absence of any burnt marijuana would tend to undermine it.⁹¹

Often, then, hindsight evidence will be most useful in evaluating witness credibility. If a witness testifies to a certain set of facts surrounding the decision, an outcome that tends to disprove those facts may reasonably cast doubt on the operation of any one of the witness’s testimonial capacities, “perception, memory, sincerity, [or] articulate-

88. This distinction does not apply, of course, when hindsight evidence is used to evaluate *objective* ex ante probability, as opposed to evaluating an actor’s conduct. The ex ante probability that an event outcome would materialize—that the toaster would explode—is itself a fact. The fact-judgment distinction emerges only when hindsight evidence bears on decisionmaking.

89. Fed. R. Evid. 401; see also James B. Thayer, Presumptions and the Law of Evidence, 3 Harv. L. Rev. 141, 143 (1889) (“Evidence, then, is any matter of fact which is furnished to a legal tribunal otherwise than by reasoning, as the basis of inference in ascertaining some other matter of fact.”).

90. See Sanchirico, *supra* note 17, at 1197 (stating outcome information may be relevant when “the fact finder has reason to believe . . . that the defendant . . . was in a position to know about the likely results of her actions”); see also Baron & Hershey, *supra* note 19, at 569 (noting outcome information “may affect a judge’s beliefs about actor information”).

91. See, e.g., *Sherman v. Holt*, No. 6:12-CV-292 (ATB), 2013 WL 6506475, at *7 (N.D.N.Y. Dec. 12, 2013) (relying on evidence of a “roach” not necessarily warm to the touch to conclude “[a]lthough it is a close question, the court finds that there is an issue of fact as to whether plaintiff’s vehicle actually smelled of burned marijuana”). But cf. *United States v. Awolowo*, No. 3:09-CR-14-1, 2010 WL 55327, at *6 n.4 (E.D. Tenn. Jan. 4, 2010) (rejecting argument that there was no probable cause when officers who said they smelled burnt and “green” marijuana found no marijuana beyond some residue, crumbs, and seeds).

ness.”⁹² Particularly when the witness is the defendant, who has an incentive to distort the information available at the time of the decision, outcome information can help the fact-finder decide whether or not to credit the witness’s testimony.

Litigants may also use hindsight evidence to prove that a defendant possessed a certain mental state.⁹³ Nearly all crimes contain an element of *mens rea*,⁹⁴ and some civil causes of action require litigants to prove a mental state. In securities fraud cases, for example, the plaintiff must prove scienter, an actual “intention ‘to deceive, manipulate, or defraud.’”⁹⁵ Although this is a less concrete inquiry, mental state is a question of fact and can be proved by any of the ordinary means of proving facts, including direct and circumstantial evidence,⁹⁶ such as hindsight evidence. Hindsight evidence is probative of mental state when a culpable state of mind is more likely to produce the outcome observed than a blameless state of mind is. This may not always be the case. For an excellent marksman, yes, intent is a better explanation for shooting someone from a hundred feet away than recklessness or negligence. But for a terrible marksman, it may not be.

In the above examples, hindsight evidence clearly bears on a fact of interest. But when a jury evaluates a party’s actions, it also makes nonfactual determinations; namely, juries exercise their *judgment* in light of found facts. When jurors make a determination that has no causal relationship to an outcome—when with full knowledge of the surrounding facts they simply decide, in their lay judgment, what behavior is acceptable and what is unacceptable—hindsight evidence is irrelevant.

For example, jurors in a negligence case must decide whether the defendant exercised “reasonable care.” As the Restatement of Torts notes, this consists of two inquiries: What were the facts relating to the actor’s conduct? And given those facts, did the defendant exercise

92. Richard D. Friedman, *Route Analysis of Credibility and Hearsay*, 96 *Yale L.J.* 667, 685 (1987).

93. See *Farmer v. Brennan*, 511 U.S. 825, 839 (1994) (describing subjective tests as “focusing on what a defendant’s mental attitude actually was (or is), rather than what it should have been (or should be)”).

94. See Model Penal Code § 2.02(1) (Am. Law Inst. 1965) (“[A] person is not guilty of an offense unless he acted purposely, knowingly, recklessly or negligently, as the law may require, with respect to each material element of the offense.”).

95. *Tellabs, Inc. v. Makor Issues & Rights, Ltd.*, 551 U.S. 308, 313 (2007) (quoting *Ernst & Ernst v. Hochfelder*, 425 U.S. 185, 193 n.12 (1976)); see also 15 U.S.C. § 78j(b) (2012) (prohibiting use of “any manipulative or deceptive device” in contravention of SEC’s rules).

96. *United States v. Rajaratnam*, 719 F.3d 139, 153 (2d Cir. 2013).

reasonable care?⁹⁷ Reasonableness is often styled as a “question of fact,”⁹⁸ but the second portion of the reasonableness inquiry, whether the action was reasonable given the circumstances, calls for an exercise of the fact-finder’s judgment.⁹⁹ As commentators have noted, “[t]he reasonable person standard is an empty vessel that jurors fill with community norms.”¹⁰⁰ The law thereby delegates this part of the inquiry to the lay judgment of the jury, capitalizing on the jurors’ values and understandings of normal human cognition.¹⁰¹

The jury’s exercise in judgment can be understood to itself contain two components: First, the jurors determine whether the defendant’s apprehension of the risk was reasonable, given his circumstances. Second, they decide whether his action was reasonable, given that risk.¹⁰²

97. See Restatement (Third) of Torts: Liab. for Physical & Emotional Harm § 8 (Am. Law Inst. 2010) (providing that when reasonable minds can differ, jury determines “facts relating to the actor’s conduct” and “whether the conduct lacks reasonable care”); see also Restatement (Second) of Torts § 328C (Am. Law Inst. 1965) (noting jury determines facts and whether those facts conform to legal standard).

98. See, e.g., *Frausto v. Cooper Tire & Rubber Co.*, No. 3–12–0761, 2014 WL 581724, at *3 (M.D. Tenn. Feb. 13, 2014) (“[W]hat care is reasonable is a question of fact, to be determined by a jury.”); *Gordon v. Alaska Pac. Bancorporation*, 753 P.2d 721, 725 (Alaska 1988) (“Whether particular conduct is negligent or reasonable is a question of fact for the jury, if reasonable minds could draw different inferences from the evidence presented.”); *Ind. Dep’t of Transp. v. Howard*, 879 N.E.2d 1119, 1125 (Ind. Ct. App. 2008) (“Whether telling the contractor to add or replace traffic control devices without more is reasonable is a question of fact.”).

99. Restatement (Third) of Torts: Liab. for Physical & Emotional Harm § 8 cmt. b (“In all, reaching a decision on the negligence issue requires an exercise of judgment by the jury.”); see also Restatement (Second) of Torts § 328C cmt. b (“[T]he jury normally is expected to determine what the general standard of conduct would require in the particular case, and so to set a particular standard of its own within the general one.”). Scholars disagree about precisely what benchmark jurors should or do use to decide reasonableness. See, e.g., Heidi Li Feldman, *Prudence, Benevolence, and Negligence: Virtue Ethics and Tort Law*, 74 *Chi.-Kent L. Rev.* 1431, 1446–50 (2000) (discussing how jurors may use thought experiments to picture a fictional reasonable person’s behavior and ask whether it matches defendant’s behavior); Gregory C. Keating, *Reasonableness and Rationality in Negligence Theory*, 48 *Stan. L. Rev.* 311, 341–82 (1996) (using social contract theory to illuminate reasonableness); Patrick J. Kelley & Laurel A. Wendt, *What Judges Tell Juries About Negligence: A Review of Pattern Jury Instructions*, 77 *Chi.-Kent L. Rev.* 587, 622 (2001) (linking community public-safety conventions to reasonableness inquiry); Alan D. Miller & Ronen Perry, *The Reasonable Person*, 87 *N.Y.U. L. Rev.* 323, 370 (2012) (noting, but rejecting, theory of reasonableness based on observation of actual behavior); Catharine Wells, *Tort Law as Corrective Justice: A Pragmatic Justification for Jury Adjudication*, 88 *Mich. L. Rev.* 2348, 2402–13 (1990) (discussing how juries’ decisions about reasonableness stem from group norms).

100. Steven Hetcher, *Non-Utilitarian Negligence Norms and the Reasonable Person Standard*, 54 *Vand. L. Rev.* 863, 864 (2001).

101. See *id.* (emphasizing “role of non-utilitarian negligence norms”); Eugene Volokh, *Tort Law vs. Privacy*, 114 *Colum. L. Rev.* 879, 909 (2014) (discussing jury’s role in inquiry).

102. See Restatement (Third) of Torts: Liab. for Physical & Emotional Harm § 3 cmt. g (noting that in some cases, “what is foreseeable concerns what the actor ‘should have

The second component—determining whether the defendant acted reasonably, given a particular likelihood of harm and the magnitude of that harm—is an exercise of pure judgment. Since it is not a factual question that relates to the probability of the outcome materializing, hindsight evidence is not logically relevant to answering it. Say a mother seeks to remove her ex-husband as custodian of their children's funds, claiming he breached his duty to manage the money prudently.¹⁰³ She establishes that he took half the money to a casino and placed it all on black at the roulette table. The outcome of the bet—win or loss—should not bear on the fact-finder's assessment of prudence.¹⁰⁴ The risk of losing the money was 52.6%,¹⁰⁵ and the outcome is independent of whether the fact-finder considers that risk reasonable or unreasonable. By the same token, the normative judgment of prudence has no *explanatory* power concerning the result, since it assumes a known probability of harm. Observing loss (or gain), then, provides no logical indication of what standard the fact-finder should set for acceptable conduct. This example is unrealistically precise, of course—jurors will rarely have risk probabilities so neatly presented—but it serves to show that hindsight evidence does not bear on the normative element of a juror's task.

Some questions, however, do not fall so neatly on one side of the fact-judgment line. Such is the case with the first component of the jury's task in a negligence case: Given the factual circumstances of the event, was the defendant's apprehension of the likelihood (and likely magnitude) of harm reasonable?

This analysis could be understood as a pure question of judgment. Once the jury determines the objective reality that surrounded the actor at the time of his decision, it must draw on its own lay judgment to decide what information a reasonable person would extract from those circumstances. In this way, the jury operates as a *standard-setter*, and hindsight evidence cannot tell it what standard to set.

Or the analysis could be a factual determination: If we believe that jurors sitting in a courtroom are imperfectly equipped to simulate the perceptions of a reasonable actor on the scene—because actually being in the defendant's situation gives people a better ability to extract

known"); id. § 3 cmt. e (noting jury should balance foreseeable risk and benefit when determining negligence); see also Restatement (Second) of Torts § 328C cmt. b (describing jury's role).

103. See, e.g., 20 Pa. Stat. and Cons. Stat. Ann. § 5312(b) (West 2005) (establishing duty of care); *Sutliff v. Sutliff*, 489 A.2d 764 (Pa. Super. Ct. 1985).

104. In reality, this bet would likely be imprudent as a matter of law. See *Sartore v. Buder*, 759 P.2d 785, 787 (Colo. App. 1988) (noting speculative investments violate reasonable prudence standard).

105. See *The Wizard of Odds, Roulette: Introduction*, <http://wizardofodds.com/games/roulette/basics> [<http://perma.cc/4PUD-XAH9>] (last updated Mar. 9, 2015) (giving odds on roulette bets).

information¹⁰⁶—then jurors cannot *set* the standard for what information a reasonable actor would extract from the circumstances. Instead, knowing a reasonable standard of information gathering exists beyond their knowledge, the jurors must decide whether the defendant's conduct met or fell short of that standard. In that case, hindsight evidence can be relevant. A reasonable person makes more accurate assessments of her circumstances than an unreasonable person; a bad outcome, then, is probative of an *unreasonable* assessment of a low probability of harm, as opposed to a *reasonable* assessment. For example, if an assault defendant claiming self-defense says that the victim's facial expression and posture led her to reasonably believe he was going to stab her, even if we have a picture of the relevant moment, hindsight evidence—the presence of a knife—might bear on the reasonableness of her perceptions.

Whether an issue falls on the fact side or the judgment side will depend on both the substantive law and the specifics of the case. In a regular negligence case, courts are likely to commit the reasonableness determination entirely to the jury's lay judgment. Judges will likely reason that jurors are fully equipped to decide reasonableness based on their everyday experience;¹⁰⁷ to hold a defendant to a higher standard than jurors would otherwise apply is an inappropriate *ex post* adjustment of the reasonableness standard.¹⁰⁸ The decision may be more difficult where the actor has some level of skill, knowledge, or experience that a lay juror or judge does not have, and the court must decide whether to let evidence speak to what a reasonable person with that trained mind would perceive. But the decision is easy if the question is one on which courts *already* take evidence. In that case, courts have decided the question is one of fact, not pure judgment. If hindsight evidence bears on that determination, it is relevant.

For example, parties often introduce evidence to determine whether a defendant exercised appropriate *professional* judgment. The Restatement of Torts notes that adjudicators should account for superior skills or knowledge when determining if a person acted reasonably.¹⁰⁹

106. See generally, but don't spend too long looking at, Malcolm Gladwell, *Blink* (2007), which extolls the virtues of the "adaptive unconscious," and Richard A. Posner, *Blinkered*, *New Republic* (Jan. 24, 2005), <http://www.newrepublic.com/article/blinkered> [<http://perma.cc/56SC-B7CD>], which harshly criticizes Gladwell's book but notes multiple literatures "demonstrate the importance of unconscious cognition."

107. See generally *Pokora v. Wabash Ry. Co.*, 292 U.S. 98, 104 (1934) (discussing when determination of reasonableness should be turned over to the jury).

108. See *Stornelli v. U.S. Gypsum Co.*, 134 F.2d 461, 462 (2d Cir. 1943) (L. Hand, J.) ("[W]e think of that common-law duty as though it were imposed before the event"); Benjamin Ewing & Douglas A. Kysar, *Prods and Pleas: Limited Government in an Era of Unlimited Harm*, 121 *Yale L.J.* 350, 419–20 (2011) (discussing vague standards imposed on greenhouse gas emissions).

109. See Restatement (Third) of Torts: Liab. for Physical & Emotional Harm § 12 (Am. Law Inst. 2010) ("If an actor has skills or knowledge that exceed those possessed by most others, these skills or knowledge are circumstances to be taken into account in

Where the defendant is a professional, then, he is liable if he fails to use the skills and knowledge “normally possessed by members of that profession or trade in good standing in similar communities.”¹¹⁰ This standard has been applied to doctors, lawyers, architects, engineers, auditors, psychiatrists, educators, and others.¹¹¹ The range of acceptable practices, then, is determined by the members of the profession,¹¹² and litigants must educate lay jurors about professional norms and practices. Indeed, in malpractice and professional negligence cases, experts testify as to the standards of the profession and in some cases, to the ultimate issue of whether the defendant’s actions complied with those standards.¹¹³ The professional standard of care can be understood as a proxy for factual circumstances of the decision—the knowledge accumulated by the profession and skills imparted to professionals.

Hindsight evidence will often be relevant to questions of professional judgment when the question is whether a specific kind of action was acceptable within professional norms, rules, or standards. Relevance requires only the plausible assumption that actions acceptable to the profession are more likely to create good outcomes than actions that are unacceptable to members of the profession. If that is the case, a bad outcome tends to prove the professional’s action violated professional norms.

determining whether the actor has behaved as a reasonably careful person.”); Restatement (Second) of Torts § 290 cmt. f (Am. Law Inst. 1965) (“If the actor has special knowledge, he is required to utilize it . . .”).

110. Jill Wieber Lens, *The (Overlooked) Consequence of Easing the Prohibition of Expert Legal Testimony in Professional Negligence Claims*, 48 U. Louisville L. Rev. 53, 64–65 (2009) (quoting Restatement (Second) of Torts § 299A).

111. See *id.* at 65 & n.76 (reciting professions); see also Laurie S. Jamieson, Note, *Educational Malpractice: A Lesson in Professional Accountability*, 32 B.C. L. Rev. 899, 902 (1991) (discussing standard of care in educational malpractice actions).

112. See Claudia E. Haupt, *Professional Speech*, 125 Yale L.J. 1238, 1244 (2016) (“It is thus the knowledge community that determines the standard of care.”); Alex Stein, *Toward a Theory of Medical Malpractice*, 97 Iowa L. Rev. 1201, 1209 (2012) [hereinafter Stein, *Medical Malpractice*] (noting prevalent doctrine looks at norms established by profession).

113. See Stein, *Medical Malpractice*, *supra* note 112, at 1209–10 (“Oftentimes, experts disagree about the applicable professional norm, and when that happens, factfinders need to resolve the disagreement.”); see also Wilburn Brewer, Jr., *Expert Witness Testimony in Legal Malpractice Cases*, 45 S.C. L. Rev. 727, 730 (1994) (“Expert testimony is used to define the standard of care to which attorneys must conform by explaining how attorneys conduct their affairs under given circumstances . . . [T]he expert may also testify about the attorney’s breach of duty by offering an opinion that . . . the attorney deviated from the standard of care.”); Jamieson, *supra* note 111, at 950–51 (describing use of expert testimony in educational malpractice actions); cf. *Adams v. Lab. Corp. of Am.*, 760 F.3d 1322, 1340–41 (11th Cir. 2014) (Garza, J., concurring) (listing circuits that accept “standard-of-care expert testimony”).

Hindsight evidence will typically be *irrelevant*, however, where the professional standard is a product of cost-benefit analysis.¹¹⁴ If the component of cost-benefit analysis at issue is the probability of the outcome materializing, then hindsight evidence will indeed bear on that probability. But if the question that the fact-finder must resolve is the appropriate *balance* of cost and benefit according to the profession, hindsight evidence has little to say because the view that strikes the best balance is not necessarily the one that produces higher odds of a good outcome. If one treatment costs \$1,000,000 and has a 25% chance of saving the life of someone who would otherwise die and another treatment costs \$50,000 and has a 15% chance of saving that person's life, the outcome of either treatment cannot tell the court which choice struck the right balance.

The key here—inherent in Rule 401 but rarely discussed when a party seeks to admit hindsight evidence¹¹⁵—is that hindsight evidence must bear on an issue of fact. Each time a party seeks to admit hindsight evidence, the judge should ask herself what fact of consequence it makes more or less likely. If the evidence could affect only the adjudicator's task of judging the legality of the defendant's actions *given* the facts, then it is irrelevant and should not be admitted. But in the many cases where hindsight evidence *does* tend to prove one party's case—where the outcome is more likely to occur given that party's proposed facts—it is relevant and admissible, subject, of course, to Rule 403 balancing of unfair prejudice against probative value.

III. PROBATIVE VALUE

All relevant evidence is presumptively admissible.¹¹⁶ But when a party seeks to exclude hindsight evidence on the basis of unfair prejudice or other Rule 403 considerations, the judge's decision to admit or exclude depends in part on the *probative value* of the evidence.¹¹⁷ Under the Bayesian framework, the probative weight of a piece of evidence corresponds to its likelihood ratio; under the inference to the best explanation framework, evidence is probative to the extent that one proposed narrative better explains the evidence than another.¹¹⁸ The basic scheme for understanding the probative value of hindsight evidence, then, was set forth in the preceding sections: Hindsight evidence

114. See Stein, *Medical Malpractice*, supra note 112, at 1204–05 (noting courts are just as competent as doctors to resolve questions of cost-benefit optimization).

115. See *infra* Part V (discussing how courts have handled hindsight evidence in four different areas of law).

116. Fed. R. Evid. 402.

117. Fed. R. Evid. 403.

118. See *supra* sections II.A.1, II.B (discussing these theories of relevance and probative value).

is more probative to the extent that one set of facts is more likely to generate the outcome than another.

Probative weight differs from relevance, however, in that the value of a piece of evidence depends on what other evidence is available.¹¹⁹ The Supreme Court discussed this distinction in *Old Chief v. United States*: Whereas a piece of evidence is relevant even if the fact it tends to prove is not in dispute,¹²⁰ an assessment of probative value considers alternative forms of proof. If substitute pieces of evidence are more probative of the fact and less prejudicial (and so will be admitted), the judge discounts the value of the evidence in light of the alternatives.¹²¹ In other words, the probative value of a piece of evidence is its *marginal* probative value, given the other available evidence.¹²² Although *Old Chief* focuses on an evidentiary alternative that points in the *same* direction as the evidence in question, the reasoning applies as well to contradictory evidence: The weight of any piece of evidence should be evaluated in light of what we already know. On its own, witness testimony that “the red car ran the stop sign” might be very probative, but the witness testimony has little value if there is video evidence of the accident, independent of which car was actually at fault.¹²³ Hindsight evidence, then, is more probative when the facts it tends to prove are not proven by better, alternative evidence.¹²⁴ It fills the gaps left by more conclusive evidence.

This Part examines the probative value of hindsight evidence in three classes of cases. First, it discusses how a fact-finder should understand the probative value of this evidence when the hindsight evidence arises in every case of a certain type, like harm in negligence cases. Second, it looks at nonbinary hindsight evidence, where the probative value of the evidence depends on the outcome’s magnitude or characteristics. Finally, it addresses the relationship between hindsight evidence and the doctrine of *res ipsa loquitur*.

A. *Inevitable Hindsight Evidence and Prior Odds*

The probative weight of hindsight evidence is less apparent where the evidence is an element of the cause of action—where it is present in

119. See *Old Chief v. United States*, 519 U.S. 172, 182–84 (1997).

120. *Id.* at 182; see also Fed. R. Evid. 401 advisory committee’s notes on proposed rules.

121. *Old Chief*, 519 U.S. at 182–84 (explaining probative value “may be calculated by comparing evidentiary alternatives”).

122. See Kahan, *supra* note 29, at 1638–39 (discussing concept of marginal value).

123. This ranking of evidence is related, but not identical to, the epistemic “best evidence” principle. See Alex Stein, *Foundations of Evidence Law* 39–40 (2005) (explaining best evidence principle).

124. See Baron & Hershey, *supra* note 19, at 571 (“[I]t can often be appropriate to use outcome information to evaluate decision quality, especially . . . when it is necessary to judge decisions by their outcomes (as fallible as this may be) simply because there is little other useful information.”).

every case and so does not *distinguish* between cases. For example, every negligence case contains an alleged injury; otherwise the plaintiff could not establish standing or state a claim.¹²⁵ A reasonable juror could understand injury to be highly probative of negligence. The rate of harm in the wake of negligent actions may be much, much higher than the rate of harm following non-negligent actions. Should this evidence of injury, present in every single negligence case, carry so much weight? Further, Rule 401 provides that relevant evidence makes a fact “more or less probable than it would be without the evidence.”¹²⁶ In cases where the evidence is inevitable given the cause of action, it seems incoherent to speak of the fact “without the evidence.” Without the evidence, there would be no viable cause of action, and the parties would not be considering the relevance of the outcome. The hindsight evidence fails to make some cases stronger than others.

But this hindsight evidence *can* yield insight into which cases get to court in the first place, so it is helpful information for the fact-finder. The rate of negligence among actions that are the subject of a negligence suit *is*, no doubt, greater than the rate of negligence among all actions. To illustrate: A certain percentage of all the drivers who change lanes in a given day are negligent, but if this set is limited to the drivers of cars that collided with another car while changing lanes, a much larger proportion were negligent. The same goes for every other action with a potential to cause harm: The rate of negligence among the subset where injury occurred is almost certainly higher than in the general population. A rational, truthseeking juror called for a negligence case, then, would start with some initial assessment—some prior probability—of the rate of negligence in the world and would then adjust that assessment upon learning that the “harm” element of the cause of action is satisfied.¹²⁷

This suggestion conflicts with several scholars—including Judge Richard Posner and Professor Edward Cheng—who have suggested that the prior odds be fixed at a specific value that signifies unbiased decisionmaking. Judge Posner and Professor Cheng have maintained that, as a normative matter, unbiased adjudication requires that prior odds be set at one to one.¹²⁸ Other theorists have rejected this fiction of

125. John C.P. Goldberg & Benjamin C. Zipursky, *The Restatement (Third) and the Place of Duty in Negligence Law*, 54 *Vand. L. Rev.* 657, 658 (2001) (noting nearly every state adheres to four-element formula of duty, breach, causation, and injury).

126. Fed. R. Evid. 401(a).

127. See Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 *Harv. L. Rev.* 1329, 1360 n.100 (1971) (asserting evidence that defines a class of cases may affect the prior probability of litigated propositions).

128. Cheng, *Burden of Proof*, *supra* note 67, at 1267 (“In civil trials, the prior probabilities as a normative matter should arguably be equal.”); Posner, *Economic Approach*, *supra* note 70, at 1514 (“Ideally we want the trier of fact to work from prior odds of 1 to 1 that the plaintiff or prosecutor has a meritorious case.”); see also Claude Fluet, *Liability Rules Under Evidentiary Uncertainty*, 30 *Int’l Rev. L. & Econ.* 1, 7 & n.23 (2010) (“I assume that the court has no pre-trial bias toward either the plaintiff’s claim or

fixed prior odds and suggested that the prior probability depends on actual odds of wrongdoing. Professor Richard Friedman has argued that one-to-one prior odds contravene the presumption of innocence in criminal cases because they are large relative to the odds a random person who has not been charged with the crime actually committed it.¹²⁹ He would set the prior odds at the likelihood a passerby on the street had committed a crime of the type charged around the time of the crime.¹³⁰ Similarly, Professor Louis Kaplow contends that under the strongest version of the conventional conception of burdens of proof, prior odds must account for the relative “frequencies of harmful and benign acts coming before the tribunal.”¹³¹ Professor Kaplow’s version of the preponderance standard would, then, depend on the type of case at hand.¹³²

The idea of fixed, even prior odds has normative appeal: It resonates with the ideal of equality before the law and the image of the scales of justice that tilt only with evidence.¹³³ And it avoids the discomfort of “[t]reating proof of one element of a negligence case ([injury]) as if it were proof of another distinct and separate element ([breach of duty]).”¹³⁴ But for purposes of rational truthseeking, beginning a case with artificially even odds makes little sense.¹³⁵ If the burden of proof in a

the defendant’s counterclaim, i.e. before hearing the evidence the court assigns probability one half to either party being right.”); Jason S. Johnston, *Bayesian Fact-Finding and Efficiency: Toward an Economic Theory of Liability Under Uncertainty*, 61 *S. Cal. L. Rev.* 137, 149 (1987) (“One way to formalize such freedom from pre-trial predilections is to impose the assumption that the juror’s pre-trial beliefs are equally balanced between guilt and innocence.”); David S. Schwartz, *A Foundation Theory of Evidence*, 100 *Geo. L.J.* 95, 149 (2011) (“A much more defensible prior probability is 0.5. Juries are uniformly required and instructed to approach cases without a predisposition in favor of either side.”).

129. Richard D. Friedman, *A Presumption of Innocence, Not of Even Odds*, 52 *Stan. L. Rev.* 873, 880–81 (2000) [hereinafter Friedman, *Presumption of Innocence*].

130. *Id.* at 881–82.

131. Louis Kaplow, *Burden of Proof*, 121 *Yale L.J.* 738, 778 (2012) [hereinafter Kaplow, *Burden of Proof*]. Professor Kaplow’s article ultimately argues for an unconventional conception under which courts should set the burden of proof to maximize social welfare, and the optimal evidence threshold may be completely different from what it is under the preponderance standard. *Id.* at 146–47. For thoughtful criticism of Professor Kaplow’s thesis, see Allen & Stein, *supra* note 47, at 579–93, and Edward K. Cheng & Michael S. Pardo, *Accuracy, Optimality and the Preponderance Standard*, 14 *Law, Probability & Risk* 193, 195–201 (2015).

132. See Kaplow, *Burden of Proof*, *supra* note 131, at 797 n.100 (discussing factors that may go into priors).

133. See Friedman, *Presumption of Innocence*, *supra* note 129, at 886–87 (“The icon of equal scales suggests symmetry and balance; presumably the scales will tilt one way or the other at the end of the case, but fairness may seem to demand that at the beginning they be on the same level.”).

134. Kenneth S. Abraham, *Self-Proving Causation*, 99 *Va. L. Rev.* 1811, 1812 (2013) (discussing negligence cases in which breach proves causation).

135. If the jurors have no knowledge of base rates, starting from even odds is ideal from a truthseeking perspective. See Cheng & Pardo, *supra* note 131, at 201–07

civil case is understood to impose an equal risk of mistaken loss on plaintiffs and defendants,¹³⁶ the jury must be free to account for all information that will help them impose an accurate judgment. That includes their lay understanding of base rates of wrongdoing and base rates of wrongdoing given outcome information. Professor Friedman's proposal for criminal cases, then, where the fact-finder starts with the base rate of wrongdoing in the population and adjusts this as he receives evidence—including the hindsight evidence that is a part of every case—is a reasonable mode of analysis for both civil and criminal cases.¹³⁷

B. *Nonbinary Hindsight Evidence*

Thus far I have discussed the probative weight of binary hindsight evidence: either there was harm or there was not, the investment did well or it did badly, the search uncovered drugs or it did not. But hindsight evidence is not limited to these dichotomies. Two features of hindsight

(demonstrating traditional preponderance of the evidence standard is optimal under these circumstances).

136. See Lempert, *New Evidence*, *supra* note 67, at 463 (“Although a civil jury should not favor one side over another, this is, in terms of my model of relevance, a matter captured by a regret matrix and not by Bayes’s Theorem.”).

137. Professor Kaplow’s proposal, that the jury account for the rates at which these cases arrive *in court*, poses prohibitive practical difficulties. This knowledge cannot be gleaned from a juror’s life experience. The only way to ascertain it is to survey prior verdicts, which jurors will not do.

Allowing the jury to rationally incorporate inevitable hindsight evidence into its prior odds will not work if a selection process artificially alters this base rate. This is the case, for example, in challenges to patent validity based on obviousness. Of all the conceivable inventions in the world, inventions that have actually or constructively been reduced to practice are almost certainly more likely to be obvious than inventions that have not been invented, since people are more likely to devise things that are easy to invent. And in every patent infringement case, the plaintiff has actually invented something. If any invention at all could be patented, a rational juror would take the fact of invention into account. However, the Patent and Trademark Office provides a screening function, issuing a patent only if it determines that the invention was not obvious in light of the prior art. See U.S. Patent & Trademark Office, *Manual of Patent Examining Procedure* § 2141 (9th ed. 2015), <http://www.uspto.gov/web/offices/pac/mpep/s2141.html> [<http://perma.cc/76VC-3Z4W>] (discussing criteria for determining non-obviousness). It is sensible, then, for judges to attempt to eliminate the fact of invention from consideration by instructing jurors to look to more objective “secondary” considerations, such as “commercial success, long felt but unsolved needs, [and the] failure of others,” when judging obviousness. *Graham v. John Deere Co.*, 383 U.S. 1, 17 (1966).

Patent examiners themselves, however, *should* incorporate the fact of invention into their prior odds that any invention is obvious. Inventors and their employers may of course forgo filing patent applications on clearly obvious inventions in order to avoid wasting time and money. Cf. Ted Sichelman & Stuart J.H. Graham, *Patenting by Entrepreneurs: An Empirical Study*, 17 *Mich. Telecomm. & Tech. L. Rev.* 111, 132–34 (2010) (noting that costs of patenting and mistaken beliefs about obviousness may discourage inventors from filing for patents). But the group of inventions that comes before a patent examiner is still likely to be disproportionately obvious, relative to all possible inventions.

evidence—magnitude and specific type of outcome—may bear on its probative weight.

First, the *magnitude* of harm may be probative. Cases with relatively worse outcomes are typically more likely to involve culpable conduct than cases with relatively less bad outcomes.¹³⁸ This conclusion follows from the simple idea that acceptable conduct tends to lead to better outcomes than unacceptable conduct, even looking only at cases where the outcome was bad. Compare three investment advisers, each of whom recommends a “safe” investment, per the client’s request. One investment loses 2% of its value, one loses 40% of its value, and the third loses 99% of its value. Competent management may frequently lead to small losses, although negligent management may lead to them with slightly greater frequency, so the 2% loss is minimally probative of negligence. But competent management is less likely to lead to large losses. And even negligent management may be very unlikely to cause complete loss; instead, complete loss may indicate that management was fraudulent. So even though the person suing his or her investment adviser must always prove a loss,¹³⁹ the magnitude of loss may be additionally probative of malfeasance.

In addition, certain *types* of negative outcomes—ways that the bad result came about and the sorts of negative results caused—can be more indicative of culpability than others. When culpable behavior leads to a certain type of result at a much higher rate than acceptable behavior, that type of result is more probative of culpability than a different type that arises from acceptable and unacceptable behavior comparably. The outcome more probative of culpability may be more severe than a qualitatively different outcome less probative of culpability—but it may not be. For example, every medical malpractice case involves injury to the plaintiff.¹⁴⁰ But discomfort caused by a sponge left in the patient’s abdomen during surgery may be more probative of malpractice¹⁴¹ than, say, death from a heart attack following surgery.¹⁴² It is extremely unlikely

138. See Baron & Hershey, *supra* note 19, at 570 (“As a general rule, it makes sense to punish actors more severely for more severe consequences; it is usually difficult to know what the actor knew, and severity of consequences is a clue as to the degree of negligence.”).

139. See, e.g., *Brown v. E.F. Hutton Grp., Inc.*, 991 F.2d 1020, 1031 (2d Cir. 1993) (noting plaintiffs must prove “detriment” in § 10(b) unsuitability claims); *Cont’l Life Ins. Co. v. Shearson Lehman Hutton, Inc.*, No. 88-9279, 1992 WL 6750, at *3 (E.D. Pa. Jan. 14, 1992) (observing negligence claims against brokers require proof of damages).

140. Kathleen Flynn Peterson, *Litigating Tort Cases* § 61:21 (2014) (listing elements of medical negligence claims).

141. Restatement (Second) of Torts § 328D cmt. d (Am. Law Inst. 1965) (“[T]here are other kinds of medical malpractice, as where a sponge is left in the plaintiff’s abdomen after an operation, where no expert is needed to tell the jury that such events do not usually occur in the absence of negligence.”).

142. See generally Giora Landesberg et al., *Perioperative Myocardial Infarction*, 119 *Circulation* 2936 (2009) (discussing incidence of heart attacks following surgery).

that a non-negligent surgical team will leave a sponge in a patient's abdomen, whereas heart attacks may result from negligent and non-negligent treatment at more similar rates. Thus the (potentially) less severe outcome is more probative of malpractice.

C. *Res Ipsa Loquitur*

The observation that certain types of outcomes may be much more probative of unacceptable action than other types underlies the doctrine of *res ipsa loquitur*. This 150-year-old doctrine,¹⁴³ as outlined in the Restatement (Third) of Torts, permits a finding of negligence when “the accident causing the plaintiff’s harm is a type of accident that ordinarily happens as a result of the negligence of a class of actors of which the defendant is the relevant member.”¹⁴⁴ It allows an inference of negligence where the plaintiff has neither alleged nor proven *specific* negligence.¹⁴⁵ The doctrine has been applied, to give a few examples, when a surgical team leaves a sponge or surgical instrument in a patient’s body,¹⁴⁶ when a car runs off an empty road,¹⁴⁷ and when a vehicle hits escaped livestock on a public street.¹⁴⁸ A jury may find the doctor, driver, or livestock owner liable without knowing anything more about the events in question.

The *res ipsa* doctrine is in part justified by this Essay’s framework. The probative value of an outcome, in a negligence case, is the probability of observing that outcome given negligent behavior divided by the probability of the outcome given non-negligent behavior. *Res ipsa* allows a jury to find liability when the probative value of hindsight evidence is so great that the likelihood of negligence, given the outcome, is high—sufficiently high to satisfy the plaintiff’s burden of production.¹⁴⁹ From an inference to the best explanation perspective, the doctrine permits a finding for the plaintiff when any explanation of the hindsight

143. See *Byrne v. Boadle* (1863) 159 Eng. Rep. 299; 2 H. & C. 722 (first recognizing doctrine of *res ipsa loquitur*).

144. Restatement (Third) of Torts: Liab. for Physical & Emotional Harm § 17 (Am. Law Inst. 2010).

145. Meiring de Villiers, *Virus Ex Machina: Res Ipsa Loquitur*, 2003 Stan. Tech. L. Rev. 1, 48.

146. See, e.g., *Baumgardner v. Yusuf*, 51 Cal. Rptr. 3d 277, 281–84 (Ct. App. 2006) (discussing *res ipsa* in several medical cases); *Haddock v. Arnspiger*, 793 S.W.2d 948, 951 (Tex. 1990) (noting *res ipsa* is widely applied in these contexts). But see Seinfeld: The Junior Mint (NBC television broadcast Mar. 18, 1993).

147. See, e.g., *Eaton v. Eaton*, 575 A.2d 858, 864–65 (N.J. 1990) (applying *res ipsa* to “one-car accident” case).

148. See, e.g., *Roberts v. Weber & Sons, Co.*, 533 N.W.2d 664, 669 (Neb. 1995) (“There are certain factual situations, as evidenced by the case at bar, wherein livestock ordinarily would not escape onto a public highway in the absence of some negligence.”).

149. See generally David Kaye, *Probability Theory Meets Res Ipsa Loquitur*, 77 Mich. L. Rev. 1456 (1979) [hereinafter Kaye, *Res Ipsa*] (discussing how high likelihood ratio must be for *res ipsa* to apply).

evidence favoring the defendant is sufficiently implausible, relative to explanations favoring the plaintiff. In this way, if judges admitted and juries considered hindsight evidence in accordance with the framework of this Essay, the need for a separate doctrine of *res ipsa loquitur* would be greatly diminished: Juries rationally considering the probative value of hindsight evidence in these extreme situations might find the plaintiff's position more probable based on the hindsight evidence alone.¹⁵⁰

But the doctrine did not originate with this understanding in mind, and the true origins of *res ipsa loquitur* suggest another benefit of introducing hindsight evidence. *Res ipsa* arose "to restore the appropriate balance between parties when one party has access to information that the other lacks."¹⁵¹ In the pre-discovery era, when plaintiffs had no way to access defendants' evidence concerning their own wrongdoing, to require the plaintiff to carry the burden in every negligence case would have created systematic pro-defendant biases.¹⁵² By putting pressure on the defendant to produce this evidence, courts improved accuracy. Thus, in earlier cases, courts would apply *res ipsa* only if there was no other evidence bearing on the events in question available to the plaintiff.¹⁵³ In the contemporary era of ample discovery, most courts applying the doctrine do not require the defendant to have superior access to information.¹⁵⁴

Still, the logic of the earlier era holds: Hindsight evidence will tend to prompt the production of better evidence. If the defendant has access to evidence bearing on her liability and the plaintiff does not, and if the plaintiff introduces hindsight evidence in his favor, the defendant will be prompted to produce evidence exonerating herself. Conversely, where the party *introducing* hindsight evidence should have better evidence, the fact-finder will consider his failure to introduce better evidence when evaluating the strength of his case.¹⁵⁵ A plaintiff with better evidence will be reluctant to rely on hindsight evidence alone. As in the specific case of *res ipsa*, a willingness to consider hindsight evidence may "smoke out" better evidence.

150. *Res ipsa* also creates an exception to the usual requirement that the plaintiff prove a specific negligent act. Richard W. Wright, *Liability for Possible Wrongs: Causation, Statistical Probability, and the Burden of Proof*, 41 *Loy. L.A. L. Rev.* 1295, 1336 (2008). Hindsight evidence cannot substitute for this feature of the doctrine.

151. Ronald J. Allen, *How Presumptions Should Be Allocated: Burdens of Proof, Uncertainty, and Ambiguity in Modern Legal Discourse*, 17 *Harv. J.L. & Pub. Pol'y* 627, 631 (1994).

152. *Id.* at 632.

153. See Ariel Porat & Alex Stein, *Tort Liability Under Uncertainty* 84 (2001).

154. See Restatement (Third) of Torts: Liab. for Physical & Emotional Harm § 17 cmt. i (Am. Law Inst. 2010) ("The elements of *res ipsa loquitur* set forth in this Section do not include the defendant's superior access to information.").

155. See Kaye, *Res Ipsa*, supra note 149, at 1475 (discussing appropriate treatment of a party's failure to come forward with evidence).

In addition to these benefits, however, hindsight evidence has a number of drawbacks that create a risk of unfair prejudice. The next Part explores the trouble with hindsight evidence.

IV. PREJUDICE, BALANCING, AND ADMISSION

With hindsight evidence comes the potential for hindsight bias. But the potential for bias is not, in and of itself, a sufficient reason for exclusion. Instead, the Federal Rules of Evidence recognize that juries may overvalue certain pieces of evidence and so the Rules provide a rule for weighing probative value against prejudice. Under Rule 403, a judge may exclude evidence only 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.”¹⁵⁶ This Part first reviews the literature on hindsight bias and its cousin, outcome bias, then discusses the best methods available for debiasing the fact-finder in different classes of cases, and finally addresses how judges should think about probative value and prejudice when deciding whether to admit hindsight evidence under Rule 403.

A. *Hindsight and Outcome Biases*

Hindsight bias—sometimes known as Monday-morning quarterbacking or the I-knew-it-all-along effect—is the tendency “not only . . . to view what has happened as having been inevitable, but also view it as having appeared ‘relatively inevitable’ before it happened.”¹⁵⁷ In other words, when people learn of an outcome, they experience an “unjustified increase in its perceived probability, for it seems to have appeared more likely than it actually was.”¹⁵⁸

Hindsight bias comes in three varieties.¹⁵⁹ Primary hindsight bias occurs when learning of an outcome increases its perceived objective ex ante likelihood of materializing.¹⁶⁰ Secondary hindsight bias manifests when a person learns of an outcome and then unknowingly overestimates the probability she herself would have assigned to that outcome ex ante.¹⁶¹ Tertiary hindsight bias arises when a person learns of an outcome and adjusts her assessment of the foreseeability of that

156. Fed. R. Evid. 403.

157. Baruch Fischhoff, For Those Condemned to Study the Past: Reflections on Historical Judgment, 4 *New Directions for Methodology Soc. & Behav. Sci.* 79, 83 (1980).

158. Baruch Fischhoff, Hindsight ≠ Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty, 1 *J. Experimental Psych.: Hum. Perception & Performance* 288 (1975), reprinted in 12 *Quality & Safety Health Care* 304, 304 (2003) [hereinafter Fischhoff, Hindsight ≠ Foresight].

159. See Kelman et al., *supra* note 64, at 252–54.

160. *Id.* at 252.

161. *Id.* at 253.

outcome, so she faults others for failing to predict it.¹⁶² Secondary bias is always and everywhere irrational: The person inaccurately assesses what she would have known.¹⁶³ But as Part II demonstrated, primary bias has a rational hindsight-evidence counterpart: Information that an outcome materialized *can be* probative of objective ex ante likelihood of the outcome.¹⁶⁴ Tertiary bias does as well: When a fact-finder has incomplete factual information about the information available to a decisionmaker ex ante, the outcome of that decision is probative of what a reasonable person would have foreseen.¹⁶⁵ However, if outcome information causes a person to *overestimate* ex ante likelihood or foreseeability, then the person experiences true hindsight *bias*.

Outcome bias, “the tendency to base assessments of a decision’s quality on its consequences,”¹⁶⁶ is closely related to hindsight bias. But whereas hindsight bias operates through the adjustment of probability assessments, outcome bias operates independent of probabilities. Even when a person *knows* how likely or foreseeable an outcome was ex ante, outcome bias may cause her to think that a decision with a bad outcome was a bad decision.¹⁶⁷ To the extent a person is tasked with judging a decision from the ex ante perspective, outcome bias is always an undesirable force: It is irrelevant to factual questions and instead changes the legally acceptable standard of behavior based on its consequences. If hindsight evidence influences a fact-finder’s decision on a nonfactual question of pure judgment,¹⁶⁸ the fact-finder has fallen victim to outcome

162. *Id.* at 253–54; see also Roesse & Vohs, *supra* note 20, at 412 (dividing hindsight bias into “memory distortion, inevitability, and foreseeability,” where “memory distortion” occurs when, after learning of an outcome, a person misremembers what he or she predicted; “foreseeability” is the same as secondary hindsight bias; and “inevitability” covers both primary and tertiary bias).

163. Kelman et al., *supra* note 64, at 253.

164. See *supra* text accompanying notes 57–61.

165. See *supra* text accompanying notes 90–96, 106; see also Kelman et al., *supra* note 64, at 253–54 (discussing rational uses of outcomes). But see *id.* at 260 (suggesting outcomes should not influence determinations of reasonableness). Indeed, Kelman and his co-authors find that when people can calculate ex ante probability easily, they do not experience hindsight bias upon learning of an outcome. See *id.* at 258 (“What we know for sure based on this experiment is that at least in dealing with situations in which ex ante probabilities can be calculated with confidence, reporting an outcome’s occurrence does *not* alter its perceived probability.”).

166. Rachlinski, *Judging in Hindsight*, *supra* note 20, at 581.

167. See Philip G. Peters, Jr., *Hindsight Bias and Tort Liability: Avoiding Premature Conclusions*, 31 *Ariz. St. L.J.* 1277, 1282 (1999) (“Outcome bias, consequently, appears even when subjects are told in advance the probability of a bad outcome.”); Doron Teichman, *The Hindsight Bias and the Law in Hindsight*, in *The Oxford Handbook of Behavioral Economics and the Law* 354, 355 (Eyal Zamir & Doron Teichman eds., 2014) (“Whereas the central point in the hindsight bias is the ex post adjustment of probability assessments, the outcome bias focuses on the question whether outcomes influence the way in which people judge the quality of a decision.”).

168. See *supra* text accompanying notes 103–105 (giving example of playing roulette with child’s funds).

bias. Although outcome bias and hindsight bias are different, they frequently work together, and many studies do not distinguish between them. Instead, most research simply investigates “whether outcome information influences evaluative judgments.”¹⁶⁹

In a series of classic studies, researchers have found that lay legal decisionmakers experience hindsight bias and outcome bias.¹⁷⁰ Mock jurors demonstrated bias in determining whether a police search of a suspect’s home was lawful,¹⁷¹ whether a municipality took adequate precautions against natural disasters,¹⁷² whether a railroad was reckless when it continued to operate on a track that had been declared hazardous,¹⁷³ whether a therapist took appropriate efforts to prevent a dangerous psychiatric patient from harming others,¹⁷⁴ and whether an invention was an obvious solution to a technical problem.¹⁷⁵

The authors of two major meta-analyses of hindsight-bias research have characterized the bias as having a “small”¹⁷⁶ or “small to medium”¹⁷⁷ effect size, but that may be misleading: The 1991 meta-analysis concluded that if a decisionmaker performs an unfamiliar task, evaluating a case where the outcome actually occurred, “a maximum of 7–27% of the population may make different decisions because of the hindsight bias.”¹⁷⁸ This high-end estimate bears out in mock-juror research on relatively close cases. In two studies of close cases, about a third of the subjects shifted with hindsight, changing which side had the majority

169. Peters, *supra* note 167, at 1283.

170. Social scientists, starting with Professor Baruch Fischhoff, have been studying hindsight bias for forty years. See Roese & Vohs, *supra* note 20, at 411 (“Hindsight bias is one of the most widely studied of decision traps, having been featured in more than 800 scholarly papers.”). A Google Scholar search for “hindsight bias” yields approximately 13,000 results. Google Scholar Search for “Hindsight Bias,” Google Scholar, <http://scholar.google.com> (selecting “Articles” and searching for “hindsight bias,” excluding citations).

171. See Casper et al., *supra* note 28, at 299 tbl.2 (finding subjects were less likely to award compensatory damages if told police found evidence of criminal conduct).

172. See Kamin & Rachlinski, *supra* note 22, at 98 (reporting only 24% of subjects in foresight condition deemed anti-flood precaution necessary, but 57% of those in hindsight condition, where flooding occurred, said city should have taken precaution).

173. See Hastie et al., *Juror Judgments*, *supra* note 57, at 605 (reporting only one-third of mock jurors said train should not operate in foresight but two-thirds deemed railroad reckless after accident occurred).

174. See Susan J. LaBine & Gary LaBine, *Determinations of Negligence and the Hindsight Bias*, 20 *Law & Hum. Behav.* 501, 510 (1996) (“Overall, 24% of the respondents (n = 23) found the therapist negligent in the violent outcome conditions; only 6% (n = 6) and 9% (n = 9) found the therapist negligent in the no violence and outcome unspecified conditions, respectively.”).

175. Mandel, *supra* note 22, at 1406–09 (explaining design and results of study of hindsight bias in patent litigation).

176. Christensen-Szalanski & Willham, *supra* note 20, at 156.

177. Guilbault et al., *supra* note 20, at 109.

178. Christensen-Szalanski & Willham, *supra* note 20, at 162.

vote.¹⁷⁹ In cases where mock jurors largely agree in foresight, however, outcome information had much less of an effect.¹⁸⁰ The bias is unlikely to create close cases out of very easy cases.

Scholars have investigated whether judicial expertise serves to mitigate the hindsight bias. Results have been mixed. The power trio of Professor Chris Guthrie, Professor Jeffrey Rachlinski, and Magistrate Judge Andrew Wistrich has conducted a series of studies examining heuristics and biases in judicial decisionmaking. In one study, federal magistrate judges who learned of the outcome of an appeal overestimated how likely they were to predict that outcome *ex ante*.¹⁸¹ By contrast, their study of 900 state and federal judges found subjects were just as likely to find probable cause for a police search in foresight as in hindsight.¹⁸² In addition, when judges participated in the study referenced above with the potentially reckless railroad, the difference between hindsight and foresight was not statistically significant.¹⁸³ These studies together provide *some* reason to believe that judges are less susceptible to the hindsight bias—or are better able to overcome it—than lay jurors are. They do not suggest, however, that judges are completely immune. Judges are only human, after all.

B. *Debiasing*

Bias invites debiasing. For nearly forty years, researchers have tried to quash the effects of the hindsight bias and foster rational decisionmaking. This has proven a “thorny task.”¹⁸⁴ To date, no techni-

179. Kamin & Rachlinski, *supra* note 22, at 98 (observing only one-quarter of foresight-condition subjects concluded precaution was unnecessary but 57% of hindsight-condition subjects found defendant liable); see also Hastie et al., *Juror Judgments*, *supra* note 57, at 605 (noting 33% of subjects disapproved of railroad’s operations in foresight but 67% judged railroad reckless in hindsight). When jurors deliberate, the side with the majority on the first ballot typically prevails. Harry Kalven, Jr. & Hans Zeisel, *The American Jury* 488 (1966).

180. See, e.g., Casper et al., *supra* note 28, at 299 tbl.2 (reporting, in unlawful-search study, that 92% of subjects found for plaintiff in foresight or with “innocent” outcome but 83% still found for plaintiff when told police uncovered incriminating evidence); LaBine & LaBine, *supra* note 174, at 510 (reporting 9% of subjects found therapist negligent in foresight and 24% found her negligent when her patient committed violent acts).

181. Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *Blinking on the Bench: How Judges Decide Cases*, 93 *Cornell L. Rev.* 1, 24–26 (2007); see also Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *Inside the Judicial Mind*, 86 *Cornell L. Rev.* 777, 784–86, 799–805 (2001) (describing study of magistrate judges in depth).

182. See Jeffrey J. Rachlinski, Chris Guthrie & Andrew J. Wistrich, *Probable Cause, Probability, and Hindsight*, 8 *J. Empirical Legal Stud.* 72, 76 (2011) (“We . . . conclude that the hindsight bias . . . does not affect [judges’] legal rulings.”).

183. Reid Hastie & W. Kip Viscusi, *What Juries Can’t Do Well: The Jury’s Performance as a Risk Manager*, 40 *Ariz. L. Rev.* 901, 906 (1998). The authors note that the judicial discrepancy between hindsight and foresight—25% versus 15%—might assume significance with a larger sample. *Id.*

184. Teichman, *supra* note 167, at 364.

que functions perfectly and eliminates hindsight bias completely.¹⁸⁵ But there is a difference between “working perfectly and working at all,”¹⁸⁶ and debiasing techniques that “work at all” can help ameliorate prejudice. This section discusses three techniques that have emerged as most promising for mitigating hindsight bias at trial. First, a consider-the-opposite strategy, in which subjects are asked to think of and explain how alternative outcomes could have occurred. Second, an emphasis on consequence severity, whereby subjects are told their decision will seriously affect another person’s life. Third, bifurcation, whereby jurors decide on liability before hearing evidence of damages.¹⁸⁷ Implementing the consider-the-opposite and consequence-severity strategies may not require special instructions or procedures: Attorney argument naturally encourages jurors to consider the opposite, and jurors are aware of the serious consequences of their actions. When judges consider the prejudicial effect of hindsight evidence for purposes of determining admissibility, then, they should account for this debiasing. Bifurcation, on the other hand, is most suitable when the hindsight evidence is inadmissible *as* hindsight evidence but is admitted for another purpose.

1. *Consider the Opposite.* — The consider-the-opposite strategy stems from the strongest explanation of hindsight bias: When people learn of an outcome, they assimilate it with everything else they know about the story to create a coherent, causal account of what happened.¹⁸⁸ Pioneer hindsight-bias researcher Professor Baruch Fischhoff proposed this theory in 1975, dubbing it “creeping determinism.”¹⁸⁹ Subsequent research has supported this view, demonstrating that people do not experience hindsight bias when the outcome was caused by “unforeseeable

185. See Rachlinski, *Judging in Hindsight*, *supra* note 20, at 587 (“Complete elimination of the bias has eluded psychologists.”).

186. David Alan Sklansky, *Evidentiary Instructions and the Jury as Other*, 65 *Stan. L. Rev.* 407, 415 (2013).

187. This last technique is less “debiasing” than “insulation” from outcome information. See Christine Jolls & Cass Sunstein, *Debiasing Through Law*, 35 *J. Legal Stud.* 199, 225 n.21 (2006) (noting authors’ use of the word “debiasing” “to refer specifically to situations in which bounded rationality diminishes, as opposed simply to having its effects on outcomes blunted by a legal rule”). Here, I use the term “debiasing” to encompass any technique that combats bias.

188. See Scott A. Hawkins & Reid Hastie, *Hindsight: Biased Judgments of Past Events After the Outcomes Are Known*, 107 *Psychol. Bull.* 311, 311 (1990) (“We provide an interpretation of the creeping determinism hypothesis in terms of inferences made to reevaluate case-specific evidence once the relevant outcome is known and conclude that it is the most common mechanism underlying observed hindsight effects.”); Rachlinski, *Judging in Hindsight*, *supra* note 20, at 584 (explaining “creeping determinism” as “people naturally integrat[ing] an outcome and the events that preceded it into a coherent story,” thereby “making [certain] circumstances seem more significant than they appeared in foresight”); Roese & Vohs, *supra* note 20, at 415 (“[S]ensemaking contributes mainly to inevitability . . .”).

189. Fischhoff, *Hindsight ≠ Foresight*, *supra* note 158, at 310.

'chance' factors"¹⁹⁰ and outcome information alone is insufficient to elicit hindsight bias without a potential causal antecedent.¹⁹¹

The consider-the-opposite strategy capitalizes on this cognitive, causal understanding of the bias by forcing decisionmakers to consider plausible alternative scenarios in which the same facts resulted in different outcomes.¹⁹² The idea is that, by "breaking down the causal links between outcome knowledge and antecedent behavior," the hindsight bias can be mitigated.¹⁹³ This method has "proven consistently effective,"¹⁹⁴ reducing hindsight bias "markedly" in most studies that have employed it.¹⁹⁵ However, the most successful implementations of this strategy have not simply encouraged decisionmakers to think of alternatives but instead actually required them to write down alternative scenarios or come up with probability estimates for given alternatives.¹⁹⁶ For example, in one study, when mock jurors learned that an auditor concluded a certain company was stable but the company failed six months later, they judged the auditor more favorably after assessing the probability of two alternative outcomes that would have allowed the company to remain solvent.¹⁹⁷ The consider-the-opposite strategy is the most promising debiasing technique for the courtroom when a judge wishes to admit or consider hindsight evidence.

190. David Wasserman, Richard O. Lempert & Reid Hastie, *Hindsight and Causality*, 17 *Personality & Soc. Psychol. Bull.* 30, 30 (1991).

191. See Jennelle E. Yopchick & Nancy S. Kim, *Hindsight Bias and Causal Reasoning: A Minimalist Approach*, 13 *Cognitive Processing* 63, 70 (2012) ("We suggest that hindering causal reasoning resulted in the disappearance of hindsight bias and promoting causal reasoning elicited its appearance."); see also John C. Anderson, D. Jordan Lowe & Philip M.J. Reckers, *Evaluation of Auditor Decisions: Hindsight Bias Effects and the Expectation Gap*, 14 *J. Econ. Psychol.* 711, 730 (1993) ("Individuals tended to give higher relevance to negative factors (cues) when the outcome was negative and to give higher relevance to positive factors (cues) when the outcome was positive.").

192. See Roese & Vohs, *supra* note 20, at 418 ("The consider-the-opposite strategy . . . illuminat[es] novel, previously unconsidered means by which the same or different outcomes might have occurred.").

193. D. Jordan Lowe & Philip M.J. Reckers, *The Effects of Hindsight Bias on Jurors' Evaluations of Auditor Decisions*, 25 *Decision Sci.* 401, 406 (1994).

194. Roese & Vohs, *supra* note 20, at 418.

195. Peters, *supra* note 167, at 1289.

196. See, e.g., Philip E. Tetlock, *Expert Political Judgment 203-04* (2005) (discussing experiment in which subjects were asked for their "expert opinion on how close we came to experiencing alternative outcomes"); Hal R. Arkes et al., *Eliminating the Hindsight Bias*, 73 *J. Applied Psychol.* 305, 306 (1988) (subjects wrote down one piece of evidence supporting each possible outcome); Lowe & Reckers, *supra* note 193, at 412 (subjects assessed alternative outcomes and provided their own); Michelle R. Nario & Nyla R. Branscombe, *Comparison Processes in Hindsight and Causal Attribution*, 21 *Personality & Soc. Psychol. Bull.* 1244, 1247 (1995) (subjects explained how antecedents could have produced alternative outcome).

197. Lowe & Reckers, *supra* note 193, at 415. Those in the debiasing condition also provided their own alternative outcome. *Id.* at 412.

There is little danger that employing consider-the-opposite will prevent jurors from appreciating the *probative* value of hindsight evidence. A juror performs exactly the right inquiry when he or she considers the likelihood of *each* possible outcome, given the factual circumstances *ex ante*; doing so should let him or her derive the correct probability of observing this outcome under each party's story.¹⁹⁸ In addition, jurors reasoning by inference to the best explanation will notice whether the alternative accounts are coherent, powerful stories or implausible explanations of the outcome.¹⁹⁹ This suggests that the consider-the-opposite strategy truly counteracts a bias—a willingness to see only one possible story—by forcing decisionmakers to consider the likelihood of more possible outcomes and arrive at better-informed probability estimates.

Where the fact-finder is a jury, the lawyers may be more effective at debiasing than the judge. Judges can hardly force jurors to write down alternatives. But this cumbersome task may not be necessary when attorneys provide alternative causal accounts of the evidence. For example, if a police officer shot someone who turned out to be unarmed,²⁰⁰ the officer's attorney might say to the jury:

The plaintiff wants you to think that there is one motion people make when they pull out a gun and another motion they make when they reach for identification. But trained police officers know this is not so simple. As the evidence has shown, through their training, they know that motions that might seem innocent to you and me and bystanders—motions that ultimately *may be* innocent, as this case tragically demonstrates—are often *not* innocent. Policemen have seen videos demonstrating how quickly a peaceful encounter can turn into a fatal shooting of an officer. That's what my client saw here: a situation that looked like countless others where a cop got shot.²⁰¹

198. In a study by Professors Ruth Beyth-Marom and Fischhoff, subjects failed to seek out information about the denominator of the Bayesian likelihood ratio—the probability of observing the evidence if the hypothesis is false—but did use the information, in conjunction with information about the numerator, when it was presented to them. Ruth Beyth-Marom & Baruch Fischhoff, *Diagnosticity and Pseudodiagnosticity*, 45 *J. Personality & Soc. Psychol.* 1185 (1983); see also James S. Liebman et al., *The Evidence of Things Not Seen: Non-Matches as Evidence of Innocence*, 98 *Iowa L. Rev.* 577, 636–42 (2013) (discussing “uniqueness fallacy,” wherein people “jump to the numerator conclusion” if numerator is high without considering denominator). Consider-the-opposite helps repair this deficit by providing information about the denominator.

199. See Roeser & Vohs, *supra* note 20, at 418 (noting that when a decisionmaker finds it difficult to think of explanations for alternatives, he or she may conclude alternatives are implausible).

200. For discussion of this example, see *supra* text accompanying notes 1–4, and *infra* text accompanying notes 239–278.

201. See Scott Fielden, *The Mind of a Cop: What They Do, and Why They Do It* 20 (2009) (noting that on the first day of training for Metro Nashville Police Department,

In this way, an attorney can flesh out an alternative scenario, allowing jurors to more easily imagine a different outcome.

Research on whether attorney argument might be effective in this way has yielded somewhat mixed results. In the one mock-juror study by researchers Merrie Jo Stallard and Debra Worthington where verbal encouragement alone *did* mitigate the bias, a defense attorney attempted “to focus subject-jurors’ attention on the time prior to the outcome’s occurrence” by warning them that the plaintiff was trying to turn them into Monday-morning quarterbacks and urging them to avoid the temptations of hindsight.²⁰² The authors suggest that the defense attorney was effective because he “provid[ed] participants with an alternative story upon which they could draw.”²⁰³ This fits with Professor Philip Tetlock’s suggestion that considering alternatives debiases because people rate *easier-to-imagine* scenarios as more likely.²⁰⁴ It fits with work showing that subjects forced into the cognitively difficult task of imagining *too many* alternative scenarios experience an *increase* in bias.²⁰⁵ And it fits with the natural reasoning process of inference to the best explanation:²⁰⁶ If people weigh evidence by how well hypotheses explain it, presenting jurors with two causal stories will foster better-informed weighing. The strongest evidence against this idea, however, is a study by Kim Kamin and Professor Rachlinski where mock jurors who heard both attorney

recruits watched videos of officers being killed during routine jobs); Seth Stoughton, Law Enforcement’s “Warrior” Problem, 128 Harv. L. Rev. Forum 225, 227 (2015) (describing officers in training watching videos of other officers being shot); see also Matt Apuzzo, Training Officers to Shoot First, and He Will Answer Questions Later, N.Y. Times (Aug. 1, 2015), <http://www.nytimes.com/2015/08/02/us/training-officers-to-shoot-first-and-he-will-answer-questions-later.html> (on file with the *Columbia Law Review*) (reporting Force Science Institute, which has videos demonstrating how quickly encounters can unfold, has “trained tens of thousands of police officers”).

202. Merrie Jo Stallard & Debra L. Worthington, Reducing the Hindsight Bias Utilizing Attorney Closing Arguments, 22 Law & Hum. Behav. 671, 675 (1998). In the foresight condition, 29% concluded a board’s decision was negligent, in contrast to 57% in the hindsight condition and 37% in the debiasing condition. When asked whether the board “should institute” or “should have instituted” its plan, 67% of the foresight group said yes, in contrast to 39% of the hindsight group and only 65% of the debiased group. *Id.* In another study, “jurors,” who heard closing arguments, exhibited less hindsight bias than “citizens,” who did not. In hindsight, 74% of “citizens” favored punitive damages, in contrast to 59% of “jurors,” but only 33% of participants in the foresight condition would have prohibited the defendant’s conduct. Hastie et al., Juror Judgments, *supra* note 57, at 607.

203. Stallard & Worthington, *supra* note 202, at 680; see also *id.* at 681 (linking this debiasing effect to availability heuristic).

204. See Tetlock, *supra* note 196, at 197, 204 (observing effects of “imaginability”).

205. Lawrence J. Sanna, Norbert Schwartz & Shevaun L. Stocker, When Debiasing Backfires: Accessible Content and Accessibility Experiences in Debiasing Hindsight, 28 J. Experimental Psychol.: Learning, Memory & Cognition 497, 500–01 (2002) (discussing “ironic effects” of generating many alternatives).

206. See *supra* section II.A.2.

arguments and debiasing instructions demonstrated strong hindsight bias.²⁰⁷

Although these studies and arguments are far from conclusive, then, they strongly suggest that attorney argument presenting alternative causal accounts of outcome evidence will decrease, though not eliminate, hindsight bias, making the bias less influential at trial than in the lab.²⁰⁸ Judges should consider this when they weigh probative value against prejudice under Rule 403.

Judges could also *encourage* jurors to consider the opposite through jury instructions.²⁰⁹ But judges rarely use instructions to debias,²¹⁰ and they may be reluctant to start. In a bench trial, however, the judge *can* ensure that the fact-finder takes the time and energy to debias herself under the consider-the-opposite strategy. While judges may be somewhat more resistant to hindsight bias than the general public, they are not immune and typically tend to overvalue hindsight evidence.²¹¹ The humble judge, then, should acknowledge his or her own fallibility²¹² and commit to painstakingly writing out at least one or two alternative scenarios with different outcomes when he or she receives hindsight evidence. As silly as the exercise may seem, judges who force themselves to do it will maximize their own accuracy.

2. *Consequence Severity.* — At least one study has suggested that people exhibit less hindsight bias when they know their decision will have

207. See Kamin & Rachlinski, *supra* note 22, at 98 (“[A]nalysis comparing the hindsight cell to the debiasing cell showed no significant differences . . .”).

208. See Peters, *supra* note 167, at 1292 (“What we do know suggests that the hindsight biases may have less influence in actual jury trials than in experimental settings.”).

209. A judge inclined to use the best available evidence to instruct jurors might tell them to:

Imagine specific alternative scenarios, including any proposed by the defendant, where the defendant makes the same choice but the outcome is different. It might help to write down these narratives. Thinking of these stories can help you avoid the temptation to be a “Monday-morning quarterback” and to second guess the defendant’s decision with the benefit of hindsight.

These instructions incorporate elements of the defense attorney’s argument in the Stallard and Worthington study. See Stallard & Worthington, *supra* note 202, at 675 (summarizing defense attorney’s argument). Professor Simon achieved some success in the coherence-based reasoning context by telling subjects simply to “take some time to seriously consider the possibility that the opposite side has a better case.” Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. Chi. L. Rev. 511, 544 (2004) [hereinafter Simon, *Cognitive Coherence*].

210. See Rachlinski, *Judging in Hindsight*, *supra* note 20, at 603 (“In fact, courts do not attempt to use judicial instructions as a means of debiasing jurors.”).

211. See *supra* notes 181–183 and accompanying text (discussing studies of hindsight bias in judicial decisionmaking).

212. See Suzanna Sherry, *Judges of Character*, 38 Wake Forest L. Rev. 793, 799 (2003) (discussing history of calls for judicial humility).

serious consequences for someone's life than when they make a judgment in the abstract. In that study, mock jurors evaluated whether an auditor was negligent in failing to detect fraud. Although simply telling the mock jurors to avoid hindsight bias had little impact, they exhibited effectively no bias when informed that under the law, "auditors found negligent may be severely fined and/or expelled from the Institute of Chartered Accountants."²¹³ In another study, subjects assessing a company's decision to operate a railroad in a hazardous location were assigned to one of two roles: a juror rendering a verdict or a citizen giving an opinion. The jurors exhibited significantly less hindsight bias than the citizens, suggesting that the seriousness of their role influenced the level of bias.²¹⁴ The jurors also heard attorney argument and judicial instructions,²¹⁵ so consider-the-opposite may have played a part in mitigating the bias.

This research, like the consider-the-opposite research, suggests that actual jurors may exhibit the hindsight bias to a lesser degree than study subjects.²¹⁶ In court, the stakes are high, and jurors are aware that their verdict will have serious consequences. Judges may consider this when evaluating the potential prejudice of hindsight evidence.

3. *Bifurcation*. — Bifurcation²¹⁷ serves to mitigate bias by insulating jurors from the full weight of outcome evidence. A number of scholars have suggested bifurcating trials to avoid hindsight bias. In negligence or medical malpractice cases, jurors could determine liability before hearing evidence of damages.²¹⁸ Alternatively, jurors could determine which, if any, actions would have been negligent before knowing which

213. Peter M. Clarkson, Craig Emby & Vanessa W.-S. Watt, *Debiasing the Outcome Effect: The Role of Instructions in an Audit Litigation Setting*, *Auditing: J. Prac. & Theory*, Sept. 2002, at 7, 18 (emphasis omitted); see also *id.* at 13 ("Based on these mean scores, it appeared that the Evaluation Consequences Instructions to all intents and purposes restored the foresight perspective.").

214. Hastie et al., *Juror Judgments*, *supra* note 57, at 607.

215. *Id.* at 603.

216. See Peters, *supra* note 167, at 1300 ("As a result [of juror motivation], hindsight bias may be easier for defense counsel to counteract in an actual trial than in an experimental setting.").

217. See Fed. R. Civ. P. 42(b) ("For convenience, to avoid prejudice, or to expedite and economize, the court may order a separate trial of one or more separate issues, claims, crossclaims, counterclaims, or third-party claims." (emphasis added)).

218. See Norman G. Poythress, Richard Wiener & Joseph E. Schumacher, *Reframing the Medical Malpractice Tort Reform Debate: Social Science Research Implications for Non-Economic Reforms*, 16 *Law & Psychol. Rev.* 65, 109 (1992) ("In our view, these findings suggest that the courts should give greater consideration to the use of bifurcated trials in medical malpractice cases."); David B. Wexler & Robert F. Schopp, *How and When to Correct for Juror Hindsight Bias in Mental Health Malpractice Litigation: Some Preliminary Observations*, 7 *Behav. Sci. & L.* 485, 496 (1989) ("[I]t may make sense to bifurcate only in those categories of cases where hindsight bias is particularly strong and persistent.").

option the defendant chose.²¹⁹ Bifurcation is not a perfect solution to the problem of hindsight bias. In negligence cases, jurors likely know that the plaintiff suffered *some* injury and may even be able to infer the type of injury.

Still, bifurcation has proven somewhat effective. In one recent study, “[j]urors who heard evidence regarding the defendant’s conduct but no evidence of the plaintiff’s injuries judged the defendant’s behavior to be less negligent than did jurors who heard about both the conduct and the injuries.”²²⁰ Studying a period when the Northern District of Illinois permitted separate trials on liability and damages, pioneering jury researchers Hans Zeisel and Thomas Callahan found that while only about a third of the unitary trials reached a verdict in the defendant’s favor, over half of the bifurcated trials resulted in a pro-defendant verdict.²²¹ A study of simulated jurors reached a similar result.²²²

When a judge decides that hindsight evidence is *not* admissible for the purposes of evaluating *ex ante* decisionmaking but the evidence *is* inevitably part of the case—such as the fact of injury in a negligence action—bifurcation is the most straightforward solution. This solution may be more feasible in some circumstances than others. The liability and damages phases of a negligence case could easily be conducted in sequence. But in, say, a patent case in which a jury has to determine whether an invention was “obvious”²²³ and where learning about the invention elicits hindsight bias, the jury will almost certainly need to learn what the invention was. Where it is feasible, however, bifurcation could effectively eliminate outcome information from jury consideration.

4. *Rules that Counteract Hindsight Bias.* — In addition to these debiasing mechanisms, judges—or lawmakers—could expand the scope of any of the rules that already serve to counteract the hindsight bias. In his seminal paper on hindsight, Professor Rachlinski catalogued these techniques, which include adopting industry norms as an acceptable standard of care,²²⁴ requiring more than the mere fact of injury in order

219. Christine Jolls, Cass R. Sunstein & Richard Thaler, A Behavioral Approach to Law and Economics, 50 *Stan. L. Rev.* 1471, 1527–28 (1998). This sort of bifurcation is best suited to cases where the defendant faced a set of concrete choices, such as operating on a patient or not operating.

220. Alison C. Smith & Edith Greene, Conduct and Its Consequences: Attempts at Debiasing Jury Judgments, 29 *Law & Hum. Behav.* 505, 522 (2005).

221. Hans Zeisel & Thomas Callahan, Split Trials and Time Saving: A Statistical Analysis, 76 *Harv. L. Rev.* 1606, 1612 *tbl.3* (1963).

222. See Irwin A. Horowitz & Kenneth S. Bordens, An Experimental Investigation of Procedural Issues in Complex Tort Trials, 14 *Law & Hum. Behav.* 269, 277 (1990) (“Unitary trials produced significantly more proplaintiff, affirmative verdicts on general causation (85.4%) than did separated trials (68.6%).”).

223. 35 U.S.C. § 103 (2012).

224. Rachlinski, *Judging in Hindsight*, *supra* note 20, at 610–13 (suggesting that where evidence shows industry norms match reasonable care standard, courts should allow compliance with norms as a defense).

to state a claim,²²⁵ adopting doctrinal rules that curtail liability for certain actors,²²⁶ and of course, excluding entire categories of evidence.²²⁷

If judges, rulemakers, or legislators are convinced that hindsight evidence is too problematic in a well-defined category of cases—either because the bias is too prejudicial or because the evidence counters policy goals—rules like these can save time and further legitimate aims of the judicial system.²²⁸ For example, courts guard against hindsight bias with respect to corporate directors through the business judgment rule, which protects them from liability for negligence.²²⁹ Courts review directors' decisions only when they act in bad faith.²³⁰ Although hindsight evidence might be probative of the failure to exercise due care, the business judgment rule places a thumb on the scale in favor of directors for policy reasons—to encourage risk-taking so shareholders can see high returns.²³¹

5. *Summary.* — The preceding discussion of hindsight bias and debiasing is far from fully satisfying. The magnitude of the bias remains unclear, with loose bounds circumscribed by existing research. In one study, debiasing attorney arguments had a large effect, whereas in another, attorney arguments and debiasing instructions failed entirely. But understanding the state of this research is better than either maintaining an ill-founded fear that the bias will dominate all decisionmaking or dismissing the bias as trivial. Judges—and the lawyers who educate them—can use the existing evidence of a bias that is substantial but not inevitably dispositive, and possibly mitigated by attorney argument, to make the best possible judgments about whether to admit hindsight evidence and whether to bifurcate a trial.

C. *Admitting Hindsight Evidence*

A judge confronted with a motion to exclude hindsight evidence must decide whether probative value is “substantially outweighed” by prejudice or other dangers when neither element is easily quantifiable. This challenge is as inevitable as it is unenviable—all exclusion motions under Rule 403 require judges to estimate and compare probative value and danger in the full context of the case.²³² Although the weighing

225. See *id.* at 616 (noting that in both negligence and federal securities cases, courts require more than mere fact of injury).

226. *Id.* at 619–23 (discussing business judgment rule).

227. *Id.* at 617–18 (discussing Federal Rule of Evidence 407, which declares subsequent remedial measures evidence inadmissible).

228. See *id.* at 610–17 (evaluating advantages and disadvantages of such rules).

229. E.g., Stephen M. Bainbridge, *Corporate Law* § 6.1, at 96 (2d ed. 2009).

230. See, e.g., *Kamin v. Am. Express Co.*, 383 N.Y.S.2d 807, 810 (Sup. Ct. 1976) (explaining standard).

231. See, e.g., Bainbridge, *supra* note 229, § 6.2 at 104 (discussing policy rationale).

232. See Kahan, *supra* note 29, at 1638–39 (citing *Old Chief v. United States*, 519 U.S. 172, 182–83 (1997)). This task may be all the more difficult if, as Professor Simon's

process is highly case specific and any quest for a single, definitive admissibility rule would be quixotic, this Essay has provided guidance for estimating the magnitude of probative value and prejudice.

When weighing probative value and prejudice, judges should not simply compare the magnitude of each but rather think about the chances that admitting the evidence will cause the jury to err in its verdict. As Professor Dan Kahan has discussed, in theory someone weighing probative value against prejudice should compare the increase in marginal likelihood of error due to overvaluation of evidence to the increase in marginal likelihood of error due to exclusion of the probative value, and the evidence should be excluded only if the latter is smaller.²³³ This principle suggests that when the case is otherwise close and hindsight evidence is of sufficient probative value that the judge expects it to change the vote of most rational jurors, the judge should tend to admit the evidence. When the case is lopsided and neither the evidence nor hindsight bias is likely to sway jurors, judges may also wish to admit hindsight evidence of even moderate value if it will not waste time.²³⁴ However, when the case is sufficiently close, the hindsight evidence sufficiently weak, and the expectation of bias sufficiently strong that the judge anticipates jurors who would not flip for the probative value will flip due to bias,²³⁵ prejudice is so substantial that the evidence should be excluded.²³⁶

The judge does not have perfect information and must estimate all of these values—the closeness of the case, the probative weight of the evidence, and the evidence’s potential to bias—to the best of his or her ability. Although judges will fall short of perfection, this is the sort of judgment we expect them to exercise regularly under Rule 403,²³⁷ and the framework in this Essay can help them at least think of probative value and prejudice in an informed, reasoned way.

research on coherence-based reasoning suggests, a piece of evidence that favors one side may influence how a fact-finder views the probative value of *other* evidence. See Simon, Cognitive Coherence, *supra* note 209, at 537–40 (discussing “effect of changes in one task variable on other variables”). Even evidence that is not prejudicial in itself, then, may have a hidden prejudicial effect.

233. Kahan, *supra* note 29, at 1634.

234. See Fed. R. Evid. 403 (permitting exclusion for “undue delay, wasting time, or needlessly presenting cumulative evidence”).

235. In estimating likely bias, the judge should account for any debiasing techniques he plans to use. See *supra* section IV.B (outlining debiasing methods).

236. Professor Kahan shows that if hindsight bias simply increases the perceived likelihood ratio of the evidence, judges should *exclude* evidence favorable to the plaintiff if the plaintiff has a *weak* case but *admit* it if the case is *close*. Kahan, *supra* note 29, at 1636–38. This heuristic is highly administrable, but it does not allow for the possibility that hindsight bias has a relatively small effect when a case is otherwise very weak.

237. See Kahan, *supra* note 29, at 1638–39 (citing *Old Chief*, 519 U.S. at 182–83).

V. APPLICATION TO AREAS OF DOCTRINAL CONFUSION

This Part applies the framework set out in the previous Parts to four examples where courts have disagreed over whether hindsight evidence is relevant. This is not a complete survey of hindsight evidence in the law. Instead, it explores a small sample of issues from widely disparate doctrinal areas, each illuminating a division within the courts that has been underexplored in the academic literature. Courts' relevance determinations in these cases seem scattershot. They sometimes appear to be infected by policy considerations and excessive concern about hindsight bias, at other times they are unreasoned declarations of relevance or irrelevance, and only occasionally are they grounded in a sound understanding of why this evidence is relevant. The discussion of each example examines the considerations of relevance—and where a jury is involved, prejudice—that should factor into the courts' determinations.

The first example is relatively clean and paradigmatic: When a plaintiff sues a police officer alleging excessive force and the officer says she shot the man because she believed he was drawing a gun, should the court admit evidence that the person who was shot did not, in fact, have a weapon? The three subsequent examples, like *Tevye's daughters*,²³⁸ increasingly deviate from the standard. In the first, relating to the enforceability of liquidated damages clauses, no jury is involved: The judge himself decides how heavily to weigh any evidence. In the second, related to placement determinations under the IDEA, administrative hearing officers, not federal judges, are the primary fact-finders, and judges defer to their factual determinations. In the third, concerning motions for judgment as a matter of law, the “hindsight evidence”—the jury's verdict—is not evidence at all; it is simply information that the judge can use to make a legal determination. But in each of these cases hindsight evidence has a role to play.

A. *Excessive Force and Self-Defense*

Although no comprehensive database of police shootings exists, every year hundreds or thousands of people are shot by police officers in the United States.²³⁹ After a shooting, either the person shot or (if that person was killed) a relative might sue the officer under section 1983,²⁴⁰ alleging that she used excessive force in violation of the Fourth Amendment. The officer might respond that she fired at the plaintiff

238. See *Fiddler on the Roof* (United Artists 1971).

239. See Wesley Lowery & Steven Rich, *Since 1976, the FBI Hasn't Counted More than 460 Fatal Police Shootings in a Year. We've Counted 463 Already in 2015.*, Wash. Post: Post Nation (July 1, 2015), <http://www.washingtonpost.com/news/post-nation/wp/2015/06/30/wp/2015/07/01/since-1976-the-fbi-hasnt-counted-more-than-460-fatal-police-shootings-in-a-year-weve-counted-461-already-in-2015/> [<http://perma.cc/QAC8-ZMV6>].

240. 42 U.S.C. § 1983 (2012).

because it appeared to her that he was about to draw a weapon.²⁴¹ Sometimes the person indeed had a gun. Other times, no weapon is found: The plaintiff's hands were in empty pockets or he had only a cellphone. In these cases, courts must decide whether to admit evidence that the plaintiff was unarmed as evidence that the officer's action was unreasonable. Although this question has recurred in the case law, the academic literature has not yet addressed it.²⁴² This section, then, introduces a new issue into legal scholarship by analyzing whether courts should consider evidence that the person who was shot did or did not have a gun. I conclude that this evidence is typically relevant and admissible.

A section 1983 claim against the police alleging excessive force is analyzed under the Fourth Amendment's "objective reasonableness" standard, which looks at whether the officer's actions were reasonable in light of the circumstances, independent of her subjective intent.²⁴³ "Reasonableness" is "judged from the perspective of a reasonable officer on the scene, rather than with the 20/20 vision of hindsight."²⁴⁴ Police may use deadly force if "they have probable cause to believe they must protect themselves or another person present from imminent death or serious physical injury."²⁴⁵

Many courts deem evidence that the plaintiff was unarmed irrelevant. The leading case on this issue, discussed briefly in the

241. Of course, in some cases, the force is more clearly justified. See, e.g., *Neuberger v. Thompson*, 124 F. App'x 703, 704 (3d Cir. 2005) (noting complaint admitted decedent pointed handgun at police and said, "You're making me do this").

242. Scholars have addressed the related question of whether the fact that plaintiff's decedent intended to commit "suicide by cop" is relevant. See Rahi Azizi, *When Individuals Seek Death at the Hands of the Police: The Legal and Policy Implications of Suicide by Cop and Why Police Officers Should Use Nonlethal Force in Dealing with Suicidal Suspects*, 41 *Golden Gate U. L. Rev.* 183, 198–208 (2011) (discussing decisions addressing whether evidence of suspect's suicidal intent is relevant to merits of § 1983 claim); Timothy P. Flynn & Robert J. Homant, 'Suicide by Police' in Section 1983 Suits: Relevance of Police Tactics, 77 *U. Det. Mercy L. Rev.* 555, 559–66 (2000) (discussing how courts have handled "pre-seizure evidence" in suicide-by-police cases). Academics have also criticized the standards by which the legal system judges police conduct. See James J. Fyfe, *The Split-Second Syndrome and Other Determinants of Police Violence*, in *Violent Transactions: The Limits of Personality* 207, 217–21 (Anne Campbell & John J. Gibbs eds., 1986) (arguing police should be responsible for avoiding need to make dangerous split-second decisions); Kathryn R. Urbonya, *Dangerous Misperceptions: Protecting Police Officers, Society, and the Fourth Amendment Right to Personal Security*, 22 *Hastings Const. L.Q.* 623, 628 (1995) (arguing "danger is not the only factor to consider in determining whether a police practice is unconstitutional").

243. *Graham v. Connor*, 490 U.S. 386, 397 (1989).

244. *Id.* at 396.

245. N.Y.C. Police Dep't, *Patrol Guide: Deadly Physical Force* (Aug. 1, 2013), http://muckrock.s3.amazonaws.com/foia_files/203-12_Deadly_Physical_Force.pdf [<http://perma.cc/9JU2-G8AD>].

Introduction, is *Sherrod v. Berry*,²⁴⁶ an in banc decision from the Seventh Circuit. The facts of *Sherrod*, according to the defendant-officer, were as follows: Officer Berry and his partner were en route to the scene of a recent robbery when they pulled over a car with two men in it.²⁴⁷ Officer Berry recognized the passenger as a man named Duckworth, who matched the description of the robber.²⁴⁸ Officer Berry and his partner exited their vehicle and approached the car, guns pointed at the occupants.²⁴⁹ Officer Berry had to order the suspects to raise their hands three times before they complied.²⁵⁰ He testified that this reluctance raised his suspicions that the two were dangerous, and he became more suspicious when he saw them looking at each other as if to say, “[W]hat are we going to do next?”²⁵¹ The officer raised his gun and approached the car.²⁵² He testified that at that point, he saw the driver, Ronald Sherrod, make a “quick movement with his hand into his coat . . . [as if] he was going to reach for a weapon.”²⁵³ Officer Berry fired at Sherrod and killed him instantly.²⁵⁴ Sherrod’s father sued, and at trial, over defendants’ objection, he introduced evidence that a search of Sherrod failed to find a weapon.²⁵⁵ The jury found in plaintiff’s favor and awarded over \$1.6 million.²⁵⁶

The Seventh Circuit voted eight to three to reverse the verdict and remand for a new trial.²⁵⁷ The majority opinion, written by Judge Coffey, held that under Rule 401, the evidence that Sherrod was unarmed was irrelevant and therefore inadmissible. It “ha[d] no place in the trial court’s or jury’s proper post-hoc analysis of the reasonableness of the actor’s judgment” because “Officer Berry’s liability [must] be determined exclusively upon an examination and weighing of the information Officer Berry possessed immediately prior to and at the very moment he fired the fatal shot.”²⁵⁸ He wrote that courts should be reluctant “to second-guess an officer’s split-second reasonable judgment to protect himself and those around him through the use of deadly force.”²⁵⁹ Reasoning that the jury should be put in the same position as Officer

246. 856 F.2d 802 (7th Cir. 1988) (en banc). *Sherrod* was decided before *Graham v. Connor* but used an equivalent substantive standard.

247. *Id.* at 803.

248. *Id.*

249. *Id.*

250. *Id.*

251. *Id.*

252. *Id.*

253. *Id.*

254. *Id.*

255. *Id.* at 804.

256. *Id.*

257. *Id.* at 803.

258. *Id.* at 805. Judge Ripple voted to reverse but on different grounds; he agreed with Judge Cummings’s dissent on the evidentiary issue. *Id.* at 808 (Ripple, J., concurring).

259. *Id.* at 805 (majority opinion).

Berry was when he shot Sherrod, the court continued, “The Sioux Indians have a prayer that asks for this wisdom: ‘Grant that I may not judge another until I have walked a mile in his moccasins.’”²⁶⁰

Judge Cummings’s impassioned dissent called the decision a “miscarriage of justice” and contrasted the majority’s account of “a good police officer in a bad situation confronting a recalcitrant suspect” with the record’s portrayal of “a bad cop with a history of unnecessary violence, acting in violation of police procedure, confronting an innocent victim.”²⁶¹ He suggested the fact that Sherrod was unarmed was relevant for two central reasons. First, it could “help[] the jury understand the type of movement Sherrod made that precipitated his death,” as “[t]he manner in which Sherrod reached into his pocket is of great consequence to the determination of the action.”²⁶² Second, the lack of a weapon could bear on the credibility of Officer Berry as a witness. The fact that Sherrod was unarmed makes it less likely that Officer Berry believed he saw Sherrod moving to draw a gun, even though it does not render that belief impossible, and it makes it less likely that Officer Berry made a reasonable observation. “Simply stated . . . the impeachment goes . . . to the probability that the officer’s version of the events is credible.”²⁶³

Several courts have followed the majority in *Sherrod*, concluding that evidence the plaintiff or decedent was unarmed when he was shot is irrelevant to the determination of an excessive force claim,²⁶⁴ and several other courts have come to the same conclusion on their own.²⁶⁵ But a few judges have resisted this trend. In one case out of the Eastern District of Pennsylvania, the court found evidence that the plaintiff was unarmed to

260. Id. at 806.

261. Id. at 808 (Cummings, J., dissenting).

262. Id. at 810 (Cummings, J., dissenting).

263. Id. at 811 (Cummings, J., dissenting).

264. See, e.g., *Morales v. Holly*, No. 1:09CV175, 2012 WL 4511068, at *6 (M.D.N.C. Sept. 28, 2012) (“[T]he Fourth Circuit has recognized that evidence that a plaintiff-suspect actually was unarmed at the time a defendant-officer used force often will have no relevance . . . because the reasonableness inquiry requires scrutiny only of the facts reasonably perceived by the defendant-officer at the time.” (citing *Greenridge v. Ruffin*, 927 F.2d 789, 792 (4th Cir. 1991); *Sherrod*, 856 F.2d 802)); *Viera v. City of El Monte*, No. CV 04-06082 MMM (Rzx), 2006 WL 6626761, at *5 (C.D. Cal. May 2, 2006) (“Based on *Sherrod* and the above-cited authorities, the court concludes that the fact Viera was unarmed is not relevant in assessing whether Lt. Hernandez reasonably used deadly force.”); *Bougress v. Mattingly*, 426 F. Supp. 2d 601, 606 (W.D. Ky. 2006) (“Initially, we note that the fact that Newby was armed at the time of the incident is not relevant to the question of whether Mattingly reasonably believed at the time that the use of deadly force was necessary.” (citing *Sherrod*, 856 F.2d 802)).

265. See *Reese v. Anderson*, 926 F.2d 494, 501 (5th Cir. 1991) (“Also irrelevant is the fact that Crawford was actually unarmed.”); see also *Anderson v. Russell*, 247 F.3d 125, 132 (4th Cir. 2001) (quoting *Reese*, 926 F.2d at 501). Although the *Reese* opinion uses the word “irrelevant,” it is not wholly clear that the court actually purports to undertake a Rule 401 evaluation. It may mean that the evidence was simply unpersuasive.

be relevant and admissible.²⁶⁶ Two officers had given contradictory testimony—one said the plaintiff thrust out with a shiny object, which was actually a cellphone, and the other said he did not²⁶⁷—and the court said that the lack of a weapon was relevant to credibility.²⁶⁸ Further, the court reasoned, a jury would need to decide whether a cellphone could be mistaken for a weapon.²⁶⁹ A dissenting judge on a recent Sixth Circuit panel would have denied summary judgment to officers who claimed that the decedent had pointed a silver object out of a car window, yelling, “I have a gun!”²⁷⁰ As the judge wrote, “The problem is, of course, there was no gun at all.”²⁷¹ He would have sent the case to a jury to evaluate the officers’ credibility.²⁷²

Under the framework set out in this Essay, the dissenters—Judge Cummings in *Sherrod* and this latter group of judges—are correct. Evidence that the decedent did not have a weapon is relevant, probative hindsight evidence in most cases. Judge Cummings’s dissent notes a question of how *Sherrod* moved.²⁷³ The hindsight evidence made it more likely that he moved in a way that a reasonable police officer would not interpret as reaching for a weapon—perhaps a slower, more casual motion, as opposed to a quick-draw motion. Judge Cummings also notes that the evidence casts doubt on Officer Berry’s overall credibility as a witness.²⁷⁴ Indeed it does: A story unlike Officer Berry’s is more likely to accompany no evidence of a weapon than Officer Berry’s story is. This reasoning will hold in many similar cases, where there is little objective evidence about what occurred and the evidence supports a story other than what the officer recounts.

Other evidence might reduce the hindsight evidence’s value so much that the judge should exclude it under Rule 403. For example, a

266. *Pelzer v. City of Philadelphia*, No. 07–CV–0038, 2011 WL 93054, at *1 (E.D. Pa. Jan. 11, 2011).

267. *Id.* at *5.

268. *Id.*

269. *Id.* at *6; see also *Cooper v. City of Chester*, No. 11–5381, 2013 WL 925067, at *2 & n.3 (E.D. Pa. Mar. 11, 2013) (deeming absence of weapon in vicinity admissible, as it contradicted defendant’s claim that “he saw [plaintiff] turning toward him with a metallic object in his hand that looked like a gun”); *Horton v. City of Harrisburg*, No. 1:06-CV-2338, 2009 WL 2225386, at *3 (M.D. Pa. July 23, 2009) (considering fact that Horton was actually unarmed when denying summary judgment to shooting officer who claimed Horton pulled small dark object from his coat when cornered).

270. *Simmonds v. Genesee Cty.*, 682 F.3d 438, 447 (6th Cir. 2012) (Merritt, J., dissenting).

271. *Id.* (Merritt, J., dissenting).

272. *Id.* at 447–48 (Merritt, J., dissenting).

273. *Sherrod v. Berry*, 856 F.2d 802, 810 (7th Cir. 1988) (en banc) (Cummings, J., dissenting).

274. *Id.* at 810–11 (Cummings, J., dissenting).

video of the shooting²⁷⁵ or uncontradicted testimony from multiple disinterested witnesses might confirm that a reasonable officer could or would have perceived that the decedent was about to draw a weapon.²⁷⁶ A judge might also deem the evidence unfairly prejudicial because of the dangers of hindsight bias or its emotional impact and exclude it on that basis.²⁷⁷ But for the most part, evidence that the decedent was unarmed seems sufficiently probative to be admissible in the absence of significant unfair prejudice.

Courts in these cases appear to base their relevance determinations not on thorough analysis of probative value but rather on policy considerations: The substantive law favors deference to the judgment of police officers,²⁷⁸ and in light of that policy, judges appear reluctant to acknowledge that hindsight evidence is relevant. This phenomenon is apparent in the *Sherrod* opinions. Judge Coffey, with his “walk a mile in his moccasins” quote, includes a determination of irrelevance as part of his deference to police.²⁷⁹ Judge Cummings, too, is tied to his narrative of a bad cop and an innocent victim.²⁸⁰ This explanation comes into starkest relief, however, when excessive force cases are juxtaposed with similar cases of self-defense.

In criminal homicide prosecutions where the defendant claims he was acting in self-defense because he believed the victim was drawing a weapon, courts almost always allow evidence that the victim was not actually armed.²⁸¹ Courts have deemed this evidence relevant and admissible for over a century.²⁸² Even though the question at issue is “whether

275. See, e.g., *Hudspeth v. City of Shreveport*, 270 F. App'x 332, 335 (5th Cir. 2008) (“The videotapes provide indisputable evidence of what transpired.”).

276. But see Dan M. Kahan, David A. Hoffman & Donald Braman, *Whose Eyes Are You Going to Believe? Scott v. Harris and the Perils of Cognitive Illiberalism*, 122 *Harv. L. Rev.* 837, 864–81, 904 (2009) (demonstrating people may divide in their perceptions of video evidence based on cultural predispositions and suggesting judges refrain from disregarding any group’s perspective).

277. See, e.g., *Sherrod*, 856 F.2d at 814 (Flaum, J., dissenting) (saying he would have held evidence should have been excluded under Rule 403 had objection been preserved).

278. See *Graham v. Connor*, 490 U.S. 386, 396–97 (1989) (setting out “reasonableness” standard).

279. *Sherrod*, 856 F.2d at 806 (majority opinion).

280. See *id.* at 808 (Cummings, J., dissenting) (“The record presents a bad cop with a history of unnecessary violence, acting in violation of police procedure, confronting an innocent victim.”).

281. *Little v. Grace*, No. 06-CV-2544, 2007 WL 2306619, at *4 (E.D. Pa. Aug. 8, 2007) (“Evidence a victim was unarmed does not make an assailant’s belief of impending harm *per se* unreasonable, but courts weigh the evidence heavily when determining if a defendant’s belief was reasonable.”).

282. See *James v. State*, 52 So. 840, 841 (Ala. 1910) (“The defendant’s evidence tended to show that deceased advanced, without provocation, upon her with his hand in the direction of his hip pocket. The state was well entitled, as bearing on that issue, to adduce evidence that deceased had no weapon of any kind.”); see also *Thompson v. State*, 29 S.W.2d 343, 345 (Tex. Crim. App. 1929) (citing “long-established rule that the state may

the defendant with sufficient reason believed that [the victim] was attempting to draw a weapon,” evidence that the victim was unarmed has “some tendency to contradict the defendant’s story that he was reaching as if for a weapon when he was shot.”²⁸³ The D.C. Circuit, charged by the Supreme Court with deciding when “uncommunicated threats” are admissible, concluded that evidence that a morgue attendant found an open penknife in the victim’s pocket after the homicide was admissible. “It is of course true that the apparent conduct of the deceased at the time of the homicide, rather than any concealed plan he may have had, bears directly on the question whether the accused acted in self-defense,” the court wrote, “but evidence that the deceased had a concealed plan of attack bears on the question what his apparent conduct was.”²⁸⁴ These courts are all correct: The fact that the deceased was unarmed is, of course, relevant and admissible. The problem is that courts fail to apply the same reasoning to excessive force cases. By applying a sound theory of hindsight evidence, courts can avoid this inconsistency.

B. *Liquidated Damages*

Although policy considerations might animate courts’ reluctance to admit hindsight evidence in the excessive force context, they likely fail to explain a similar dispute in one area of contract law. Parties to a contract may insert a liquidated damages clause stipulating the amount of damages they will pay in the event of a breach. Courts typically enforce these provisions unless they are penalty clauses in disguise, designed not to compensate but to prevent the parties from breaching by threatening punishment.²⁸⁵ This determination often turns on whether the stipulated amount was a reasonable forecast of actual damages.²⁸⁶ This section examines whether and when courts should use actual damages as hindsight evidence to assess whether the contract’s forecast was reasonable, concluding that this information is often valuable. There has

prove that deceased was unarmed or owned no pistol as disproving the claim of the accused that he made a demonstration as if to draw a gun”); *State v. Arrington*, 106 S.E. 445, 447 (W. Va. 1921) (“[W]here self-defense is an issue in a trial for homicide, evidence that the deceased was unarmed at the time of the killing is admissible.”); E.T. Branch, Branch’s Annotated Penal Code of the State of Texas § 1931 (1916) (“The State is entitled to prove that the deceased was unarmed at the time of the homicide in support of the theory that he was not the aggressor or that he was not reaching for a weapon when he was killed.”).

283. *State v. Abrams*, 223 P. 301, 302 (Kan. 1924).

284. *Griffin v. United States*, 183 F.2d 990, 992 (D.C. Cir. 1950).

285. 172 *Van Duzer Realty Corp. v. Globe Alumni Student Assistance Ass’n*, 25 N.E.3d 952, 957 (N.Y. 2014); see also Charles J. Goetz & Robert E. Scott, *Liquidated Damages, Penalties and the Just Compensation Principle*, 77 *Colum. L. Rev.* 554, 555 (1977) (“[C]ourts have . . . refused enforcement where the clause agreed upon is held to be *in terrorem* . . .”).

286. Restatement (Second) of Contracts § 356 (Am. Law Inst. 1981); Goetz & Scott, *supra* note 285, at 554.

been some limited debate in the literature about the use of outcomes in these cases, with some scholars suggesting that outcome information is irrelevant or improper,²⁸⁷ some noting only that hindsight bias poses a problem in these cases and judges should compensate for it,²⁸⁸ and at least one author indicating that disproportionate actual damages can evince *ex ante* unreasonableness.²⁸⁹ This section advances that discussion.

Liquidated damages clauses offer several legitimate benefits. They allow parties to manage their risk, they circumvent the costs of haggling over consequential damages in court, and they permit parties to account for intangible damages.²⁹⁰ But stipulated damages clauses also may stem from a disparity in bargaining power and may function as improper *in terrorem* clauses that deter the weaker party from breaching by threatening punishment.²⁹¹

Courts will enforce a liquidated damages clause only if it is reasonable, as measured by an objective standard.²⁹² They apply a two-pronged test, asking first, whether the amount of stipulated damages is a reasonable forecast of just compensation and second, whether the harm caused by the breach is difficult to measure.²⁹³

Courts and commentators mostly agree that reasonableness should be judged at the time the contract was formed, not after breach and

287. See Ann Morales Olazábal, *Formal and Operative Rules in Overliquidation Per Se Cases*, 41 *Am. Bus. L.J.* 503, 522 (2004) (“Clearly the use of hindsight in the process of judicial determination of reasonableness runs contrary to principles of party autonomy, permitting judicial paternalism in the form of voiding a clause based on just compensation grounded in *ex post* facts and circumstances.”).

288. See Robert A. Hillman, *The Limits of Behavioral Decision Theory in Legal Analysis: The Case of Liquidated Damages*, 85 *Cornell L. Rev.* 717, 735–36 (2000) (discussing judges’ cognitive processes and potential remedies); Jeffrey J. Rachlinski, *The “New” Law and Psychology: A Reply to Critics, Skeptics, and Cautious Supporters*, 85 *Cornell L. Rev.* 739, 761 (2000) (responding to Professor Hillman’s suggestions).

289. See Melvin Aron Eisenberg, *The Limits of Cognition and the Limits of Contract*, 47 *Stan. L. Rev.* 211, 234 (1995) (“A second-look standard for liquidated damages provisions is justified not because a second look may show that a provision was unconscionable, but because it may show that the provision was in all likelihood the product of defective cognition.”).

290. *Wassenaar v. Panos*, 331 N.W.2d 357, 362 (Wis. 1983).

291. *Id.*; Goetz & Scott, *supra* note 285, at 555.

292. *Wassenaar*, 331 N.W.2d at 362–63.

293. See, e.g., *Magill v. Watson*, 409 S.W.3d 673, 679 (Tex. App. 2013) (explaining two-part test); *Walter Implement, Inc. v. Focht*, 730 P.2d 1340, 1343 (Wash. 1987) (same). This formulation follows both the UCC and Restatement (Second) of Contracts. See Restatement (Second) of Contracts § 356 (Am. Law Inst. 1981) (requiring amount of stipulated damages to be “reasonable in the light of the anticipated or actual loss caused by the breach and the difficulties of proof of loss”); see also U.C.C. § 2-718(1) (Am. Law Inst. 2005) (requiring amount to be “reasonable in the light of the anticipated or actual harm caused by the breach, the difficulties of proof of loss, and the inconvenience or nonfeasibility of otherwise obtaining an adequate remedy”).

injury.²⁹⁴ But courts disagree about the role of actual damages in determining whether anticipated damages were reasonable. In particular, where the nonbreaching party hardly suffers at all, should courts consider the absence of damages in evaluating *ex ante* reasonableness? This situation may arise, for example, when a buyer breaks a real estate contract and the seller resells at a higher price;²⁹⁵ when an employer breaks an employment contract and the employee quickly finds a new job at a higher salary;²⁹⁶ or when a manufacturer delivers equipment late but it would not have been used in the interim.²⁹⁷

Most commentators and judges divide judicial responses into two categories: the prospective or single-look approach and the retrospective or second-look approach.²⁹⁸ Courts using the prospective approach evaluate reasonableness at the time of the contract²⁹⁹ “only in the context of information available at the time of contracting.”³⁰⁰ Courts using the retrospective approach compare actual losses sustained to the stipulated damages and evaluate reasonableness in light of any large discrepancies between them.³⁰¹ But a better taxonomy would divide jurisdictions into three nonmutually exclusive groups based on whether they evaluate reasonableness from an *ex ante* or *ex post* perspective *and* whether they look to actual damages in their evaluation.³⁰²

First are the courts that evaluate reasonableness prospectively and “ignore or exclude evidence of actual damages sustained,”³⁰³ as they

294. See *Wallace Real Estate Inv., Inc. v. Groves*, 881 P.2d 1010, 1016 (Wash. 1994) (citing commentators); 24 Richard A. Lord, *Williston on Contracts* § 65:17 (4th ed. 2002) (“The more popular view is that the reasonableness of a liquidated damages clause should be determined as of the time the contract was executed, not with the benefit of hindsight.”).

295. See, e.g., *Kelly v. Marx*, 694 N.E.2d 869, 870–71 (Mass. 1998) (stating buyer failed to buy house for \$355,000 and seller sold it to third party for \$360,000), *rev'd*, 705 N.E.2d 1114 (Mass. 1999).

296. See, e.g., *Guiliano v. Cleo, Inc.*, 995 S.W.2d 88, 92–93 (Tenn. 1999) (noting plaintiff was relieved from duties of job with \$103,000 salary and accepted new employment at \$110,000 salary).

297. See, e.g., *United States v. Bethlehem Steel Co.*, 205 U.S. 105, 117 (1907) (noting War Department could not have used gun carriages had they been delivered on time).

298. See, e.g., *Kelly*, 694 N.E.2d at 873–74 (enumerating courts taking each approach); *Guiliano*, 995 S.W.2d at 98–99 (discussing both approaches); 24 Lord, *supra* note 294, § 65:17 (same); *Olazábal*, *supra* note 287, at 519–22 (same).

299. See, e.g., *Kelly*, 694 N.E.2d at 870 (describing “single look’ approach”).

300. *Olazábal*, *supra* note 287, at 520.

301. *Id.* at 521; see also *Guiliano*, 995 S.W.2d at 99 (“If the liquidated sum greatly exceeds the amount of actual damages, then courts . . . will treat the estimated sum as a penalty and will limit recovery to the actual damages.”).

302. See, e.g., Ian R. Macneil, *Power of Contract and Agreed Remedies*, 47 *Cornell L.Q.* 495, 504 (1962) (noting three possible approaches to actual losses: using them only to determine what was reasonable at the time of contracting, not requiring plaintiff to show losses at all, or requiring reasonableness in relation to actual losses).

303. *Olazábal*, *supra* note 287, at 520.

believe “the amount of actual damages at the time of breach is of little or no significance to the recovery of liquidated damages.”³⁰⁴ This exclusionary group deems the hindsight evidence wholly irrelevant. For example, in *Frick Co. v. Rubel Corp.*, the trial court, evaluating the reasonableness of a liquidated damages clause, refused to admit evidence that the stipulated damages were grossly disproportionate to actual damages.³⁰⁵ Affirming, Judge Learned Hand wrote that although he disagreed, the other two members of the panel believed “that evidence as to the *actual* loss was not material to the issue of the losses in *contemplation*.”³⁰⁶ Thus, while Judge Hand would have considered hindsight evidence, the majority found it irrelevant to the prospective approach.

Second, some courts employ a retrospective approach that focuses on actual damages for purposes other than determining the reasonableness of *ex ante* damages. Courts concerned with the “justice and equity of enforcement”³⁰⁷ may decline to enforce liquidated damages clauses if they seem unfair or unreasonable *ex post*. The most common manifestation of this is the “no actual injury defense,”³⁰⁸ which the Restatement endorses.³⁰⁹ Courts that recognize this equitable defense will refuse to enforce a liquidated damages clause,³¹⁰ “no matter how reasonable the estimate of damages was at the time of contracting,”³¹¹ if the plaintiff did not actually suffer a loss.

Third, a number of courts, most notably the Wisconsin Supreme Court, use some variation of a prospective–retrospective approach, where they consider actual damages “as evidence helpful in determining what was reasonable at the time of contracting.”³¹² These courts care about *ex ante* reasonableness; their primary concern is not with equity or with

304. *Guiliano*, 995 S.W.2d at 99.

305. 62 F.2d 765, 767–68 (2d Cir. 1933).

306. *Id.* at 768. In *Frick*, the plaintiff sued when the defendant was not ready to accept plaintiff’s ice-making machines at the prescribed date. *Id.* at 765.

307. 11 Joseph M. Perillo, *Corbin on Contracts* § 58.11 (rev. ed. 2005) (citing cases from Connecticut, Illinois, Iowa, Minnesota, Missouri, New York, Ohio, and Pennsylvania).

308. Larry A. Dimatteo, *A Theory of Efficient Penalty: Eliminating the Law of Liquidated Damages*, 38 *Am. Bus. L.J.* 633, 663 (2001).

309. Restatement (Second) of Contracts § 356 cmt. b (Am. Law Inst. 1981) (“If, to take an extreme case, it is clear that no loss at all has occurred, a provision fixing a substantial sum as damages is unenforceable.”).

310. See *Norwalk Door Closer Co. v. Eagle Lock & Screw Co.*, 220 A.2d 263, 268 (Conn. 1966) (“[E]quitable principles will be invoked to deny recovery when the facts make it apparent that no damage has been suffered.”).

311. *Yockey v. Horn*, 880 F.2d 945, 953 (7th Cir. 1989).

312. *Wassenaar v. Panos*, 331 N.W.2d 357, 364 (Wis. 1983); see also 11 Perillo, *supra* note 307, § 58.11 (“Where the sum named in a contract to be paid on a breach is wholly disproportionate to the actual damages sustained, the courts will deem the parties to have intended to stipulate for a mere penalty to secure performance, and not for a liquidation of the damages.”); Macneil, *supra* note 302, at 505 (“The courts in many more cases, however, have stated that the actual loss was irrelevant, except perhaps as evidence on the question of reasonableness at the time of contracting.”).

principles of just compensation.³¹³ The idea here is that a very large difference between stipulated damages and actual damages tends to show that the stipulated damages amount was an unreasonable forecast, whereas a small difference between stipulated and actual damages tends to show that the forecast was *ex ante* reasonable.³¹⁴ Judge Hand would have taken this approach in *Frick*, using actual damages as evidence of contemplated damages.³¹⁵ Courts that use this combined approach rarely discuss precisely *how* actual damages bear on the reasonableness of the clause or how heavily they weight different disparities between stipulated and actual damages. Instead, courts typically say that if the two are “grossly disproportionate,” the clause is an unenforceable penalty.³¹⁶ These courts, then, use hindsight evidence in a rough way, without necessarily examining the specifics of the case closely. If the evidence is extremely strong, the judge will declare the clause *ex ante* unreasonable.

This Essay does not speak to the propriety of the second approach of using actual damages for purposes other than looking at *ex ante* reasonableness.³¹⁷ But for courts that are interested primarily in reasonableness at the time of contracting, this Essay’s framework suggests that between the first and third approach, the third is superior. Whether a liquidated damages clause constitutes a penalty is a question of law for a judge, but it is based on underlying facts and circumstances;³¹⁸ it is

313. But see *Wassenaar*, 331 N.W.2d at 366–67 (noting courts will refuse to enforce stipulated damages clause in absence of damages *both* because it would violate compensation principle and because harm is probative of reasonableness).

314. See *Zidell, Inc. v. Pac. N. Marine Corp.*, 744 F. Supp. 982, 987 (D. Or. 1990) (“[S]everal factors indicate that the \$1100 figure was a reasonable forecast of actual damages Third, \$1100 per day is not ‘grossly disproportionate’ to the fair market rate of \$833 for the period at the end of the lease.”).

315. *Frick Co. v. Rubel Corp.*, 62 F.2d 765, 768 (2d Cir. 1933).

316. *Truck Rent-A-Ctr., Inc. v. Puritan Farms 2d, Inc.*, 361 N.E.2d 1015, 1018 (N.Y. 1977) (“It is plain that a provision which requires, in the event of contractual breach, the payment of a sum of money grossly disproportionate to the amount of actual damages provides for penalty and is unenforceable.”); see also *Zidell*, 744 F. Supp. at 987 (“Defendant may satisfy this burden by showing that the liquidated damages are ‘grossly disproportionate’ to actual damages.”); *Wassenaar*, 331 N.W.2d at 364 (“If the damages provided for in the contract are grossly disproportionate to the actual harm sustained, the courts usually conclude that the parties’ original expectations were unreasonable.”).

317. For criticisms of this approach, see Olazábal, *supra* note 287, at 522 (referring to “judicial paternalism”); James Arthur Weisfield, Note, “Keep the Change!”: A Critique of the No Actual Injury Defense to Liquidated Damages, 65 Wash. L. Rev. 977, 978 (1990) (“The *Lind* defense of ‘no actual injury’ eradicates the benefits of liquidating damages.”); see also Goetz & Scott, *supra* note 285, at 557–58 (“We argue that agreed damage measures and *in terrorem* provisions represent, under many circumstances, the most efficient means by which parties can insure against the otherwise non-compensable consequences of breach.”).

318. See 172 *Van Duzer Realty Corp. v. Globe Alumni Student Assistance Ass’n*, 25 N.E.3d 952, 957 (N.Y. 2014) (noting this determination is a question of law); *Wassenaar*, 331 N.W.2d at 360–61 (concluding validity is a question of law for judge, typically “derived

rarely a pure question of judgment. The judge frequently lacks the specialized knowledge of one or both parties who enter into a contract—the parties' experience in their respective industries may give them special insight into what level of damages would compensate and what level would deter. Further demonstrating that reasonableness is often based on facts that merit evidentiary proof, courts sometimes take expert evidence on reasonableness of a liquidated damages clause.³¹⁹ And actual damages are indeed probative of whether the stipulated amount was a reasonable forecast. A knowledgeable party has some ability to predict damages accurately, so reasonable stipulations are more likely to approximate actual damages than unreasonably low stipulations. Courts that universally deem this evidence irrelevant are mistaken.

Although any disparity between actual and stipulated damages can be probative of *ex ante* unreasonableness, and in some cases a judge may be so well informed that the question reduces to one of pure judgment, the oft-used rough rule that grossly disproportionate stipulated damages will not be enforced is a reasonable and flexible heuristic. A slight difference between stipulated and actual damages is minimally probative, as it is difficult to predict damages with absolute precision. But a very large difference—particularly one that sheds light on foreseeable reasons why damages might have been smaller than those stipulated—may be highly probative. Further, courts may rationally conclude that the grossly disproportionate rule is economically efficient, as it mitigates the need to hire and listen to experts.³²⁰

Judges, not juries, decide whether liquidated damages clauses are enforceable, so Rule 403 does not apply to these cases. However, judges should remain aware of the effects of hindsight bias and outcome bias and do their best to debias themselves, writing down alternative outcomes in order to “consider the opposite.”³²¹ Their fear of hindsight

from a resolution of disputed facts or inferences” including “existence and extent of the anticipated and actual injury to the nonbreaching party”).

319. See, e.g., *Wallace Real Estate Inv., Inc. v. Groves*, 881 P.2d 1010, 1018 (Wash. 1994) (“At trial, an economics professor testified to the reasonableness of the \$15,000 extension payments.”).

320. The grossly disproportionate approach is least useful in cases where damages will be either very large or negligible, with little in between. For example, if a government places an advance order for wartime equipment, a late shipment will cause substantial loss, unless the war has ended before the delivery date. See, e.g., *United States v. Bethlehem Steel Co.*, 205 U.S. 105, 121 (1907) (ordering enforcement of liquidated damages provision even though war ended and “practically no damage accrued to the Government”). In that case, the end of war does tend to prove that peace was probable by the due date, and the defendant may plausibly argue that the possibility of peace should have been incorporated into the liquidated damages clause. But the absence of damages says nothing about what the damages would have been had the war continued. In that case, the “grossly disproportionate” standard makes little sense; it should instead be used when the *amount* of foreseeable damages is at issue.

321. See *supra* section IV.B.1.

bias, however, should not prevent them from extracting useful information from hindsight evidence.

C. *Individuals with Disabilities Education Act*

Juries are also absent from proceedings under the IDEA, but in these cases, as in liquidated damages cases, judges and hearing officers can use the hindsight-evidence framework to their benefit. The IDEA³²² requires local educational agencies to provide each disabled child with a program that is “reasonably calculated to enable the child to receive educational benefits.”³²³ This section addresses a set of cases where the school district or the child’s parents attempt to introduce evidence of the student’s progress or lack thereof—hindsight evidence—to demonstrate whether the district met this standard. There is an active debate in the courts over how to use this hindsight evidence, with the Second Circuit weighing in most recently, deeming it categorically inadmissible.³²⁴ Only one piece of scholarship has addressed the question directly.³²⁵ This section concludes that in many cases, but not all, this evidence is relevant and judges and administrative hearing officers should put it to cautious use.

Under the IDEA, each student with disabilities receives an individualized education program (IEP),³²⁶ which sets forth annual goals and the services the child will receive.³²⁷ The IEP is prepared at a meeting of the child’s parents, the child’s teacher, a representative of the local educational agency who is qualified to provide specially designed instruction, and sometimes the child himself and other individuals familiar with the child.³²⁸ This IEP must provide education and services that

322. Pub. L. No. 101-476, 104 Stat. 1141 (1990) (codified as amended at 20 U.S.C. §§ 1400–1482 (2012)).

323. *Bd. of Educ. v. Rowley*, 458 U.S. 176, 207 (1982) (discussing IDEA’s predecessor statute, the Education of the Handicapped Act). The IDEA and its predecessors condition certain federal financial assistance on a state establishing a qualified policy that provides a free appropriate public education to students with disabilities. *Id.* at 180–81.

324. See *R.E. v. N.Y.C. Dep’t of Educ.*, 694 F.3d 167, 187 (2d Cir. 2012) (“Parents who end up placing their children in public school cannot later use evidence that their child did not make progress under the IEP in order to show that it was deficient from the outset.”).

325. See Fan, *supra* note 10, at 1539–44 (arguing “actual-progress evidence” should be admitted as relevant); cf. Martin A. Kotler, *Distrust and Disclosure in Special Education Law*, 119 Penn St. L. Rev. 485, 524–25 (2014) (suggesting that improvement at private-school placement should be a strong indicator of whether placement was appropriate).

326. See Fan, *supra* note 10, at 1503 n.3 (“The reader should be aware that neither the IDEA nor its previous incarnations lack acronyms.”).

327. See *R.E.*, 694 F.3d at 182 (detailing one child’s IEP).

328. 20 U.S.C. § 1414(d)(1)(B) (2012); *Bd. of Educ. v. Rowley*, 458 U.S. 176, 182 (1982).

are “reasonably calculated to enable the child to receive educational benefits.”³²⁹

Parents who believe their child’s IEP is inappropriate may institute an administrative due process hearing with an impartial hearing officer,³³⁰ who receives documentary evidence and testimony and decides whether the IEP was appropriate. A party unhappy with the administrative decision may file a civil action in federal district court.³³¹ The court reviews the administrative record, hears any additional evidence a party wishes to present, and then bases its decision on the preponderance of the evidence.³³² The court determines whether the IEP meets the IDEA’s minimal requirements, but “questions of methodology are for resolution by the States.”³³³ In addition, the court must afford the administrative proceedings “due weight.”³³⁴

Parties may seek to admit hindsight evidence at the administrative or district-court level. Parents trying to prove that the IEP was inappropriate may argue that lack of progress demonstrates that the IEP was *ex ante* unreasonable.³³⁵ Conversely, a school district may attempt to introduce evidence of progress to show *ex ante* reasonableness.³³⁶

IDEA litigation is a tricky context for analysis of an evidence question. Evidence is largely introduced at the administrative hearing, where the Federal Rules of Evidence do not apply.³³⁷ In the district court there is no jury, and the judge considers the case on the administrative

329. *Rowley*, 458 U.S. at 207. The IDEA also has a number of procedural requirements for preparing the IEP. *Id.* at 206.

330. See 20 U.S.C. §§ 1415(b)(6), 1415(f)(1), 1415(i)(2)(A) (listing procedural rights). States may implement a one-tier system with just an administrative hearing officer or a two-tier system where a local hearing officer’s decision is reviewed at the state level. See *id.* § 1415(f)(1); see also Andriy Krahmal, Perry A. Zirkel & Emily J. Kirk, “Additional Evidence” Under the Individuals with Disabilities Education Act: The Need for Rigor, 9 *Tex. J.C.L. & C.R.* 201, 202 (2004) (noting choice between one- and two-tier systems).

331. 20 U.S.C. § 1415(g).

332. *Id.* § 1415(i)(2)(C).

333. *Rowley*, 458 U.S. at 208.

334. *Id.* at 206.

335. See, e.g., *Lessard v. Wilton-Lyndeborough Coop. Sch. Dist.*, 518 F.3d 18, 29 (1st Cir. 2008) (noting parents’ argument that methodology in IEP was inappropriate because daughter made little progress).

336. See, e.g., *CJN v. Minneapolis Pub. Sch.*, 323 F.3d 630, 638 (8th Cir. 2003) (highlighting district’s evidence of student progressing academically at average rate). In addition, because parents seeking tuition reimbursement for private school must show that their alternative placement was proper, evidence of progress may show that their own school choice was reasonable. See *Sch. Comm. of Burlington v. Dep’t of Educ.*, 471 U.S. 359, 370 (1985) (discussing retroactive reimbursement); *Berger v. Medina City Sch. Dist.*, 348 F.3d 513, 522 (6th Cir. 2003) (recognizing evidence of progress in private placement as relevant but not sufficient to show placement was appropriate).

337. Susan G. Clark, *Judicial Review and the Admission of “Additional Evidence” Under the IDEA: An Unusual Mixture of Discretion and Deference*, 201 *Educ. L. Rep.* 823, 825 (2005).

record, although he may consider “additional evidence” presented by the parties.³³⁸ But the federal courts may reverse administrative determinations that improperly rely on irrelevant evidence,³³⁹ so the Federal Rules’ principles of relevance and probative value ultimately guide evidence requirements at the administrative level.³⁴⁰ So the following analysis should assist judges, hearing officers, and parties arguing before them.

Courts have split—fanned out, really—over whether adjudicators should consider hindsight evidence. Courts largely agree that an IEP should be judged prospectively, at the time it was created,³⁴¹ but they disagree on the implications of this *ex ante* perspective for hindsight evidence.³⁴²

On the most restrictive end, the Second Circuit recently stated—in dicta, but in no uncertain terms—that because an IEP should be judged prospectively, parents “cannot . . . use evidence that their child did not make progress under the IEP in order to show that it was deficient from the outset.”³⁴³ The Seventh and Ninth Circuits have also severely cabined the use of hindsight evidence.³⁴⁴

338. 20 U.S.C. § 1415(i)(2)(C) (2012).

339. See, e.g., *D.F. ex rel. N.F. v. Ramapo Cent. Sch. Dist.*, 430 F.3d 595, 599 (2d Cir. 2005) (“[W]e cannot determine at this stage whether the approach taken by *both* the district court and the [state review officer] was improper.”).

340. Fan, *supra* note 10, at 1517; see also Perry A. Zirkel, Zorka Karanxha & Anastasia D’Angelo, *Creeping Judicialization in Special Education Hearings?: An Exploratory Study*, 27 *J. Nat’l Ass’n Admin. L. Judiciary* 27, 44 (2007) (finding hearings are becoming increasingly judicialized); Perry A. Zirkel, *Expert Witnesses in Impartial Hearings Under the Individuals with Disabilities Education Act*, 298 *Educ. L. Rep.* 648, 651 (2014) (setting forth criteria of reliability and relevance for expert testimony); cf. N.Y. State Educ. Dep’t Office of Vocational & Educ. Servs. for Individuals with Disabilities, *Impartial Hearing Process for Students with Disabilities* (2001), <http://www.p12.nysed.gov/specialed/publications/policy/ihprocessguide.htm> [<http://perma.cc/66AR-NWJ7>] (allowing hearing officer to exclude irrelevant or duplicative witnesses).

341. See, e.g., *Roland M. v. Concord Sch. Comm.*, 910 F.2d 983, 992 (1st Cir. 1990) (“An IEP is a snapshot, not a retrospective. In striving for ‘appropriateness,’ an IEP must take into account what was, and was not, objectively reasonable when the snapshot was taken, that is, at the time the IEP was promulgated.”).

342. For a more comprehensive analysis of the circuits’ case law, see Fan, *supra* note 10, at 1526–39.

343. *R.E. v. N.Y.C. Dep’t of Educ.*, 694 F.3d 167, 187 (2d Cir. 2012). The Second Circuit had previously punted on this question, remanding it to the district court and hinting that the First, Third, and Ninth Circuits had “held that inquiry into whether an IEP is valid is a necessarily *prospective* analysis, and that consideration of proof of whether an IEP meaningfully contributed to the child’s education is not altogether proper.” *D.F.*, 430 F.3d at 598.

344. See *M.B. ex rel. Berns v. Hamilton Se. Sch.*, 668 F.3d 851, 863 (7th Cir. 2011) (holding it was reasonable to discount report showing regression over school year because IEPs must be evaluated only prospectively, not in hindsight); *J.W. ex rel. J.E.W. v. Fresno Unified Sch. Dist.*, 626 F.3d 431, 454 (9th Cir. 2010) (rejecting student’s argument contending lack of progress shows IEP was inadequate because it “ignores the ‘snapshot’

The First and Third Circuits take an intermediate stance. The First Circuit has acknowledged that actual progress may demonstrate that an IEP is appropriate but a failure to progress does not demonstrate that an IEP was inappropriate.³⁴⁵ The Third Circuit has issued the most thoughtful opinions on the hindsight-evidence question, concluding that “Monday Morning Quarterbacking” is inappropriate but courts may consider ex post progress only for the purpose of determining whether the original IEP was reasonably calculated to afford some benefit.³⁴⁶

The Sixth, Eighth, and particularly the Fourth and Fifth Circuits have been more liberal about allowing hindsight evidence. For example, the Fourth Circuit has said that “progress, or the lack thereof,” is important but not dispositive.³⁴⁷ And the Fifth Circuit looks to demonstrated benefits of an IEP to determine whether it is appropriate, calling progress a “critical factor[.]”³⁴⁸

But *is* this hindsight evidence probative in these cases? Under the framework set out herein, it must bear on an unresolved factual question, on which the parties may seek to persuade the decisionmaker. If the other evidence in the case puts the judge or hearing officer in the exact same place as the IEP team, then reasonableness is a question of judgment, not provable fact, and it would be improper to judge the district in light of the outcome.³⁴⁹ This is rarely the case—hindsight evidence is typically probative.

In many actions, there are two related, disputed facts that hindsight evidence can address: the child’s needs, as gleaned from personal interaction with him or her,³⁵⁰ and the range of educational methods and

evaluation of an IEP”); *Adams ex rel. Adams v. Oregon*, 195 F.3d 1141, 1149 (9th Cir. 1999) (holding district court should not have asked whether program was adequate in light of student’s progress).

345. *Lessard v. Wilton-Lyndeborough Coop. Sch. Dist.*, 518 F.3d 18, 29 (1st Cir. 2008).

346. *Fuhrmann ex rel. Fuhrmann v. E. Hanover Bd. of Educ.*, 993 F.2d 1031, 1040 (3d Cir. 1993) (Garth, J.); see also *Susan N. v. Wilson Sch. Dist.*, 70 F.3d 751, 762 (3d Cir. 1995) (stating Judge Garth’s opinion in *Fuhrmann* controls); *Carlisle Area Sch. v. Scott P. ex rel. Bess P.*, 62 F.3d 520, 530 (3d Cir. 1995) (“[A]ny lack of progress . . . does not render that IEP inappropriate”).

347. *M.S. ex rel. Simchick v. Fairfax Cty. Sch. Bd.*, 553 F.3d 315, 327 (4th Cir. 2009); see also *MM ex rel. DM v. Sch. Dist.*, 303 F.3d 523, 532 (4th Cir. 2002) (calling actual progress an “objective factor”).

348. *Hous. Indep. Sch. Dist. v. VP ex rel. Juan P.*, 582 F.3d 576, 588 (5th Cir. 2009); see also *Cypress-Fairbanks Indep. Sch. Dist. v. Michael F.*, 118 F.3d 245, 253 (5th Cir. 1997). For the Sixth and Eighth Circuits’ views, see *Berger v. Medina City Sch. Dist.*, 348 F.3d 513, 522 (6th Cir. 2003); *CJN v. Minneapolis Pub. Sch.*, 323 F.3d 630, 642 (8th Cir. 2003). The D.C. Circuit recently used actual-progress evidence to conclude that parents’ alternative private placement was “reasonably calculated” as well as “necessary” to benefit the student. *Leggett v. District of Columbia*, 793 F.3d 59, 70–72 (D.C. Cir. 2015).

349. See *supra* section II.B (explaining when hindsight evidence is not relevant).

350. The IEP team includes the child’s teacher and parents and potentially others familiar with the child, and those interactions inform the IEP team’s decisionmaking. See 20 U.S.C. § 1414(d)(1)(B) (2012); *Laura F. Rothstein & Julia Rothstein, Disabilities and*

services that an educational expert could deem appropriate to address the individual child's needs.³⁵¹ The answers to both of these questions affect the student's progress. If the IEP was tailored to an accurate understanding of the child and his or her needs, it is more likely to produce progress than an IEP tailored to a flawed portrait of the child. And if the IEP provided for educational methods and services that are acceptable to experts, it is more likely to produce progress than an IEP drawn inexpertly. So evidence of progress is relevant to finding these facts; failure to progress tends to prove that the school did not appropriately identify or address the child's needs.

Hearing officers regularly receive expert testimony on precisely these two questions: what needs the child evinced in personal interaction³⁵² and what educational services were appropriate for the child.³⁵³ Familiarity with the child and specialized expertise are the two central criteria by which hearing officers and judges determine the weight of expert testimony.³⁵⁴ This demonstrates that professional judgment is an

the Law § 2:19 (4th ed. 2015) (noting parents are important to IEP development process because they are so familiar with "personalities, habits, abilities, interests, likes and dislikes, problems, and needs of their children"). The judge does not have this level of interaction.

351. The IEP team includes educational experts, including a representative who is "qualified" to provide or supervise specialized instruction and someone "who can interpret the instructional implications of evaluation results." See 20 U.S.C. § 1414(d)(1)(B). Although judges may not choose between accepted methodologies, in most circuits they may determine which educational approaches meet the requirement of bare adequacy. See, e.g., *Lessard v. Wilton-Lyndeborough Coop. Sch. Dist.*, 518 F.3d 18, 29 (1st Cir. 2008) ("[C]ourts are entrusted with ascertaining the adequacy of an IEP's educational components but not with weighing the comparative merit of the components when stacked against other heuristic methods."); *Roland M. v. Concord Sch. Comm.*, 910 F.2d 983, 992 (1st Cir. 1990) (approving of courts' reluctance to overturn state education agency's judgment call "at least where it can be shown that 'the IEP proposed by the school district is based upon an accepted, proven methodology'" (quoting *Lachman v. Ill. State Bd. of Educ.*, 852 F.2d 290, 297 (7th Cir. 1988))).

352. See, e.g., *S.H. v. State-Operated Sch. Dist. of Newark*, 336 F.3d 260, 267 (3d Cir. 2003) (noting ALJ concluded expert was sufficiently informed about student to account for her needs); *Gellert v. D.C. Pub. Sch.*, 435 F. Supp. 2d 18, 25 (D.D.C. 2006) (noting expert "who conducted weekly group and individual therapy sessions" with child for ten months testified child had trouble tuning out extraneous noise).

353. See, e.g., *Neosho R-V Sch. Dist. v. Clark*, 315 F.3d 1022, 1025, 1028 (8th Cir. 2003) (noting expert testimony that student required behavior management plan and papers attached to IEP were insufficient); *MM ex rel. DM v. Sch. Dist.*, 303 F.3d 523, 531-32 (4th Cir. 2002) (noting hearing officer heard conflicting testimony as to whether IEP was appropriate and decided it was not); *Adams ex rel. Adams v. Oregon*, 195 F.3d 1141, 1149-50 (9th Cir. 1999) (discussing adequacy of district's program and parents' proposed program).

354. See, e.g., *J.W. ex rel. J.E.W. v. Fresno Unified Sch. Dist.*, 626 F.3d 431, 445 (9th Cir. 2010) (comparing witnesses based on familiarity and expertise); *Heather S. v. Wisconsin*, 125 F.3d 1045, 1057 (7th Cir. 1997) ("We note that the deference is to trained educators, not necessarily psychologists."); *Oberti ex rel. Oberti v. Bd. of Educ. of Clementon Sch. Dist.*, 995 F.2d 1204, 1210 (3d Cir. 1993) ("The ALJ disregarded Dr. McGregor's testimony because, unlike the School District's witnesses, she did not have daily experience with

important consideration in IDEA cases—adjudicators may resolve the case based on how a reasonable expert would have processed and responded to the available information. Hindsight evidence should be considered to the extent that it is probative of how a reasonable expert would have assessed and responded to the child's needs.

One caveat: Although this hindsight evidence can always be relevant at the administrative-hearing level, a number of circuits have cabined federal court *review* of administrative decisions so much that, in certain cases, hindsight evidence may be reduced to irrelevance. The Second Circuit, for example, limits federal court review “to an examination of “objective” evidence indicating whether the child is likely to make progress or regress under the proposed plan.”³⁵⁵ The court must also defer completely to administrative determinations of educational methodology.³⁵⁶ Hindsight evidence becomes irrelevant under this stifling standard because the court may not evaluate reasonableness. There is an escape hatch, however: If the district court hears “additional evidence,” it may accord the administrator's findings less deference.³⁵⁷ So a court that receives additional evidence on the IEP's appropriateness may weigh hindsight evidence in its determination.³⁵⁸

In sum, the Third Circuit, which considers hindsight evidence only to the extent that it is relevant to determining whether the IEP was reasonably calculated to enable the child to benefit, has the right idea.³⁵⁹ But there is even more to it than that: Only when hindsight evidence bears on a fact that the IEP team had access to but the judge does not—

Rafael.”); K.S. ex rel. P.S. v. Fremont Unified Sch. Dist., 679 F. Supp. 2d 1046, 1056 (N.D. Cal. 2009) (approving of ALJ's decision to afford more weight to testimony of experts who took child's specific characteristics into account); A.S. ex rel. P.B.S. v. Bd. of Educ., 245 F. Supp. 2d 417, 429–30 (D. Conn. 2001), *aff'd mem.*, 47 F. App'x 615 (2d Cir. 2002) (weighing expert testimony in accordance with special education expertise and familiarity with student); see also Zirkel, *supra* note 340, at 652 (“The basic criteria are: 1. the relevant specialized expertise 2. the relevant familiarity with the child . . .” (footnote omitted)).

355. M.H. ex rel. P.H. v. N.Y.C. Dep't of Educ., 685 F.3d 217, 242–43 (2d Cir. 2012) (quoting *Walczak v. Fla. Union Free Sch. Dist.*, 142 F.3d 119, 130 (2d Cir. 1998)). This “objective evidence” includes previous test scores and “similar objective criteria.” *Walczak*, 142 F.3d at 130.

356. M.H., 685 F.3d at 243 (citing *Grim v. Rhinebeck Cent. Sch. Dist.*, 346 F.3d 377, 383 (2d Cir. 2003)). Some circuits, like the Ninth, give courts discretion as to how much deference to give state educational agencies. See *JW.*, 626 F.3d at 438 (noting level of deference is a matter for judicial discretion). Most use a “modified *de novo* review” standard, in which the administrator's factual findings are “considered *prima facie* correct.” *S.H.*, 336 F.3d at 270 (agreeing with Fourth, Sixth, and Tenth Circuits and adopting “modified *de novo*” standard of review in Third Circuit).

357. M.H., 685 F.3d at 244.

358. See, e.g., *Cypress-Fairbanks Indep. Sch. Dist. v. Michael F. ex rel. Barry F.*, 118 F.3d 245, 253–56 (5th Cir. 1997) (determining additional evidence, including of student progress, supported district-court decision contrary to administrative decision).

359. See Fan, *supra* note 10, at 1540 (recommending courts follow Third Circuit).

most frequently the child's apparent needs or expert knowledge—is evidence of progress probative. Of course, judges and hearing officers who encounter this evidence should account for hindsight bias when they evaluate probative value. But considering student progress when it is relevant is not Monday-morning quarterbacking—it is weighing legitimate evidence.

D. *Judgment as a Matter of Law*

This final example deviates even further from the Rule 403 template. In fact, it does not concern evidence at all. Instead, this example shows how judges can apply an understanding of the probative value of hindsight evidence to other aspects of legal decisionmaking.

Rule 50 of the Federal Rules of Civil Procedure allows a party to move for judgment as a matter of law before the case is submitted to a jury³⁶⁰ and to renew its motion after the jury returns a verdict.³⁶¹ A judge may grant the motion if he or she concludes that “a *reasonable jury* would not have a legally sufficient evidentiary basis to find” for the nonmoving party.³⁶² Although the standard is formally the same before and after the jury verdict, one circumstance changes: After a verdict, the judge, who must decide what a “reasonable jury” could do, knows what an *actual* jury did. Should the court use that information as a fact-finder might use hindsight evidence?

In a rift previously unrecognized in the literature,³⁶³ circuits have divided on whether to treat a Rule 50(b) motion, made after a verdict, differently from a Rule 50(a) motion, made before the verdict. Most circuits handle the two motions identically, noting that the two are evaluated under the same standard.³⁶⁴ In particularly clear language, the

360. Fed. R. Civ. P. 50(a).

361. Fed. R. Civ. P. 50(b).

362. Fed. R. Civ. P. 50(a) (emphasis added).

363. Professor Martin Redish has noted that as a practical matter, courts might be more reluctant to grant judgment as a matter of law than to grant summary judgment. Martin H. Redish, Summary Judgment and the Vanishing Trial: Implications of the Litigation Matrix, 57 *Stan. L. Rev.* 1329, 1341 (2005).

364. See, e.g., *Hubbard v. BankAtlantic Bancorp, Inc.*, 688 F.3d 713, 716, 724 (11th Cir. 2012) (stating standard is same before and after jury verdict); *Morningstar v. Worthy*, 454 F. App'x 391, 400 (6th Cir. 2011) (citing one standard for Rule 50(a) and Rule 50(b) motions); *Fisher v. City of San Jose*, 509 F.3d 952, 957 (9th Cir. 2007), rev'd, 558 F.3d 1069 (9th Cir. 2009) (en banc) (“It is thus Rule 50(a) that sets out the standard for granting Rule 50(b) motions . . .”); *ABT Bldg. Prods. Corp. v. Nat'l Union Fire Ins. Co. of Pittsburgh*, 472 F.3d 99, 113 (4th Cir. 2006) (stating Rule 50(a) standard after Rule 50(b) motion was denied); *Raiczyk v. Ocean Cty. Veterinary Hosp.*, 377 F.3d 266, 269 (3d Cir. 2004) (stating standard for Rule 50(b) without noting difference between two standards); *Acevedo-Garcia v. Monroig*, 351 F.3d 547, 565 (1st Cir. 2003) (noting one standard for “reviewing the denial of a motion for directed verdict or for judgment notwithstanding the verdict” (quoting *Gutierrez-Rodriguez v. Cartagena*, 882 F.2d 553, 558 (1st Cir. 1989))); *Weese v. Schukman*, 98 F.3d 542, 547 (10th Cir. 1996) (“[W]e review a judgment as a

Eleventh Circuit has stated, “When a court considers a motion for judgment as a matter of law—even after the jury has rendered a verdict—only the sufficiency of the evidence matters. The jury’s findings are irrelevant.”³⁶⁵ The Second and Fifth Circuits, however, use a slightly different standard after a jury has returned a verdict. In the Second Circuit, a Rule 50 movant’s burden “is particularly heavy after the jury has . . . actually returned its verdict,”³⁶⁶ and the Fifth Circuit’s “standard of review with respect to a jury verdict is especially deferential.”³⁶⁷ A few circuits have employed ambiguous language, noting that it is difficult to overturn a jury verdict without suggesting that they evaluate Rule 50(b) motions under a different standard.³⁶⁸

Again, what a jury actually decided is not evidence proper, in that it is not submitted to the court for purposes of proof.³⁶⁹ Nor does this hindsight judgment meaningfully implicate hindsight bias anxiety. Our legal system so values juries that deference to their judgment mitigates anxiety that the judge has assumed an improper role.³⁷⁰ But the modes of reasoning discussed earlier—Bayesian updating and inference to the best explanation—apply generally to fact-finding, and they may serve a judge interested in whether a jury verdict is probative of what a reasonable jury could decide.

Under the hindsight-evidence framework, the usefulness of this information depends on how the court understands its inquiry. Part II developed the fact–judgment distinction: Hindsight evidence is useful for finding facts, but it should not play a role in a determination committed

matter of law under the same standard regardless whether the judgment is rendered before or after the jury renders its verdict . . .”).

365. *Hubbard*, 688 F.3d at 716 (citation omitted) (concluding district court erred by explicitly relying on jury’s findings when it granted judgment as matter of law).

366. *Cross v. N.Y.C. Transit Auth.*, 417 F.3d 241, 248 (2d Cir. 2005); see also *Cash v. County of Erie*, 654 F.3d 324, 332–33 (2d Cir. 2011) (quoting *Cross*, 417 F.3d at 248).

367. *Brown v. Bryan County*, 219 F.3d 450, 456 (5th Cir. 2000); see also *Brown v. Sudduth*, 675 F.3d 472, 477 (5th Cir. 2012) (quoting *Brown v. Bryan County*, 219 F.3d at 456). But see *Allen v. Radio One of Tex. II, L.L.C.*, 515 F. App’x 295, 301 (5th Cir. 2013) (“The court applies the same standard provided under Rule 50(a) to evaluate a Rule 50(b) motion . . .”).

368. See *Craig Outdoor Advert., Inc. v. Viacom Outdoor, Inc.*, 528 F.3d 1001, 1009 (8th Cir. 2008) (“A jury verdict is entitled to extreme deference, and we will not set it aside unless no reasonable jury could have reached the same verdict based on the evidence submitted.” (citation omitted)); *Muldrow ex rel. Estate of Muldrow v. Re-Direct, Inc.*, 493 F.3d 160, 165 (D.C. Cir. 2007) (“We do not, however, lightly disturb a jury verdict.”); *Zelinski v. Columbia 300, Inc.*, 335 F.3d 633, 638 (7th Cir. 2003) (“Overturning a jury verdict is a ‘hard row to hoe.’” (quoting *Sheehan v. Donlen Corp.*, 173 F.3d 1039, 1043 (7th Cir. 1999))).

369. See Thayer, *supra* note 89, at 143 (“Evidence . . . is any matter of fact which is furnished to a legal tribunal otherwise than by reasoning, as the basis of inference in ascertaining some other matter of fact. And the law of evidence is the law which has to do with the furnishing of this matter of fact.”).

370. See, e.g., *Glasser v. United States*, 315 U.S. 60, 84 (1942) (referring to jury as “prized shield against oppression”).

to the judgment of the fact-finder.³⁷¹ This fact–judgment distinction maps rather neatly onto two ways of understanding the reasonable-jury inquiry on a motion for judgment as a matter of law.

On the one hand, reasonableness or rationality³⁷² could be a *positive* inquiry, where the judge defines rational decisionmaking as typical human decisionmaking and so takes decisions made by a sufficient number of actual juries to be rational.³⁷³ In an individual case, the judge decides whether the decision was sufficiently typical of jury behavior to be reasonable or whether it was an aberration. A number of scholars have defended this understanding of the reasonable jury standard.³⁷⁴ The positive inquiry is a form of fact-finding: The judge aims to determine what the prior probability was that a jury, given the evidence presented, would come out a certain way. And evidence of the actual jury’s verdict is relevant. The probability of observing the jury’s verdict if it is a typical, highly probable decision is greater than the probability of observing the decision if only occasional outlier juries would arrive at it.³⁷⁵

On the other hand, reasonableness could be a *normative* inquiry, where the judge decides what range of jury decisions are acceptable under some principled understanding of the law and legal process, without reference to what juries actually do.³⁷⁶ The judge determines whether our legal system should countenance making certain inferences from certain evidence. A different group of scholars has defended these normative conceptions of the reasonable jury.³⁷⁷ The normative inquiry is

371. See *supra* section II.B (outlining fact–judgment distinction).

372. Suja A. Thomas, *The Fallacy of Dispositive Procedure*, 50 B.C. L. Rev. 759, 760 (2009) (noting courts use “reasonable juror” and “rational juror” interchangeably).

373. Cf. Miller & Perry, *supra* note 99, at 370 (noting positive paradigm “defines the reasonable person in accordance with empirically observed practice”).

374. See, e.g., Kahan, Hoffman & Braman, *supra* note 276, at 883–87 (grounding a brand of positive inquiry in notions of legitimacy); Thomas, *supra* note 372, at 776 (arguing courts should consider the variety of jurors’ characteristics and experiences); see also Michael W. Pfautz, Note, *What Would a Reasonable Jury Do? Jury Verdicts Following Summary Judgment Reversals*, 115 Colum. L. Rev. 1255, 1258 (2015) (implying scholars can learn about behavior of *reasonable* juries by observing *actual* juries).

375. Another positive understanding of reasonableness would understand the world to contain mostly reasonable juries and a few outlier juries, with any decision made by one of the reasonable juries taken to be reasonable. The actual jury’s verdict is still probative under that account, as the probability of observing the verdict is greater if some proportion of reasonable juries, and not just the unreasonable juries, could arrive at that verdict.

376. See Miller & Perry, *supra* note 99, at 326 (“The so-called Hand formula is a normative definition of reasonableness.”).

377. See, e.g., Kenneth S. Abraham, *The Forms and Functions of Tort Law* 96 (4th ed. 2012) (“In granting a motion for a directed verdict on the negligence issue, a court is not ruling that a jury *would not* find for the defendant, but that the jury *should not* and therefore *may not* so find, because it would be unreasonable to do so.”); cf. Miller & Perry, *supra* note 99, at 326–27 (arguing for normative conception of “reasonable person” standard in tort law).

an exercise in judgment, not in fact-finding: The judge is indifferent to the probability that an actual jury would come out one way or another and seeks only to determine what a jury *should* do. Under this view, the jury's verdict is not relevant to the judge's decision.

There are other reasons why judges might wish to defer to the judgment of an actual jury—it embodies community norms and provides legitimacy to court judgments, to name two.³⁷⁸ However, for judges who stick to the Rule 50 standard, attempting to figure out whether a “reasonable jury” could reach a verdict for the nonmoving party, the actual verdict may act as hindsight evidence. I do not take a position here on whether the positive or normative account is superior;³⁷⁹ I note only that the actual verdict does not bear on reasonableness if the judge has a normative conception of the issue, but it does if the judge's understanding is positivist.

CONCLUSION

Hindsight evidence is valuable when and because it improves accuracy in litigation. It gets us closer to the truth. Truthseeking may be a good in itself, but accuracy also has instrumental value. Ex post accuracy influences ex ante behavior—when courts identify wrongdoing and force wrongdoers to internalize the costs of their actions, the legal system deters bad conduct.³⁸⁰ To the extent hindsight evidence furthers accuracy in adjudication, it increases the expected cost of wrongdoing and decreases the expected cost of benign conduct, deterring wrongful conduct and lessening the chilling of acceptable behavior. Further, in many cases, the introduction of hindsight evidence induces the party with better access to evidence to bring forward that evidence. In this way,

378. But see Jason M. Solomon, *The Political Puzzle of the Civil Jury*, 61 *Emory L.J.* 1331, 1353–87 (2012) (arguing benefits of civil jury to fostering legitimacy and constituting community are overstated).

379. Professor Michael Pardo has suggested that a normative understanding of probative value, such as that inherent in Rule 403 determinations, entails a normative understanding of the “reasonable jury,” since it implies “there is more to the value of evidence than subjective beliefs.” Michael S. Pardo, *The Nature and Purpose of Evidence Theory*, 66 *Vand. L. Rev.* 547, 568–69 (2013); see also *id.* at 594–96 (discussing how normative account is necessary to make “micro” (admissibility questions) and “macro” (sufficiency of evidence questions) levels of evidence theory cohere). This view is appealing: It would be odd to weigh widespread bias negatively in deciding whether to admit evidence, then to weigh it positively in deciding whether the evidence is sufficient. To the extent that “applications of the proof standards are essentially assessments of the probative value of the evidence as a whole,” *id.* at 594–95, it might be contradictory to evaluate the probative value of a collection of evidence differently from the probative value of any one piece of evidence. But if Rule 403 and Rule 50 serve different goals, these seemingly contradictory understandings may be reconcilable. Or the best understanding could have both subjective and objective components. A full analysis of this complicated issue lies beyond the scope of this Essay.

380. Posner, *Economic Approach*, *supra* note 70, at 1483.

hindsight evidence jibes with the “best evidence principle,” which requires litigants, collectively, to produce the most probative evidence reasonably available.³⁸¹

This Essay conceptualizes hindsight evidence as a category of evidence, and it discusses the proper consideration of this evidence under Rule 403. Hindsight evidence may rarely be the deciding factor in a case. But thinking about hindsight evidence in terms of its probative value and not *only* its prejudice fosters a system in which courts are incrementally more likely to achieve the right result. And by improving accuracy—if only incrementally—hindsight evidence furthers a basic goal of adjudication.

381. See Dale A. Nance, *The Best Evidence Principle*, 73 *Iowa L. Rev.* 227, 233 (1988) (“[T]he best evidence principle, as particularly applied to trials, expresses the obligation of litigants to provide evidence that will best facilitate this central task of accurately resolving disputed issues of fact.”).