

RULES OF WEIGHT

Charles L. Barzun*

A central assumption of modern evidence law is that its rules are rules of admissibility only. That is, they tell judges whether or not a given piece of evidence may be viewed by the factfinder, but they do not purport to tell the finder of fact how to evaluate the evidence once admitted. One can imagine, however, a system of rules that helps factfinders weigh evidence by instructing them, for instance, that the law considers a class of evidence (say, hearsay) to be of “low weight.” In fact, such rules—rules of weight—are an old idea with roots in Roman law. But they have long been ignored by evidence scholars or, when considered, judged to be anachronistic and deeply inconsistent with a system of trial by jury.

This Article argues that such hostility to rules of weight is unjustified and that their use should be taken seriously as a possible direction for evidence reform. Given that jury trials are now increasingly rare and that, when a jury is used, its discretion is already constrained in a number of ways, the orthodox view of rules of weight now itself seems outdated. Furthermore, there are reasons to think that such rules could be beneficial for forensic factfinding. The past use of them by courts, their current role in administrative adjudication, and recent research in cognitive psychology all suggest ways in which rules of weight could make factfinding fairer, more efficient, and, most important, more accurate. Such benefits make the Supreme Court’s recent condemnation of the use of rules of weight in the administrative context that much more difficult to justify.

© 2008 Charles L. Barzun. Individuals and nonprofit institutions may reproduce and distribute copies of this Article in any format, at or below cost, for educational purposes, so long as each copy identifies the author, provides a citation to the *Notre Dame Law Review*, and includes this provision and copyright notice.

* Climenko Fellow and Lecturer on Law, Harvard Law School; J.D./M.A., University of Virginia; A.B., Harvard University. I would like to thank the following people for their helpful feedback and comments on this or earlier drafts: Sy Damle, Richard Fallon, Aron Fischer, Charles Fischette, Sydney Foster, Jack Goldsmith, Jim Greiner, Daryl Levinson, Greg Mitchell, Jennifer Mnookin, Dale Nance, Caleb Nelson, Jedediah Purdy, George Rutherglen, Frederick Schauer, Jed Shugerman, Alex Stein, Simon Stern, and G.E. White. I would also like to thank Logan Steiner for her helpful research assistance.

INTRODUCTION	1958
I. RULES OF WEIGHT AND THE JURY SYSTEM.....	1964
A. <i>The Early History of Rules of Weight</i>	1964
B. <i>Bentham, Thayer, and "Free Proof"</i>	1966
C. <i>Moore and His Critics</i>	1968
D. <i>A Response to Wigmore</i>	1972
1. Discretion-Limiting Devices at Common Law	1972
2. Rules of Weight and Nonjury Factfinding.....	1975
II. THE CASE FOR RULES OF WEIGHT	1979
A. <i>Rules in General</i>	1979
B. <i>Rules of Weight</i>	1981
1. Corroboration Rules and Sufficiency Rules	1981
2. Presumptions and Admissibility Rules.....	1985
C. <i>Potential Benefits of Rules of Weight</i>	1986
1. Accuracy of Verdicts	1987
a. Base Rates and Weight	1988
b. Application to Evidence Law	1993
2. Fairness and Risk Allocation.....	1997
3. Efficiency and Agency Adjudication	2001
III. A DEEPER CHALLENGE: <i>ALLENTOWN MACK V. NLRB</i>	2005
A. <i>Epistemic Rationale</i>	2008
B. <i>Moral Rationale</i>	2012
CONCLUSION	2017

INTRODUCTION

Modern evidence law tends to be an all-or-nothing affair. Most rules of evidence are admissibility rules that tell judges simply whether or not a given piece of evidence may be viewed by the factfinder. They do not purport to tell the finder of fact how to evaluate the evidence once admitted. One thus frequently hears trial courts respond to evidentiary objections by saying that a given issue “goes to weight, not admissibility.” The point made is that determining admissibility presents a question of law for the judge while evaluating the evidence is a task for the jury to conduct solely in its discretion.

But one can imagine a factfinding process structured quite differently. Rather than providing merely for rules of exclusion, one could design a system of rules that guides the factfinder’s evaluation of the evidence by specifying the probative value the factfinder ought to attach to a given piece of evidence. Thus, for instance, one might admit all hearsay evidence as an *inferior grade* of evidence to which the factfinder is instructed to accord little weight. Or one could have a rule that treats eyewitness identifications made under certain condi-

tions as requiring corroboration in order to sustain a conviction. In short, one could imagine a factfinding process governed by rules of *weight* instead of, or in addition to, rules of exclusion.

Indeed, when dealing with questions of law, courts do something quite like this. The idea of assigning weights to different classes of evidence, for example, is familiar in the context of statutory interpretation. Although some deny the legitimacy of using the legislative history of a statute to interpret its meaning,¹ those who do draw on it generally consider it to be relevant evidence of legislative intent but a categorically less reliable indicator of it than the statute's text.² One function of rules of weight is to structure a comparable hierarchy of evidentiary sources.

Rules of weight are hardly a new idea. The most famous examples of them come from Roman law, where the testimony of at least two witnesses was required for the conviction of certain crimes.³ And the first treatise on the common law of evidence, published in 1754, ranked categories of evidence according to their probative value.⁴ But there has been virtually no serious discussion of them for almost a century. The last effort to offer a sustained treatment of rules of weight was a little-known 1908 treatise written by a lawyer named Charles Moore.⁵ In *A Treatise on Facts or the Weight and Value of Evidence*, Moore compiled thousands of cases that he argued could be, and often had already been, cited as legal authority on such factual issues as whether experiencing pain or sudden shock enhances or distorts one's memory of events or whether a witness who has previously given contradictory testimony is to be believed.⁶

The century's leading evidence scholar, John Henry Wigmore, immediately condemned Moore's treatise as "moral treason" for its suggestion that the jury's discretion to determine the weight and credibility of the evidence could be in any way constrained by rules of law.⁷

1 See, e.g., Antonin Scalia, *Common-Law Courts in a Civil-Law System: The Role of United States Federal Courts in Interpreting the Constitution and Laws*, in *A MATTER OF INTERPRETATION* 3, 3–48 (Amy Gutmann ed., 1997); John F. Manning, *Textualism as a Nondelegation Doctrine*, 97 *COLUM. L. REV.* 673, 686 (1997).

2 See, e.g., William N. Eskridge, Jr. & Philip P. Frickey, *Statutory Interpretation as Practical Reasoning*, 42 *STAN. L. REV.* 321, 353–56 (1990).

3 John Henry Wigmore, *Required Numbers of Witnesses; A Brief History of the Numerical System in England*, 15 *HARV. L. REV.* 83, 84 (1901).

4 GEOFFREY GILBERT, *THE LAW OF EVIDENCE* 4–5 (photo. reprint 1979) (1754).

5 CHARLES C. MOORE, *A TREATISE ON FACTS OR THE WEIGHT AND VALUE OF EVIDENCE* (1908).

6 See 1 *id.* at v–viii; 2 *id.* § 740, at 809–10; *id.* § 1079, at 1217.

7 John H. Wigmore, Book Review, 3 *ILL. L. REV.* 477, 477–78 (1909) (reviewing MOORE, *supra* note 5).

And Wigmore's view has become conventional wisdom.⁸ Since Moore's time, rules of weight have been for the most part ignored entirely or, when considered, dismissed as anachronistic or at odds with a system of trial by jury.⁹

But such malign neglect of rules of weight seems unjustified. For one thing, the claim that rules of weight illegitimately usurp the jury's role is difficult to reconcile with the fact that courts already employ several relatively uncontroversial evidentiary doctrines, including admissibility rules, presumptions, and sufficiency rules, that constrain and channel the jury's discretion in one way or another.¹⁰

Furthermore, even if rules of weight were deeply inconsistent with trial by jury, their wholesale dismissal would still not be warranted. As leading evidence scholars have long maintained, jury trials should no longer serve as the exclusive paradigm of forensic factfinding because they account for only a small fraction of overall civil and criminal dispositions.¹¹ Even more adjudication takes place in admin-

8 WILLIAM TWINING, *RETHINKING EVIDENCE* 71 (Cambridge Univ. Press 2d ed. 2006) (1990) ("Wigmore's view has generally prevailed. Moore seems to have made little impact and was soon forgotten.").

9 See, e.g., *id.* at 70 (suggesting that Moore made a "category mistake" in thinking that courts could offer authoritative guidance on factual issues and that Moore's "conception of the enterprise was rather confused"); Jennifer Mnookin, *Bifurcation and the Law of Evidence*, 155 U. PA. L. REV. PENNUMBRA 134, 142 (2006), <http://www.pennumbra.com/responses/12-2006/Mnookin.pdf> (noting that rules of weight seem "deeply at odds with the very institution out of which evidence rules partly emerged: the jury as fact-finder"). There are a few exceptions to this generalization. See, e.g., MIRJAN DAMAŠKA, *EVIDENCE LAW ADRIFT* 19 (1997) (defending the Roman system of proof against criticisms that its use of rules of weight was excessively "mechanical"); ALEX STEIN, *FOUNDATIONS OF EVIDENCE LAW* 242 (2005) (endorsing the use of corroboration rules, one common type of rule of weight).

10 DAMAŠKA, *supra* note 9, at 19 (noting that despite the aspirations to a system of "free proof," Anglo-American evidence law regulates and structures the factfinding process in a number of ways). Corroboration rules, however, remain controversial. See, e.g., *Carmell v. Texas*, 529 U.S. 513, 574–75 (2000) (Ginsburg J., dissenting) (criticizing a Texas corroboration requirement on the ground that "[o]ur system of justice rests on the general assumption that the truth is not to be determined merely by the number of witnesses on each side of a controversy. In gauging the truth of conflicting evidence, a jury has no simple formulation of weights and measures on which to rely. The touchstone is always credibility; the ultimate measure of testimonial worth is quality and not quantity" (quoting *Weiler v. United States*, 323 U.S. 606, 608 (1945))). Despite the critical language, the *Weiler* Court upheld the corroboration rule in perjury cases. See *Weiler*, 323 U.S. at 609–11.

11 See DAMAŠKA, *supra* note 9, at 128–29 (noting that "jury trials continue to be employed both as the institutional background for reflection on evidence and as a benchmark for shaping evidentiary rules," but that "this fixation on classical court organization—an increasing anachronism in contemporary conditions—cannot last

istrative agencies, where rules of weight are quite common. Regulations promulgated by the Social Security Administration, for instance, demand that more weight be accorded to the testimony of the treating physician than to that of other experts in adjudicating benefits claims.¹² Agency practice, then, may offer a clue as to how factfinding can be regulated in a system of nonbifurcated decisionmaking.¹³ Given that more and more adjudication is nonbifurcated in this way, it seems worthwhile to reconsider rules of weight and to investigate their relative merits.

This Article is an effort to do just that. Its purpose is both to clarify what rules of weight are and to argue that they should be taken seriously as a possible direction for evidence reform. It does not claim that rules of weight are appropriate in all contexts and for all purposes; nor does it call for the adoption of any particular rule. But it does argue that the conventional view, according to which evidence law consists exclusively of rules of admissibility, is an overly cramped one, because rules of weight have the potential to make factfinding more fair, more efficient and—most important—more *accurate*.

Recent work by cognitive psychologists, for instance, suggests that people, even when well trained in the relevant field, are typically poor at assigning weight to evidence when making diagnostic and predictive judgments.¹⁴ And although many scholars have discussed various ways in which the “heuristics and biases” research program may apply to forensic factfinding, none has proposed rules of weight as a possi-

much longer”); TWINING, *supra* note 8, at 169 (“There is a tendency in the orthodox literature on evidence to treat the contested jury trial as the paradigm case of all trials. This is understandable, though misleading . . .”); Roger C. Park & Michael J. Saks, *Evidence Scholarship Reconsidered: Results of the Interdisciplinary Turn*, 47 B.C. L. REV. 949, 956 (2006) (explaining the relative decline of doctrinal evidence scholarship as a consequence of the decreasing significance of the jury trial and noting that the structural changes identified by Professor Damaška “might make the current rules obsolete”). Several articles present statistics and analysis relating to the decline in jury trials. See generally Marc Galanter, *The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts*, 1 J. EMPIRICAL LEGAL STUD. 459, 460–77 (2004) (reporting on the precipitous decline in the occurrence of trials in American civil litigation); Kent D. Syverud, *ADR and the Decline of the American Civil Jury*, 44 UCLA L. REV. 1935, 1942–43 (1997) (arguing that defects in the civil jury system and the growth of alternative dispute resolution methods are central causes of the collapse of the American jury system).

12 See 20 C.F.R. § 404.1527(d)(2) (2007).

13 The suggestion that we ought to look to administrative adjudication as a source of guidance for evidence reform is hardly novel. See, e.g., Kenneth Culp Davis, *Evidence Reform: The Administrative Process Leads the Way*, 34 MINN. L. REV. 581 (1950).

14 See *infra* Part II.C.1.

ble remedy for the cognitive defects identified.¹⁵ Again, this neglect is

15 For the most part, evidence scholars have considered the biases and heuristics research program to be relevant to evidence law and perhaps a justification for particular reform efforts. See, e.g., Erica Beecher-Monas, *Heuristics, Biases, and the Importance of Gatekeeping*, 2003 MICH. ST. L. REV. 987, 1003 (arguing that the possibility that jurors will suffer from the “dilution effect” in part justifies the judge’s gatekeeper role in admitting expert testimony); Victor J. Gold, *Federal Rule of Evidence 403: Observations on the Nature of Unfairly Prejudicial Evidence*, 58 WASH. L. REV. 497, 511–24 (1983) (suggesting that the heuristics and biases research is relevant to evidence law); Kevin Jon Heller, *The Cognitive Psychology of Circumstantial Evidence*, 105 MICH. L. REV. 241, 244–45 (2006) (arguing that the heuristics research may explain why jurors systematically overvalue the probative value of direct evidence in comparison to circumstantial evidence); Jonathan J. Koehler & Daniel N. Shaviro, *Veridical Verdicts: Increasing Verdict Accuracy Through the Use of Overtly Probabilistic Evidence and Methods*, 75 CORNELL L. REV. 247, 272–77 (1990) (arguing that heuristics and biases research shows that people ignore base rates); Thomas D. Lyon & Jonathan J. Koehler, *The Relevance Ratio: Evaluating the Probative Value of Expert Testimony in Child Sexual Abuses*, 82 CORNELL L. REV. 43, 65 n.81 (1996) (noting that such research suggests that people may misestimate the probative value of evidence that a victim suffered from a cluster of symptoms relevant to child sexual abuse); Joëlle Anne Moreno, *Translating “Visions of Rationality” into Specific Legal Reforms*, 2003 MICH. ST. L. REV. 1175, 1177–81 (discussing the relevance of such research to the ways jurors make decisions); J.J. Prescott & Sonja Starr, *Improving Criminal Jury Decision Making After the Blakely Revolution*, 2006 U. ILL. L. REV. 301, 333–45 (discussing a variety of cognitive errors that may affect factfinding by jurors in sentencing proceedings); Jeffrey J. Rachlinski, *Heuristics and Biases in the Courts: Ignorance or Adaptation?*, 79 OR. L. REV. 61, 93 (2000) (suggesting that the hindsight bias and “representativeness heuristic” affect legal decisionmaking but that judges are more aware of such errors in juries than in themselves); D. Michael Risinger & Michael J. Saks, *Rationality, Research and Leviathan: Law Enforcement-Sponsored Research and the Criminal Process*, 2003 MICH. ST. L. REV. 1023, 1036–50 (discussing heuristics research and its bearing on conceptions of rationality relevant to evidence law); Daniel Shaviro, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV. L. REV. 530, 542–43 (1989) (arguing that aversions to the use of naked statistical evidence may be best explained as a “moral heuristic” that reduces complexity and cognitive dissonance); Andrew J. Wistrich et al., *Can Judges Ignore Inadmissible Information? The Difficulty of Deliberately Disregarding*, 153 U. PA. L. REV. 1251, 1291 (2005) (conducting a study that confirmed that judges may suffer from “anchoring effects” in assessing the appropriate amount of damages to award). However, some scholars have remained skeptical about either the intrinsic validity of the research or its usefulness for understanding or reforming evidence law. See, e.g., L. Jonathan Cohen, *Can Human Irrationality Be Experimentally Demonstrated?*, 4 BEHAV. & BRAIN SCI. 317, 317–18, 326–30 (1981) (arguing that cognitive psychologists, particularly Kahneman and Tversky, may be guilty of misinterpreting their own data, and expressing doubt at the very idea of demonstrating that humans are systematically incompetent reasoners); Gregory Mitchell, *Mapping Evidence Law*, 2003 MICH. ST. L. REV. 1065, 1082 (arguing that the heuristics and biases research is not as broadly applicable as some have thought and suggesting instead that an approach more sensitive to context is warranted); Chris William Sanchirico, *Character Evidence and the Object of Trial*, 101 COLUM. L. REV. 1227, 1242 (2001) (rejecting the argument that the representative-

somewhat surprising, because using rules that attempt to grade the probative value of evidence seems like an obvious solution.¹⁶

The structure of this Article is straightforward. Part I presents and responds to the primary objection to rules of weight, namely that they are deeply incompatible with a system of trial by jury. I show that, rhetoric aside, the jury's discretion is in fact cabined in a number of ways and that, in any case, jury trials constitute only a small portion of adjudications.¹⁷ Part II then lays out the case for rules of weight. It begins by clarifying what exactly rules of weight are and how they differ from other evidentiary devices. It then argues that using such rules has the potential to increase the efficiency, fairness, and accuracy of adjudicative factfinding. Part III considers a far deeper challenge to rules of weight. It analyzes the Supreme Court's discussion of rules of weight in the case of *Allentown Mack Sales & Service, Inc. v. NLRB*,¹⁸ in which the Court appears to have concluded that such rules violate fundamental principles of due process.¹⁹ I argue that even under sympathetic interpretations, the Court's reasoning fails to justify a blanket exclusion of all rules of weight in all adjudicatory contexts. Finally, in conclusion, I suggest that my analysis of rules of weight shows how drawing comparisons between methods of finding facts and interpreting law may illuminate both practices.

ness heuristic justifies the exclusion of character evidence); Peter Tillers, *What Is Wrong with Character Evidence?*, 49 HASTINGS L.J. 781, 791 (1998) (criticizing the suggestion that the heuristics line of research implies that "all people—or practically all people, in any event—are incapable of judging the true value of character evidence"); Charles Yablon, *The Meaning of Probability Judgments: An Essay on the Use and Misuse of Behavioral Economics*, 2004 U. ILL. L. REV. 899, 927–29 (criticizing the suggestion that the base rate problems studied by Kahneman and Tversky necessarily imply human irrationality). Finally, at least one scholar has suggested a way in which the cognitive errors identified by psychologists may be valuable for evidentiary purposes. See Chris William Sanchirico, *Evidence, Procedure, and the Upside of Cognitive Error*, 57 STAN. L. REV. 291, 298 (2004) (suggesting that certain cognitive errors may be instrumentally valuable insofar as they make it more difficult for witnesses to lie).

16 See Mnookin, *supra* note 9, at 142 (noting, in response to Professor Schauer's argument that the cognitive defects from which judges likely suffer may justify the application of admissibility rules even in bench trials, that "rules of assessment" are likely better suited to the task).

17 In so doing, I draw in part from Professor Damaška's analysis. See DAMAŠKA, *supra* note 9, at 17–20.

18 522 U.S. 359 (1998).

19 See *id.* at 380.

I. RULES OF WEIGHT AND THE JURY SYSTEM

A. *The Early History of Rules of Weight*

Rules of weight have their roots in medieval Roman law. Under Roman law, and later under Continental civil law, the testimony of one witness was typically insufficient to prove most crimes.²⁰ The number of witnesses required to prove various acts varied, but for many crimes, such as murder, the testimony of at least two witnesses was required for the “full proof” necessary to sustain a conviction.²¹ One witness’ testimony, however, was sufficient to establish “half proof,” which authorized the court to have the defendant tortured until he confessed to the crime.²² The seemingly naïve epistemological assumption on which this system appeared to rest, namely that any testimony given under oath was as probative as any other, not to mention its close institutional affiliation with a system of torture, makes the Roman system of proof easy to caricature as draconian. But the purpose of such rules was, at least in theory, the humane one of offering criminal defendants some protection against the otherwise unchecked power of judges.²³ The judge, for instance, was rarely required to make *positive* findings against a defendant because of the number of witnesses.²⁴

The common law for the most part has not been receptive to such “corroboration rules,” which appear to be in some tension with a system in which the jury assesses the credibility of the evidence. But there are some notable exceptions. The United States Constitution, for instance, requires the testimony of two witnesses for a treason conviction.²⁵ And federal courts and many states require two witnesses to sustain a perjury conviction.²⁶ Corroborative evidence has also been

20 Wigmore, *supra* note 3, at 84.

21 See Barbara J. Shapiro, “Fact” and the Proof of Fact in Anglo-American Law (c. 1500–1850), in *HOW LAW KNOWS* 28, 31 (Austin Sarat et al. eds., 2007) (discussing the development of medieval practices from many influences, including Roman law).

22 *Id.*

23 WILLIAM WILLS, AN ESSAY ON THE RATIONALE OF CIRCUMSTANTIAL EVIDENCE 34–35 (London, Longman et al. 1838); see also John H. Langbein, *Torture and Plea Bargaining*, 46 U. CHI. L. REV. 3, 4 (1978) (“In the history of Western culture no legal system has ever made a more valiant effort to perfect its safeguards and thereby to exclude completely the possibility of a mistaken conviction.”).

24 DAMAŠKA, *supra* note 9, at 19–20.

25 U.S. CONST. art. III, § 3 (“No Person shall be convicted of Treason unless on the Testimony of two Witnesses to the same overt Act, or on Confession in open Court.”).

26 See *Weiler v. United States*, 323 U.S. 606, 609–11 (1945) (upholding the two-witness rule for perjury); *Hammer v. United States*, 271 U.S. 620, 626 (1926) (“The

required to sustain convictions based on the testimony of the victims of sexual abuse or the testimony of an accomplice to a crime.²⁷ Furthermore, the first treatise writer on the common law of evidence endorsed the use of such rules.

Lord Chief Baron Gilbert's *The Law of Evidence*²⁸ is considered the first treatise on evidence law.²⁹ Gilbert is now best known both for establishing evidence as an independent branch of law and for basing it on the "best evidence principle," according to which the central function of evidence law is to secure the best available evidence.³⁰ Gilbert opened his treatise by declaring that "[t]he first Thing to be treated of, is the Evidence that ought to be offer[e]d to the Jury, and by what Rules of Probability it ought to be weigh[e]d and consider[e]d."³¹ The burden of evidence law was thus to ensure that the evidence adduced by parties in a legal dispute was accorded precisely the amount of weight that its degree of probability required, thereby

general rule in prosecutions for perjury is that the uncorroborated oath of one witness is not enough to establish the falsity of the testimony of the accused set forth in the indictment as perjury. The application of that rule in federal and state courts is well nigh universal. The rule has long prevailed, and no enactment in derogation of it has come to our attention." (footnote omitted)). Some states have abandoned the two-witness rule for perjury convictions. *See, e.g.*, Commonwealth v. Silva, 516 N.E.2d 161, 165 (Mass. 1987) ("Hereafter, the Commonwealth may secure a conviction of perjury where it is able to offer evidence of perjury and corroboration of that evidence sufficient to establish the defendant's guilt beyond a reasonable doubt."). But other states continue to require corroboration of a single witness' testimony. *See, e.g.*, COLO. REV. STAT. ANN. § 18-8-506 (West 2004) ("In any prosecution for perjury or false swearing, except a prosecution based upon inconsistent statements . . . , falsity of a statement may not be established solely through contradiction by the uncorroborated testimony of a single witness."); N.Y. PENAL LAW § 210.50 (McKinney 1998) ("In any prosecution for perjury, except a prosecution based upon inconsistent statements pursuant to section 210.20, or in any prosecution for making an apparently sworn false statement, or making a punishable false written statement, falsity of a statement may not be established by the uncorroborated testimony of a single witness.").

27 *See* LAW COMM'N, LAW COM. NO. 202, CRIMINAL LAW: CORROBORATION OF EVIDENCE IN CRIMINAL TRIALS 3-7 (1991) [hereinafter CORROBORATION OF EVIDENCE] (recommending the abolition of England's accomplice testimony corroboration rule); Irving Younger, *The Requirement of Corroboration in Prosecutions for Sex Offenses in New York*, 40 FORDHAM L. REV. 263, 264-66 (1971) (surveying states with corroboration rules for sexual offenses).

28 GILBERT, *supra* note 4.

29 TWINING, *supra* note 8, at 38-39; Barbara J. Shapiro, "To a Moral Certainty": *Theories of Knowledge and Anglo-American Juries 1600-1850*, 38 HASTINGS L.J. 153, 176 (1986).

30 GILBERT, *supra* note 4, at 1-2. For the leading modern-day proponent of the best evidence principle, see Dale A. Nance, *The Best Evidence Principle*, 73 IOWA L. REV. 227 (1988).

31 GILBERT, *supra* note 4, at 1.

enabling the factfinder to make the “most exact Di[sc]ernment” of the rights at issue.³²

To this end, Gilbert established an elaborate hierarchy in which he classified types of evidence according to their relative weight. Written evidence was considered superior to unwritten evidence,³³ and within the category of written evidence, sealed documents were superior to unsealed documents and publicly sealed documents superior to privately sealed ones.³⁴ Similarly, when witnesses testified, two witnesses testifying under oath constituted evidence of a higher “degree of credibility” than merely one witness and “[o]ne affirmative Witness countervails the Proof of several Negative.”³⁵ This latter rule was justified on the ground that those testifying in the negative may “know not of the Matter,” in which case the two testimonies do not contradict each other.³⁶ But Gilbert’s attempt to subject factfinding to such an elaborate set of rules provoked a bitter attack by Jeremy Bentham.

B. Bentham, Thayer, and “Free Proof”

In his massive, five-volume *Rationale of Judicial Evidence* and in other shorter works, Jeremy Bentham expounded a view of evidence that has been called “the most ambitious and fully developed theory of evidence and proof in the history of legal thought.”³⁷ Bentham’s primary target was Gilbert, and his goal was to show why all rules of evidence should be abolished.

Bentham denied Gilbert’s premise that evidence law was designed to ensure the production of the best evidence available.³⁸ Instead, for Bentham, the direct end of evidence law—and indeed all procedural law—was rather to secure “rectitude of decision,” so that legal obligations would only be imposed upon those “on whom the legislator intended that they should be conferred and imposed.”³⁹ This principle was to be limited only by the second, “collateral” goal of avoiding “unnecessary delay, vexation, and expense.”⁴⁰ In Bentham’s view, virtually all rules of evidence only frustrated attempts to achieve these goals. “To find infallible rules for evidence, rules which insure a

32 *Id.* at 1–2.

33 *See id.* at 4–5.

34 *See id.* at 15–16.

35 *Id.* at 108–10.

36 *Id.* at 110.

37 TWINING, *supra* note 8, at 42.

38 *See* WILLIAM TWINING, THEORIES OF EVIDENCE 32 (1985).

39 1 JEREMY BENTHAM, RATIONALE OF JUDICIAL EVIDENCE 34 (Fred B. Rothman & Co. 1995) (1827).

40 *Id.*

just decision," Bentham insisted, "is, from the nature of things, absolutely impossible; but the human mind is too apt to establish rules which only increase the probabilities of a bad decision."⁴¹ He advocated instead a "natural system" of evidence in which judges found facts entirely unencumbered by rules. Doing so would facilitate the only "mode of searching out the truth," namely "see everything that is to be seen; hear every body who is likely to know any thing about the matter."⁴² Thus, Bentham's system of "free proof" had little room for rules of weight *or* rules of admissibility.⁴³

Bentham did not deny that factfinders may need help in evaluating evidence. Indeed, in place of exclusionary rules, he advocated a "system of instruction," which would serve as a "gentle and rational substitute" for a system of exclusion.⁴⁴ Under this system, "the judge has before him the instruction, which, in its nature, cannot be so much as intended to serve as a guide to the understanding of a judge, without also serving as a check upon his will."⁴⁵ Bentham discussed at length the form and substance such instructions might take. He catalogued, for instance, the various types of interests—wealth, power, reputation, sympathy, disgrace—that tend to affect the reliability of witness testimony.⁴⁶

Still, Bentham established "free proof" as the new baseline for evaluating evidence law. The idea that any rule of evidence requires affirmative justification has endured in Anglo-American evidence scholarship, and the doctrinal trend has been toward fewer and narrower exclusionary rules.⁴⁷ Thus, Bentham's "anti-nomian thesis" has been partially vindicated: the modern approach looks askance at rules of weight, but continues to apply admissibility rules.⁴⁸

This view finds the clearest expression in the work and teachings of James Bradley Thayer, who was, along with his student Wigmore, one of the founders of modern evidence law.⁴⁹ Thayer exerted tremendous influence on later scholars, partly through teaching and

41 JEREMY BENTHAM, *A TREATISE ON JUDICIAL EVIDENCE* 180 (Fred B. Rotham & Co. 1981) (1825).

42 5 BENTHAM, *supra* note 39, at 743.

43 *Id.* Bentham himself never used the phrase "free proof," but the term has come to be associated with his antinomian approach. See TWINING, *supra* note 8, at 209.

44 5 BENTHAM, *supra* note 39, at 615.

45 *Id.* at 616.

46 *See id.* at 629.

47 *See* STEIN, *supra* note 9, at 108; TWINING, *supra* note 38, at 69.

48 TWINING, *supra* note 8, at 44 ("In respect of rules of weight, Bentham's views have largely prevailed").

49 *See id.* at 61–62.

partly through his mostly historical *Preliminary Treatise on Evidence at Common Law*.⁵⁰

For Thayer, the purpose of evidence law could be stated simply. The “main errand of the law of evidence,” he explained, was “to determine not so much what is admissible in proof, as what is inadmissible.”⁵¹ Whether evidence ought to be excluded was determined by two factors. First, by relevance, which was “an affair of logic and experience,” and second, “by the law of evidence, which declares whether any given matter which is logically probative is excluded.”⁵² Evidence law in turn required “(1) that nothing is to be received which is not logically probative of some matter requiring to be proved; and (2) that everything which is thus probative should come in, unless a clear ground of policy or law excludes it.”⁵³ In other words, in contrast to “old formal and mechanical systems,” evidence law consisted merely in a narrow set of admissibility rules, which served as grounds for excluding otherwise relevant evidence.⁵⁴

Thayer’s view has become orthodoxy. Under the modern view, evidence law includes only exclusionary rules designed to determine which evidence the factfinder may consider, but it does not purport to guide the factfinding process. According to Thayer, it “does not undertake to regulate the processes of reasoning or argument, except as helping to discriminate and select the material of fact upon which these are to operate.”⁵⁵ To do otherwise would be to ask the impossible: “[t]he law has no mandamus to the logical faculty; it orders no body to draw inferences.”⁵⁶ “For reasoning,” Thayer insisted, “there is no law other than the laws of thought.”⁵⁷

C. Moore and His Critics

The orthodox view, however, is not the only way to treat questions of fact. And in Thayer’s own time, one treatise writer offered a different approach. In 1908, a writer and lawyer named Charles Moore compiled thousands of cases in his work, *A Treatise on Facts or the Weight and Value of Evidence*.⁵⁸ According to Moore, the cases col-

50 JAMES BRADLEY THAYER, *A PRELIMINARY TREATISE ON EVIDENCE AT THE COMMON LAW* (Boston, Little, Brown & Co. 1898).

51 *Id.* at 268.

52 *Id.* at 269.

53 *Id.* at 530.

54 *See id.* at 265.

55 *Id.* at 263.

56 *Id.* at 313 n.1.

57 *Id.* at 314.

58 MOORE, *supra* note 5.

lected contained rules for weighing evidence and assessing the credibility of witnesses. Moore made plain this ambition in his very first sentences:

Arguments on any question of fact can be supported by reference to judicial authorities—on both sides—as fully as arguments on questions of law are thus fortified. The design of this work is to facilitate the preparation for trial, the argument, and the decision of questions of fact, by exhibiting what has been said by United States, Canadian, and English judges concerning the causes of trustworthiness and untrustworthiness of evidence, and the rules for determining its probative weight.⁵⁹

In these cases, Moore explained, the reader could find the “deliberate convictions of unprejudiced minds, enlightened in numerous instances by vast experience on the bench.”⁶⁰ Such experience had revealed both general principles and specific rules according to which courts could determine the degree of credibility various forms of testimony warranted.

Moore on Facts, as the work came to be known, is a treatise unlike any other written before or since.⁶¹ Because its purpose is to assist attorneys and courts in evaluating facts, its chapters list the various factual circumstances courts have been called upon to consider. The topics covered range from “Uncontradicted Testimony,”⁶² “Memory,”⁶³ and “Bias of Witnesses”⁶⁴ to “Sound and Hearing,”⁶⁵ “Taste, Smell, and Touch,”⁶⁶ and “The Weather.”⁶⁷ Some of the section headings in *Moore on Facts* are so specific, so obviously outdated, or so just plain odd that they are more likely to strike the modern reader as a source of humor than as a repository of legal wisdom.⁶⁸ Furthermore,

59 1 *id.* at v.

60 *Id.*

61 Its closest cousins are WILLS, *supra* note 23, and JAMES RAM, A TREATISE ON FACTS AS SUBJECTS OF INQUIRY BY A JURY (Fred B. Rothman & Co. 2d ed. 1982) (1873), but neither of these works is nearly as systematic in method or as comprehensive in scope.

62 1 MOORE, *supra* note 5, at 109.

63 2 *id.* at 786.

64 *Id.* at 1224.

65 1 *id.* at 231.

66 *Id.* at 358.

67 *Id.* at 453.

68 See, e.g., *id.* § 151, at 197 (“Course of Bullets Striking Human Body—Fact versus Theory”); *id.* § 161, at 206 (“Travels of a Toddling Infant”); *id.* § 164, at 208 (“Incredibly Quick Comprehension of Long Document”); *id.* § 202, at 245 (“Church Bells—Waywardness of Sound—Sensitiveness of Hearers”); *id.* § 206, at 248 (“Sounds Attributed to Sexual Intercourse”); *id.* § 215, at 254 (“Creptitation of Injured Bones”);

some others are downright offensive: Moore classified witnesses, for instance, not only by race ("Hindus,"⁶⁹ "Negroes,"⁷⁰ and "Orientals, in General,"⁷¹), but also on the basis of such derogatory category labels as "Prostitutes,"⁷² "Opium Fiends,"⁷³ and, worst of all, "Lawyers."⁷⁴

Still, Moore showed keen awareness of some of the psychological defects that infect witness testimony. Much attention has been paid recently to the mountains of psychological research produced over the last few decades showing that eyewitness testimony is far less reliable than most people (and thus most jurors) tend to think.⁷⁵ But Moore showed that courts had long recognized this fact. He even offered explanations for the phenomenon, such as the ease with which a person's attention may be diverted,⁷⁶ the tendency to supplant observation with unconscious inference,⁷⁷ and the capacity of a witness' religious, political, or pecuniary interest to affect unconsciously what he or she observes.⁷⁸ Moore was equally aware of the problems associated with memory. He at one point asserts that he has "discovered no case where a judge has had the hardihood to extol the tenacity of memory" and goes on to discuss the various factors affecting memory.⁷⁹ Because of these and other problems, Moore was concerned that juries might often accord too much weight to testimony "intrinsically weak" in virtue of the witness' "well-known integrity and prudent judgment."⁸⁰ Indeed, the very first section of the treatise is entitled "Inclination to Give Too Much Weight to Testimony."⁸¹

Moreover, Moore found support in the case law for what appear to be rules of weight or credibility. Cases are cited for the proposition that a person's memory of his or her own actions is more reliable than

id. § 309, at 312 ("Person Dazzled by Lights Shining in His Face"); *id.* § 379, at 359 ("Educated Taste of Kentuckians"); *id.* § 439, at 410 ("Estimate of Speed by Intoxicated Person"); *id.* § 503, at 472 ("Can Fog Coexist with a Gale?"); 2 *id.* § 1001, at 1145 ("Sneering Levity"); *id.* § 1031, at 1164 ("Are Unchaste Women Less Veracious than Unchaste Men?").

69 2 *id.* § 1024, at 1160.

70 *Id.* § 1025, at 1160.

71 *Id.* § 1022, at 1159.

72 *Id.* § 1032, at 1165.

73 *Id.* § 1020, at 1157.

74 *Id.* § 1019, at 1156.

75 See *infra* Part II.C.1.

76 2 MOORE, *supra* note 5, § 693, at 747.

77 *Id.* § 701, at 757.

78 *Id.* § 702, at 758.

79 *Id.* § 728, at 795.

80 1 *id.* § 17, at 22-23.

81 *Id.* § 1, at 2.

another's recollection of them;⁸² that estimates of short periods of time in which a great deal has occurred are typically inexact;⁸³ that testimony about a conversation with a deceased person is considered weak;⁸⁴ that the testimony of an uncontradicted, disinterested, and unimpeached witness deserves at least some weight;⁸⁵ that the testimony of a witness who earlier gave contradictory testimony must be given zero weight;⁸⁶ and that affidavits are liable to abuse because they are rarely drafted by the witness.⁸⁷

Moore's treatise received some favorable reviews,⁸⁸ but its most important reviewer condemned it as "moral treason."⁸⁹ John Henry Wigmore, who wrote the leading twentieth-century treatise on evidence law, described Moore's work as "partly good," but also partly "what we should consider as bad as possible."⁹⁰ Though admiring of the book's broad scope and potential utility for lawyers, Wigmore saw "nothing but harm" in Moore's suggestion that there were "*rules of law which determine the weight or credibility of a piece of evidence.*"⁹¹ Such a proposition, he insisted, "is not known to the orthodox and traditional common law," and its adoption would "wreck our whole system of proof."⁹² The reason was simple:

If there is one thing for which the common law system of judge and jury stands, it is that the rules of evidence, as determined and applied by the judge, are rules of admissibility alone, and for the judge alone; the weight or credibility is for the jurors untrammelled by any rules of law.⁹³

Today Moore has been almost entirely forgotten, and the only scholar to discuss his work in any detail has endorsed Wigmore's nega-

82 2 *id.* § 783, at 870.

83 *Id.* § 863, at 992.

84 *Id.* § 877, at 1014.

85 1 *id.* § 98, at 148–49.

86 2 *id.* §§ 1073–1074, at 1209.

87 *Id.* § 941, at 1097.

88 See, e.g., Charles F. Chamberlayne, Book Review, 20 GREEN BAG 627, 627 (1908) (reviewing MOORE, *supra* note 5) ("This two-volume work is destined to become a classic . . ."); Kenneth M. Spence, Book Review, 9 COLUM. L. REV. 370, 371 (1909) (reviewing MOORE, *supra* note 5) ("[I]t is not too much to say that [Moore] has added to the existing and available law of evidence almost an entire branch, the usefulness of which can hardly be overestimated.").

89 Wigmore, *supra* note 7, at 478.

90 *Id.* at 477.

91 *Id.*

92 *Id.*

93 *Id.*

tive assessment of it.⁹⁴ William Twining praises Moore's treatise as "a rich and fascinating compendium," which is "full of delights and insights," but he says it was based on a "category mistake."⁹⁵ In suggesting that case law could offer courts "authoritative guidance on the weight and value of particular kinds of evidence," Moore demonstrated that his "conception of the enterprise was rather confused."⁹⁶ History thus appears to have judged Wigmore's critique dispositive. No one since Moore has put forth a serious effort to defend rules of weight or explain how they might be used.

D. *A Response to Wigmore*

Such neglect is unfortunate, for Wigmore's wholesale rejection of rules of weight on the ground that they are inconsistent with trial by jury was not justified. First, the argument proves far too much because the jury's role as finder of fact has been, and continues to be, constrained in a number of respects. Second, there are a variety of adjudicative contexts in which the finder of fact is not a jury.

1. Discretion-Limiting Devices at Common Law

Although modern evidence law purports to be consistent with Thayer's proclamation that "law has no mandamus to the logical faculty,"⁹⁷ several types of rules control and manage the information to which the factfinder has access.⁹⁸ First and most obviously, admissibility rules routinely exclude probative evidence. And although many such rules exclude evidence for reasons of policy irrespective of its probative value, some evidence is excluded precisely on the ground that the factfinder is likely to ascribe it more weight than it deserves. The hearsay rule is a classic example of such a rule, the theory being that the inability to cross-examine the declarant under oath reduces the factfinder's ability to assess accurately the reliability of his or her alleged statements.⁹⁹ Furthermore, courts often admit evidence for

94 See TWINING, *supra* note 8, at 71.

95 *Id.* at 70-71.

96 *Id.* at 70.

97 THAYER, *supra* note 50, at 313 n.1.

98 Professor Damaška has made this point well: "Common law procedure is usually regarded as the citadel of free evaluation of evidence. . . . But this account of the matter is incomplete and seriously misleading. . . . Common law was never averse to instruments aimed at constraining the fact finder's freedom in processing evidence." DAMAŠKA, *supra* note 9, at 17-18.

99 See FED. R. EVID. 801-806; DAMAŠKA, *supra* note 9, at 15 (identifying reliability concerns as one of the traditional justifications for the hearsay rule); RICHARD O. LEMPET ET AL., *A MODERN APPROACH TO EVIDENCE* 501 (3d ed. 2000) (same).

one purpose but not another, with the result that the factfinder is instructed to consider, say, a criminal defendant's reputation for dishonesty for the purpose of impeachment but not as substantive evidence that he robbed a bank.¹⁰⁰ In this way, as Professor Mirjan Damaška puts it, Anglo-American evidence law seeks to "regulat[e] legally permissible lines of inference."¹⁰¹

Just as the most common admissibility rules limit discretionary factfinding, so too do substantive rules of law. Whenever a court decides a question "as a matter of law," the effect is to withdraw from the jury's consideration what would otherwise be disputable factual questions and decide them according to a rule of law. In fact, Oliver Wendell Holmes argued that such was a chief virtue of the common law method of decisionmaking. On certain kinds of recurring questions, he maintained, courts should "gradually . . . acquire a fund of experience which enables [the judge] to represent the common sense of the community in ordinary instances far better than an average jury."¹⁰² Holmes famously put this principle in practice in *Baltimore & Ohio Railroad Co. v. Goodman*,¹⁰³ a case in which the Supreme Court endorsed the notorious "stop, look, and listen" rule for railroad crossing cases.¹⁰⁴ Under this rule, rather than ask a jury whether the defendant had adduced evidence sufficient to establish that the plaintiff was contributorily negligent when crossing a railroad track, so long as the evidence made clear that the plaintiff did not stop and get out of his car to determine whether the train was "dangerously near," the court could settle the issue as a matter of law.¹⁰⁵ This particular rule was soon limited by the Court,¹⁰⁶ but the practice is pervasive: whenever courts decide cases on the merits at the pleading stage or on a motion for summary judgment, they generate rules whose effect is to convert potential questions of fact for juries into questions of law for courts.

Also common—but more controversial in some contexts—are legal presumptions. These rules encourage or require factfinders to

100 Compare FED. R. EVID. 404(a) (excluding character evidence generally), with *id.* 608(a) (permitting character evidence for purposes of impeachment).

101 DAMAŠKA, *supra* note 9, at 18.

102 OLIVER WENDELL HOLMES, JR., *THE COMMON LAW* 124 (Little, Brown & Co. 1923) (1881).

103 275 U.S. 66 (1927).

104 *See id.* at 69–70.

105 *Id.* at 70.

106 *See Pokora v. Wabash Ry. Co.*, 292 U.S. 98, 102–06 (1934) (discussing the conflicts caused by the *Goodman* rule and deciding that it cannot be applied generally as a rule of law).

infer a fact from circumstantial evidence of that fact. Typical presumptions in the civil context include the presumption that a letter mailed to an address was received by the addressee, that the driver of a vehicle that strikes another vehicle from behind is negligent, that a person who remains absent from friends and family for seven years without explanation or communication is dead, or that a will that cannot be found was revoked by the testator.¹⁰⁷ Legal presumptions have been more controversial in the criminal context. In a line of cases in the 1970s and 1980s, the Supreme Court struck down several legislative presumptions that made it easier for juries to convict on circumstantial evidence.¹⁰⁸ Precisely how much evidence is required to defeat a presumption is sometimes unclear and has been the source of much debate,¹⁰⁹ but there is little doubt

107 See GRAHAM C. LILLY, AN INTRODUCTION TO THE LAW OF EVIDENCE § 3.3, at 63–64 (3d ed. 1996).

108 See, e.g., *Francis v. Franklin*, 471 U.S. 307, 316 (1985) (condemning an instruction that “[t]he law presumes that a person intends the ordinary consequences of his voluntary acts’” on the ground that the jury could have construed it as mandatory (quoting *Sandstrom v. Montana*, 442 U.S. 510, 515 (1979))); *Sandstrom*, 442 U.S. at 515 (striking down a mandatory presumption similar to that at issue in *Francis* as applied to a homicide defendant); *County Court v. Allen*, 442 U.S. 140, 163–67 (1979) (upholding a permissive presumption that the defendant illegally possessed a firearm from the fact that it was in her car); *Mullaney v. Wilbur*, 421 U.S. 684, 701–04 (1975) (striking down a state law presumption that shifted the burden to the accused to prove provocation in a murder trial); *Turner v. United States*, 396 U.S. 398, 405–19 (1970) (striking down a mandatory presumption that the defendant knowingly imported cocaine from abroad from the fact of his possession of heroin, but upholding a comparable presumption from heroin possession). For criticism of the Court’s approach and an analysis of presumptions as burden-shifting devices, see Ronald J. Allen, *Structuring Jury Decisionmaking in Criminal Cases: A Unified Constitutional Approach to Evidentiary Devices*, 94 HARV. L. REV. 321, 321–39 (1980); John Calvin Jeffries, Jr. & Paul B. Stephan III, *Defenses, Presumptions, and Burden of Proof in the Criminal Law*, 88 YALE L.J. 1325, 1365–73 (1979) (suggesting a framework to evaluate the constitutionality of defenses and presumptions in the criminal law as they relate to the burden of proof beyond a reasonable doubt).

109 The issue at the center of the so-called “Thayer-Morgan debate” is essentially whether the function of a presumption is to shift the burden of persuasion or just the burden of production. See, e.g., EDMUND MORRIS MORGAN, *SOME PROBLEMS OF PROOF UNDER THE ANGLO-AMERICAN SYSTEM OF LITIGATION* 81 (1956) (“[S]ensible procedure would require the abandonment of the Thayer doctrine as to presumptions and concede that the establishment of facts which create a presumption does in some situations and should in most cases fix the burden of persuasion.”); Leslie J. Harris, *Constitutional Limits on Criminal Presumptions as an Expression of Changing Concepts of Fundamental Fairness*, 77 J. CRIM. L. & CRIMINOLOGY 308, 314 n.22 (1986) (summarizing the debate and citing relevant sources).

that their use, at least in the civil context, is widespread and uncontroversial.¹¹⁰

Corroboration rules are similar to presumptions, except that they apply to direct testimony rather than circumstantial evidence, and they typically result in either the admission of evidence or the dismissal of the case.¹¹¹ These rules generally require that certain types of testimony be corroborated in order to sustain a conviction. As mentioned above, such rules are no longer particularly common, but they continue to be used in perjury and treason prosecutions as well as in those for certain sexual offenses.¹¹² The Supreme Court recently considered whether the Constitution's Ex Post Facto Clause applied to such a rule. In *Carmell v. Texas*,¹¹³ the Court split 5–4 over whether the expansion of an exception to a Texas corroboration requirement for the testimony of victims of sexual abuse could be constitutionally applied to the defendant.¹¹⁴ Interestingly, the majority and dissent could not even agree as to how to characterize the rule. The majority dubbed it a “sufficiency of the evidence rule,”¹¹⁵ while the dissent saw it as a traditional admissibility rule.¹¹⁶ In the next Part, we will see why a better term would be a rule of weight.¹¹⁷

In short, if rules of weight are inconsistent with trial by jury, many of the evidentiary devices that have become well accepted are equally so.

2. Rules of Weight and Nonjury Factfinding

Moore himself responded to Wigmore in part by noting that there are also several adjudicative contexts in which juries are not the finder of fact.¹¹⁸ In bench trials, on motions for a new trial or motions challenging the sufficiency of the evidence, or in administrative adjudication (where rules of weight are common), judges are frequently called upon to evaluate evidence—even if they are not always “finding facts” as a formal matter. Moore's point carries even more

110 See DAMASKA, *supra* note 9, at 19 (listing presumptions as one of the “common law's arsenal of instruments designed to structure the deliberative process”).

111 See *id.* at 18 (mentioning corroboration rules as one of the common law limitations on free proof); TWINING, *supra* note 8, at 44 (citing corroboration rules as one of the few contexts in which rules of weight are employed in the common law system).

112 See *supra* Part I.A.

113 529 U.S. 513 (2000).

114 See *id.* at 518–19, 552.

115 See *id.* at 546–47.

116 See *id.* at 556 (Ginsburg, J., dissenting).

117 See *infra* Part II.B.

118 See Charles C. Moore, *Correspondence*, 3 ILL. L. REV. 583, 584 (1909).

weight today because jury trials have become exceedingly rare. In 2002, for instance, only 1.2% of all civil dispositions in federal district courts resulted from jury trials.¹¹⁹ Even by 1962, however, that percentage had dwindled to 5.5%.¹²⁰

Perhaps not surprisingly, then, since Moore's writing, various appellate and trial courts sitting as finders of fact have drawn on several of the rules of weight Moore collected in his treatise. Examples include the rule that inherently incredible testimony is generally not to be believed;¹²¹ that a witness' memory of his own actions is superior to that of others;¹²² that testimony about a decedent's alleged statements against interest is generally unreliable;¹²³ that the testimony of seamen exonerating their master has little weight;¹²⁴ that the testimony of an uncontradicted, disinterested witness must be accorded

119 Galanter, *supra* note 11, at 463 tbl.1.

120 *Id.* at 462 tbl.1.

121 *See, e.g.*, Jackson v. United States, 353 F.2d 862, 867 (D.C. Cir. 1965) ("The doctrine that appellate courts must reverse findings based upon 'inherently incredible' testimony has long been accepted in this jurisdiction."); Trippett v. Virgin Islands, No. D.C. CRIM. APP. 2004/03, 2004 WL 2988560, at *3 (D.V.I. 2004) ("Where, however, the testimony of a witness is incredible, inherently or physically impossible and unbelievable, inherently improbable and irreconcilable with, or contrary to physical facts and common observation and experience, where it is so opposed to all reasonable probabilities as to be manifestly false, or is contrary to the laws of nature or to well-known scientific principles . . . , it is to be disregarded as being without evidentiary value even though uncontradicted." (alteration in original) (quoting Hollis v. Scott, 516 So. 2d 576, 578-79 (Ala. 1987))); Whelen v. Osgoodby, 50 A. 692, 694 (N.J. Ch. 1901) ("The statement that a man under certain circumstances did something which we know from experience that not one in a thousand would do under the same circumstances is discredited by the inherent improbability of the statement.").

122 *See, e.g.*, Spear v. United R.Rs. of S.F., 117 P. 956, 965 (Cal. Dist. Ct. App. 1911) ("[P]robability of a higher degree of attention and interest gives rise to a presumption of considerable force that a person's recollection of his act and of attendant circumstances is more definite and trustworthy than another person's recollection of it."); State v. Long, 108 A. 36, 37 (Del. Ct. Oyer & Terminer 1919) ("The actor usually knows better than the observer what he did or did not do, and his testimony is generally entitled to greater weight.").

123 *See, e.g.*, Herbert v. Lankershim, 71 P.2d 220, 230 (Cal. 1937) (en banc) ("On the subject of oral admissions, unless corroborated by satisfactory evidence, this court . . . rates it as the weakest of testimony that can be produced." (citation omitted)).

124 *See, e.g.*, The Aakre, 122 F.2d 469, 476 (2d Cir. 1941) ("Cargo has a slim chance at best, for it must rely exclusively on the testimony of the ship's officers and crew as to what preceded the casualty; and admiralty courts have often recognized that such witnesses almost invariably will testify on behalf of their ship."); Willis v. Pa. R.R. Co., 122 F.2d 248, 250-51 (2d Cir. 1941) (Frank J., dissenting) (applying this rule in a railroad negligence case).

some weight by a jury;¹²⁵ and that a jury may disregard a witness' testimony when it contradicts the witness' own conduct.¹²⁶ In each of these cases, the court cited a case or treatise (often Moore's) in support of its generalization about the reliability of certain forms of testimony.

And today, administrative agencies quite commonly rely on rules of weight. For instance, according to a well-known rule that the Social Security Administration uses in adjudicating benefits claims, the testimony of the treating physician is generally deemed to have more weight than that of other physicians.¹²⁷ Similarly, the Court of Federal Claims has held that the records of a family physician deserve more weight than oral testimony offered later.¹²⁸ The United States

125 See, e.g., *Chesapeake & Ohio Ry. Co. v. Martin*, 283 U.S. 209, 216 (1931) ("We recognize the general rule, of course, as stated by both courts below, that the question of the credibility of witnesses is one for the jury alone; but this does not mean that the jury is at liberty, under the guise of passing upon the credibility of a witness, to disregard his testimony, when from no reasonable point of view is it open to doubt."); *Cruz-Vargas v. R.J. Reynolds Tobacco Co.*, 348 F.3d 271, 277 (1st Cir. 2003) (noting that a witness' "testimony is furthermore neither improbable nor contradicted, thus falling under the established principle that 'testimony concerning a simple fact, capable of contradiction, not incredible, and standing uncontradicted, unimpeached, or in no way discredited by cross-examination, must be permitted to stand'" (quoting *Chi., Rock Island & Pac. Ry. Co. v. Howell*, 401 F.2d 752, 754 (10th Cir. 1968))); *Broad. Music, Inc. v. Havana Madrid Rest. Corp.*, 175 F.2d 77, 81 (2d Cir. 1949) (upholding the judgment of a trial court on the ground that the court "had the 'uncontradicted witness' rule in mind, and, accordingly, pointed to the witness' interest and to the lack of corroboration in order to bring the case within the exception").

126 See, e.g., *Old Republic Ins. Co. v. Alexander*, 436 S.W.2d 829, 836 (Ark. 1969) ("When the conduct of any witness is clearly inconsistent with his testimony and not satisfactorily explained, the trier of facts is justified in disbelieving the testimony."); *Field v. Koonce*, 12 S.W.2d 772, 775 (Ark. 1929) ("'Conduct of a witness clearly inconsistent with his testimony and not satisfactorily explained is one of the most fatal species of impeachment.'" (quoting 2 MOORE, *supra* note 5, § 1136)).

127 See 20 C.F.R. § 404.1527(d)(2) (2007) ("Generally, we give more weight to opinions from your treating sources, since these sources are likely to be the medical professionals most able to provide a detailed, longitudinal picture of your medical impairment(s) and may bring a unique perspective to the medical evidence that cannot be obtained from the objective medical findings alone or from reports of individual examinations, such as consultative examinations or brief hospitalizations."); *Robinson v. Barnhart*, 366 F.3d 1078, 1084 (10th Cir. 2004) ("The opinion of an examining physician is generally entitled to less weight than that of a treating physician, and the opinion of an agency physician who has never seen the claimant is entitled to the least weight of all."). *But see Black & Decker Disability Plan v. Nord*, 538 U.S. 822, 829–34 (2003) (holding that the treating physician rule does not apply to ERISA claims).

128 See, e.g., *Estate of Arrowood v. Sec'y of the Dep't of Health & Human Servs.*, 28 Fed. Cl. 453, 458 (1993) (noting that the rules of evidence do not apply to special

International Trade Commission has held that the testimony of the employees of an owner of an alleged trademark is not accorded much weight.¹²⁹ For analogous reasons, the Environmental Protection Agency generally gives little weight to self-serving statements by corporate officers about a company's ability to pay a fine.¹³⁰ All of these doctrines constitute efforts by agencies to establish rules or guidelines as to how much probative value certain types of testimony ought to be accorded.

* * *

The above discussion does not establish that rules of weight are a pervasive feature of common law adjudication. Outside the agency context, the number of cases in which civil and criminal courts have employed rules of weight in the way Moore envisioned is relatively low. But it does suggest that despite our rhetoric about the importance of juries, most adjudication in this country takes place without them. And when a jury is used, its discretion is curbed in a variety of ways. Given such a state of affairs, analyzing the virtues and vices of rules of weight seems worthwhile. Of course, the relative significance of the jury is not equal in all adjudicative contexts, so to the extent that rules of weight do limit jury discretion, their use may well require limitation in some contexts more than others. Using certain rules of weight in criminal trials, for instance, may implicate Sixth Amendment concerns not present elsewhere. But unless one doubts the constitutionality of virtually any limitation on the right to a jury trial,¹³¹

masters and that "a record created in the normal course by an employee of the family physician should receive more weight than oral testimony given years after the event"; see also *Hoag v. Sec'y of Dep't of Health & Human Servs.*, No. 94-67V, 1998 WL 408783, at *3 (Fed. Cl. Apr. 22, 1998) ("The Court of Appeals for the Federal Circuit has counseled the Special Masters that medical records should be given more weight than conflicting oral testimony offered after the fact.").

129 See, e.g., *Certain Braiding Machs.*, USITC Pub. 1435, Inv. No. 337-TA-130 (Oct. 1983) ("Generally speaking, little weight is accorded to the testimony of an employee of the owner of the alleged trademark because of the potential risk that the perceptions conveyed are colored by bias.").

130 See, e.g., *Zaclon, Inc.*, No. RCRA-05-2004-0019, 2006 EPA ALJ LEXIS 23, at *23 (May 23, 2006) ("[S]elf-serving testimony by corporate officers, uncorroborated by documentation, is generally given little weight regarding inability to pay."); *F & K Plating Co.*, 2 E.A.D. 443, 449 (1987) (noting that "although unsupported self-serving testimony is generally entitled to little weight," the presiding officer had met his burden of proving that he could not pay the fine); *Cent. Paint & Body Shop, Inc.*, 2 E.A.D. 309, 315 (1987) ("Self-serving declarations are entitled to little weight.").

131 See, e.g., *Suja A. Thomas, Why Summary Judgment Is Unconstitutional*, 93 VA. L. REV. 139, 145-60 (2007) (arguing that the common law in 1791 had no mechanism

using such rules does not seem to threaten the very foundations of Anglo-American evidence law in the way Wigmore feared.

There are, however, other powerful objections to the use of rules of weight that do not depend on the jury's role in our system. Thus, while recognizing that judges and juries are actors with different intellectual capacities who possess distinct institutional advantages and disadvantages,¹³² the balance of this Article takes up the more general question of whether using rules to weigh the credibility of testimony may be beneficial for adjudicative factfinding, irrespective of whether the factfinder is a judge or jury.

II. THE CASE FOR RULES OF WEIGHT

Answering this question requires first clarifying what rules of weight are and how exactly they differ from other, related evidentiary devices. It then requires showing how such rules may improve adjudicative factfinding by making it more accurate, fair, and efficient. Let me begin by explaining how the term "rule" is used throughout.

A. *Rules in General*

Rules are, first and foremost, generalizations. As Professor Fred Schauer puts it, they "speak to types and not to particulars."¹³³ But rules are not just generalizations, they are *entrenched* generalizations. To say that a generalization is *entrenched* is to say that it exerts normative force—that is, provides a reason for action—even in circumstances where following the rule would not appear to fulfill its underlying justification.¹³⁴ Importantly, this conception of a rule is sufficiently broad to include two types of directives with which rules are sometimes contrasted.¹³⁵ First, it includes so-called "rules of

similar to summary judgment and the practice is unconstitutional in violating the Seventh Amendment).

132 Most obviously, judges are repeat players and thus may develop skills and habits that jurors likely do not. See Mitchell, *supra* note 15, at 1122 ("[J]udges are effectively repeat players for most judgment and decision problems, whereas juries are not.").

133 FREDERICK SCHAUER, *PLAYING BY THE RULES* 18 (1991).

134 See *id.* at 47–49.

135 For a proponent of the view that a directive only qualifies as a rule if it applies in every case, see RONALD DWORKIN, *TAKING RIGHTS SERIOUSLY* 24 (1977) ("Rules are applicable in an all-or-nothing fashion. If the facts a rule stipulates are given, then either the rule is valid, in which case the answer it supplies must be accepted, or it is not, in which case it contributes nothing to the decision."). For a discussion distinguishing between rules and standards, see Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 *DUKE L.J.* 557 (1992).

thumb”—that is, rules that do not necessarily require application in every case in which the factual predicate for the rule holds. As Professor Schauer explains, if a rule of thumb requires that the rule-follower attain a higher degree of certainty before deviating from the rule, then it exerts normative force.¹³⁶ Whether rules of thumb count as rules matters because many rules of both weight and admissibility are rules of thumb.¹³⁷

More important, this conception of rules includes directives that are framed in broad or open-ended language and are thus frequently dubbed “standards.” This inclusion is warranted because the key issue in determining whether a decision can accurately be described as rule-guided is whether the rule itself makes a difference to the decisionmaker, not simply whether it is verbally specific.¹³⁸ To use the classic example, if the paradigmatic rule is a fifty-five miles per hour speed limit and the equivalent standard requires that people “drive reasonably,” a directive that instructs drivers to “drive very slowly” on a wide-open road may very well do normative work and could thus qualify as a rule despite the vagueness of its terms.¹³⁹ In fact, many rules of weight are of this sort insofar as they instruct the factfinder to accord “low weight” to a certain class of evidence or simply assert that it is “generally unreliable.”¹⁴⁰ In short, whether a rule of weight ought to be an absolute rule or a mere rule of thumb and whether it ought to be framed in precise language or vague language are important factors to be considered in devising it, but such considerations point to variations among, not alternatives to, rules of weight.

This definition also allows us to distinguish rules of weight from the closely related legal doctrine of judicial comment, which permits judges to give their opinion to the jury about evidence presented at

136 See SCHAUER, *supra* note 133, at 106.

137 See, e.g., FED. R. EVID. 611(b) (limiting the scope of cross examination to those matters covered in direct examination or matters affecting witness credibility but subjecting such limitation to the discretion of the judge, who may “permit inquiry into additional matters as if on direct examination”); *infra* notes 154–56 and accompanying text (discussing rules of weight that are rules of thumb).

138 See SCHAUER, *supra* note 133, at 104 n.35 (noting that “the impact of rules on the behavior of decision-makers is more than can be captured in a dimension of specificity and vagueness”).

139 Thus, even Professor Kaplow, whose conception of a rule places less significance on a directive’s generality than on whether its content is determined before or after its application, stresses that a law that uses broad language, such as one that prohibits “vulgar behavior,” may plausibly be interpreted as a rule to the extent that rule-followers would agree on what likely constitutes such behavior. See Kaplow, *supra* note 135, at 601.

140 See *supra* text accompanying notes 121–26.

trial.¹⁴¹ A judge who simply opines to the jury that “witness A strikes me as untrustworthy and his testimony utterly implausible” is in no sense applying a *rule*. The difference is obvious but important, because whereas judicial comment only makes sense in the context of a bifurcated tribunal (since one can hardly “comment” to one’s self), rules of weight can be employed whether the factfinder and the legal decisionmaker are the same or not. Furthermore, objections that hold for one practice may not necessarily hold for the other. Indeed, it is precisely in virtue of their status as entrenched generalizations that many critics object to rules of weight.¹⁴²

B. Rules of Weight

In Part I, I surveyed a variety of evidentiary devices that constrain the jury’s discretion in one way or another. There I also promised to explain why a corroboration rule was better described as a rule of weight than a sufficiency of the evidence rule. Here I will make good on that promise and also show how rules of weight differ from legal presumptions and admissibility rules. A more precise conception of rules of weight will thus emerge.

1. Corroboration Rules and Sufficiency Rules

At first glance, a corroboration rule that prevents a case from reaching a jury unless a certain type of testimony is corroborated seems to be a quite different sort of rule than one that allows evidence to be submitted to the jury with an instruction to accord it “low weight.” Whereas the latter rule attempts to guide the factfinding process, the former seems to bypass that process altogether. Perhaps for this reason, the Court in *Carmell* called the Texas corroboration rule at issue in that case a “sufficiency of the evidence rule.”¹⁴³ The term is

141 Professor Renée Lerner has documented well the decline of judicial comment in this country. See generally Renée Lettow Lerner, *The Transformation of the American Civil Trial: The Silent Judge*, 42 WM. & MARY L. REV. 195, 241–61 (2000) (documenting the decline of judicial comment in this country).

142 But for others, judicial comment is only legitimate insofar as it entails the application of a rule of weight. See *Spalding v. Lowe*, 23 N.W. 46, 48 (Mich. 1885) (“As a general rule, it is improper for the trial judge to instruct the jury that the evidence of one witness is deserving of more weight than that of another. In doing so he invades the province of the jury, whose function it is to determine from the evidence whether any fact in issue is sufficiently proved or not *He may, however, define the weight which the law attaches to a whole class of testimony*,—for instance, that of accomplices,—but he may not single out certain testimony and tell the jury it is entitled to much or little weight.” (emphasis added)).

143 *Carmell v. Texas*, 529 U.S. 513, 517–18 (2000).

not inaccurate, because the rule purports to make a judgment as to the conditions that must hold for a given type of evidence (in this case, the testimony of a victim of sexual assault) to be sufficient to sustain a conviction. Still, the term is somewhat misleading both because its meaning is ambiguous and because not all corroboration rules are sufficiency rules.

The term is ambiguous because courts distinguish between the legal and factual sufficiency of evidence. A verdict is said to be *legally* insufficient when, for instance, the court improperly instructs the jury so that its judgment about the evidence is premised on a misunderstanding or omission of one of the statutory elements of the crime.¹⁴⁴ A verdict is *factually* insufficient when the weight of the evidence does not support the conviction even under a proper view of the law, perhaps because the only evidence offered was the testimony of one interested witness and a few pieces of circumstantial evidence.¹⁴⁵ This distinction holds even though, under *Jackson v. Virginia*,¹⁴⁶ it is technically legal error for a jury to convict a defendant when the verdict is not supported by the weight of the evidence.¹⁴⁷

The corroboration rule in *Carmell* is concerned only with the *factual* sufficiency of the evidence. That is, it reflects the legislature's concern about the evidentiary reliability of a certain class of testimony and its preference for placing the risk of factfinding error associated with such testimony on the prosecution rather than the criminal

144 See, e.g., *Yates v. United States*, 354 U.S. 298, 312 (1957) (reversing a jury's conviction under the Smith Act on the ground that the only evidence offered in support of one of the elements occurred at a time beyond the applicable statute of limitations), *overruled in part* by *Burks v. United States*, 437 U.S. 1, 10 (1978).

145 See *Griffin v. United States*, 502 U.S. 46, 58–59 (1991).

146 443 U.S. 307 (1979).

147 See *id.* at 318–19; see also *Griffin*, 502 U.S. at 58–59 (“Insufficiency of proof . . . is legal error.”). Texas is one state where courts have distinguished carefully between legal and factual insufficiency. Compare *Mar. Overseas Corp. v. Ellis*, 971 S.W.2d 402, 407 (Tex. 1998) (explaining, in reversing a trial court on factual insufficiency grounds, that “the court of appeals must detail all the evidence relevant to the issue and clearly state why the jury’s finding is factually insufficient or so against the great weight and preponderance of the evidence that it is manifestly unjust”), with *Merrell Dow Pharm., Inc. v. Havner*, 953 S.W.2d 706, 711 (Tex. 1997) (noting that a court may reverse a judgment on a legal insufficiency or “no evidence” challenge where “(a) there is a complete absence of evidence of a vital fact, (b) the court is barred by rules of law or of evidence from giving weight to the only evidence offered to prove a vital fact, (c) the evidence offered to prove a vital fact is no more than a mere scintilla, or (d) the evidence conclusively establishes the opposite of the vital fact”). For further information on this distinction, see Harvey Brown, *Procedural Issues Under Daubert*, 36 HOUS. L. REV. 1133, 1191 (1999) (noting the distinction between the two types of insufficiency).

defendant.¹⁴⁸ It does not reflect a policy judgment as to what conduct ought to be punished and what ought not; surely the state of Texas would like *all* sexual offenders to be punished if possible.¹⁴⁹ Of course, as this example makes clear, to say that a corroboration rule is primarily concerned with the reliability of testimony does not mean that its sole purpose is to increase the accuracy of verdicts. The rule may be designed not only to reduce factfinding errors, but also to allocate fairly the risks of such errors or to minimize the cost of reducing them.¹⁵⁰ Still, the purpose is evidently to regulate to the factfinding process.

A *legal* sufficiency of the evidence rule, on the other hand, would simply amount to a substantive rule of law. A court might hold, for instance, that evidence showing that a defendant who secured a victim's consent to get into his car by fraud or deceit is insufficient to constitute "forced asportation" for the purposes of a kidnapping statute.¹⁵¹ Such a holding may be called a "sufficiency of the evidence rule" because, if there is no evidence of physical force, then the verdict is in error. But this is true not because of any defect in the evidence or the jury's evaluation of it, but because courts have decided as a matter of substantive law that such facts do not amount to forced asportation.¹⁵²

Corroboration rules such as those in *Carmell* are, then, factual sufficiency of the evidence rules, but they are also rules of weight that serve a function very similar to those rules—such as those cited in Part I from Moore's treatise—that call for certain evidence to be admitted but only as a lower grade of evidence.¹⁵³ The difference between the two types of rules is exaggerated by the fact that in Anglo-American law, factfinding takes place in a bifurcated tribunal so that the judge must make an initial determination as to whether there is sufficient evidence to get to a jury. But it is not difficult to see that a

148 See STEIN, *supra* note 9, at 208.

149 Perhaps it is more accurate to say that Texas would like to punish all sex offenders possible at a reasonable social cost.

150 See STEIN, *supra* note 9, at 1.

151 In this context, asportation means the taking of a person. The example comes from *People v. Green*, 609 P.2d 468, 507 (Cal. 1980) (en banc) ("To constitute consent on the part of a person to a criminal act or transaction, he must act freely and voluntarily and *not under the influence of fraud, threats, force or duress.*" (internal quotation marks and citation omitted)).

152 See *People v. Guiton*, 847 P.2d 45, 52 (Cal. 1993) (en banc) (concluding that a verdict is legally insufficient where "the inadequacy is legal, not merely factual, that is, when the facts do not state a crime under the applicable statute . . . absent a basis in the record to find that the verdict was actually based on a valid ground").

153 See *supra* notes 58–87 and accompanying text.

rule that deems a certain type of testimony to be “of low weight” or “generally unreliable” will in most cases be functionally analogous to a rule that deems the same testimony “*generally* insufficient to sustain a verdict on its own.”

And this points to the second reason why corroboration rules are better described as rules of weight than as sufficiency of the evidence rules. Not all corroboration rules are, strictly speaking, sufficiency rules. Instead, some corroboration rules serve as *rules of thumb* as to the likely adequacy of the evidence in a way comparable to the rules described in the previous paragraph. For instance, one corroboration requirement that existed in England until recently required that, if the testimony of an accomplice was uncorroborated by other evidence, the judge was to instruct the jury that it was “dangerous” to convict purely on the basis of such testimony.¹⁵⁴ Similarly, the Board of Immigration Appeals has interpreted the regulations governing asylum petitions to require that “where it is reasonable to expect corroborating evidence for certain alleged facts [e.g., of persecution in the applicant’s country of origin] pertaining to the specifics of an applicant’s claim, such evidence should be provided.”¹⁵⁵ But if the applicant cannot produce such corroborating evidence, the result is not a denied application; instead, the applicant has the opportunity to explain her failure to do so.¹⁵⁶

Thus, the relevant distinction is not between rules of weight, on the one hand, and corroboration and sufficiency rules, on the other. Rather, the distinction is between those rules of weight that call for evidence to be admitted to the factfinder, but only as a lower grade of evidence, and those rules whose application deems a particular class of evidence to be insufficient to sustain a verdict *in all cases*. We may call the former type *defeasible* rules of weight and the latter *absolute* rules of weight. Whether a rule of weight deserves to be called a “corroboration rule” simply depends on whether or not the rule issues even when corroborating evidence is presented.

154 See CORROBORATION OF EVIDENCE, *supra* note 27, at 3–7.

155 S-M-J, 21 I. & N. Dec. 722, 725 (1997). The relevant regulations can be found at 8 C.F.R. § 208.13(a) (2007).

156 See S-M-J, 21 I. & N. Dec. at 725. Of course, the distinction can be a fine one. The BIA rule could plausibly be reinterpreted as a type of sufficiency rule according to which an application must be denied where it is reasonable to expect corroborating evidence, where none is provided, *and where no explanation for the failure to provide it is offered*.

2. Presumptions and Admissibility Rules

It may appear that rules of weight are simply a subclass of legal presumptions. Like rules of weight, presumptions may also be either mandatory or permissive.¹⁵⁷ And, to be sure, a hypothetical rule of weight that deems DNA evidence to be of more weight than fingerprint evidence could plausibly be described as a “presumption” that the former is more reliable than the latter on the issue of identification. In fact, though, most legal presumptions work somewhat differently than rules of weight. A presumption encourages or requires the factfinder to draw particular inferences from one fact in order to establish the existence of another fact *regardless of the type of evidence used* to establish the first fact. Rules of weight, on the other hand, apply to specific types of evidence, such as written affidavits¹⁵⁸ or certain classes of witness testimony.¹⁵⁹ In other words, the purpose of rules of weight, but not legal presumptions, is to grade the reliability of certain classes of evidence.

Of course, the purpose of some common admissibility rules, such as the hearsay rule or the original documents rule,¹⁶⁰ is also to account for the reliability of evidence. Conversely, a corroboration rule might be interpreted as a type of conditional admissibility rule since, if the corroboration requirement is met, the testimony is admitted without any special instruction or qualification.¹⁶¹ So it may seem that rules of weight ought to be classed as a type of admissibility rule.

But rules of weight differ in important respects from admissibility rules. First, many admissibility rules are not about evidentiary weight at all. Those admissibility rules whose purpose is entirely unrelated to factfinding (at least in the case at hand), such as the various evidentiary privileges, which exclude information on the ground that admitting it might have harmful policy consequences, are not rules of

157 In fact, there are more types of presumptions than just these two. See LEMPERT ET AL., *supra* note 99, at 1295–306 (analyzing various types of presumptions and explaining how each functions relative to the burdens of production and persuasion).

158 See 2 MOORE, *supra* note 5, § 941, at 1097–99.

159 See *supra* notes 121–26.

160 See FED. R. EVID. 801–802 (defining and generally excluding hearsay evidence); *id.* 804(b) (enumerating hearsay exceptions); *id.* 1001 (requiring the “best evidence” available); *id.* 1004 (providing exceptions to the original document rules).

161 See *Carmell v. Texas*, 529 U.S. 513, 564 (2000) (Ginsburg, J., dissenting) (arguing that a Texas corroboration rule was best interpreted as a conditional rule of admissibility according to which “the testimony of the victim shall be inadmissible to prove the defendant’s guilt unless corroborated” (internal quotation marks omitted)).

weight.¹⁶² Second, and more important, whereas rules of admissibility offer only one remedy when applied—exclusion of the evidence from the factfinding process—rules of weight admit of two further possibilities: the evidence may be admitted as a lower grade of evidence (defeasible rules of weight) or it may settle the issue as a matter of law (absolute rules of weight).

Table 1 summarizes the above discussion. Only the rules in the lower-middle and right-hand cells are rules of weight.

TABLE 1.

		<i>Consequences of Applying Rule</i>		
		<i>Excludes Evidence</i>	<i>Admits as Lower Form of Proof</i>	<i>Settles Issue as Matter of Law</i>
<i>Purpose of Rule</i>	<i>Policy/ Substantive</i>	Admissibility Rules (e.g., Privileges)	X [†]	Substantive Rules of Law or “legal sufficiency of the evidence rules” (e.g., fraud is not forced asportation)
	<i>Weight/ Evidentiary</i>	Admissibility Rules (e.g., Hearsay)	<i>Defeasible Rules of Weight</i> (e.g., Moore’s rules, BIA corroboration rule)	<i>Absolute Rules of Weight</i> or “factual sufficiency of the evidence rules” (e.g., <i>Carmell</i> corroboration rule)

† Whether rules of weight do and should serve independent policy goals is for the most part not considered in this Article, though I raise the question in the conclusion.

C. *Potential Benefits of Rules of Weight*

The dominant trend in evidence scholarship over the past couple of centuries has been toward a system of “free proof” in which rules play an increasingly minor role.¹⁶³ Recently, however, some scholars have recognized that rules may play a crucial function in regulating factfinding. These scholars argue that the traditional virtues of rule-based decisionmaking in the context of substantive law apply with equal force in the factfinding domain. Below I explain how rules in

162 See, e.g., FED. R. EVID. 403 (excluding evidence with a prejudicial effect that substantially outweighs its probative value); see also STEIN, *supra* note 9, at 30 (arguing that, for the reason stated in the text, such rules ought to be excluded from the domain of evidence law).

163 See STEIN, *supra* note 9, at 108 (noting that the “abolitionist claim has permeated legal discourse for approximately two centuries”).

general may be beneficial for factfinding and then argue that rules of weight, specifically, may help improve the accuracy, fairness, and efficiency of factfinding.

1. Accuracy of Verdicts

Rules are typically said to be over- and underinclusive with respect to their underlying rationales—and thus lead to less accurate decisions than do standards. Professor Louis Kaplow, however, has shown that such a judgment assumes that rules are simple, which is not necessarily the case. If a rule is sufficiently complex and takes into account all relevant factors bearing on a question, Kaplow explains, it may be that it results in more accurate decisions than a more open-ended standard, especially if the rule-follower, when left employing such a standard, would not appropriately consider all the relevant factors.¹⁶⁴ Thus, rules will improve the accuracy of decisions on balance in contexts where the decisionmaker is prone to make more errors when using his or her discretion than result from any over- or under-inclusiveness of rules.¹⁶⁵

Adjudicative factfinding may well be such a domain. On almost any account of evidence law, a primary, though certainly not exclusive, goal of evidence law is and ought to be the production of accurate verdicts.¹⁶⁶ And the danger that jurors will misevaluate the evidence and reach inaccurate verdicts is the classic justification for rules of admissibility.¹⁶⁷ But Professor Schauer has suggested recently that it is not only jurors who would benefit from the discretion-constraining effect of rules.¹⁶⁸ Citing an impressive body of cognitive psychology literature showing both that individuals are prone to a variety

164 See Kaplow, *supra* note 135, at 594.

165 See SCHAUER, *supra* note 133, at 150.

166 See TWINING, *supra* note 8, at 76 tbl.1 (listing as one of the common assumptions of the “rationalist tradition” of evidence scholarship the view that “[t]he pursuit of truth (i.e. seeking to maximize accuracy in fact-determination) is to be given a high, but not necessarily an overriding, priority in relation to other values, such as the security of the state, the protection of family relationships or the curbing of coercive methods of interrogation”); see also LARRY LAUDAN, *TRUTH, ERROR, AND CRIMINAL LAW* 1 (2006) (arguing that one of the “basic aims” of criminal procedural law is “to find out the truth about a crime and thus avoid false verdicts”). But see STEIN, *supra* note 9, at x (“[T]he key function of evidence law is to apportion the risk of error in conditions of uncertainty, rather than facilitate the discovery of the truth.”).

167 See CHRISTOPHER B. MUELLER & LAIRD C. KIRKPATRICK, *EVIDENCE UNDER THE RULES* 1 (5th ed. 2004) (“[M]istrust of juries is the single overriding reason for the law of evidence.”).

168 See Frederick Schauer, *On the Supposed Jury-Dependence of Evidence Law*, 155 U. PA. L. REV. 165, 187 (2006).

of cognitive errors and that experts in particular tend to overestimate their own professional judgments, Schauer challenges the conventional wisdom that rules of admissibility need only apply in jury trials.¹⁶⁹

Schauer does not himself suggest using rules of weight in factfinding, but doing so would likely be a far more effective remedy for the problem he identifies.¹⁷⁰ For when a judge serves as both legal decisionmaker and factfinder, it may be impossible for her to ignore entirely a piece of evidence she has deemed inadmissible. Instead, judges probably discount such evidence, noting its potential probative value while recognizing its weaker credentials.¹⁷¹ Thus, such situations seem to call naturally for rules of weight, rather than rules of exclusion. Still, any argument that rules of weight would be effective in this way requires a more thorough explanation of what cognitive defects judges (and juries) might suffer from and how these rules could cure such defects.

a. Base Rates and Weight

The idea that humans are not perfectly rational creatures is hardly new, but only in the last thirty years have researchers rigorously analyzed the precise ways in which people systematically fail to satisfy certain normative standards of rationality. By now the impressive body of “heuristics and biases” research within cognitive psychology, most famously developed by Daniel Kahneman and Amos Tversky, is old news among legal scholars.¹⁷² It has inspired a whole new field of behavioral economics, which focuses on the ways in which people’s actual consumption choices and behavior deviate from the traditional models of rational preference maximization used in decision theory and economics.¹⁷³ It is even older news among evidence scholars,

169 See *id.* at 186–92.

170 In her response to Schauer’s piece, Professor Mnookin makes this point. See Mnookin, *supra* note 9, at 142 (noting that “it might be that without a bifurcated tribunal, the only rules of evidence that could genuinely operate as rules during a bench trial would be rules of assessment rather than admissibility”).

171 For empirical support for this claim, see *infra* note 204 and accompanying text.

172 See Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 *SCIENCE* 1124, 1124–31 (1974). A search for “Tversky /s Kahneman” within US Law Reviews and Journals, Combined yielded 1357 results on LexisNexis; “availability heuristic” yielded 514; “hindsight bias” yielded 621; “endowment effect” yielded 674; “framing effect” yielded 403; “representativeness heuristic”—the most relevant heuristic for our purposes—yielded 218 (last accessed Mar. 26, 2008).

173 See, e.g., Christine Jolls & Cass R. Sunstein, *Debiasing Through Law*, 35 *J. LEGAL STUD.* 199, 203 (2006); Jeffrey J. Rachlinski, *The Uncertain Psychological Case for Paternalism*, 97 *Nw. U. L. REV.* 1165, 1170–75 (2003).

who have long recognized its potential impact for adjudicative factfinding.¹⁷⁴

To summarize crudely, this research suggests that when people make predictions or probabilistic estimates, they focus too much on particulars and not enough on abstract information.¹⁷⁵ More specifically, when people are asked to make conditional probability estimates (i.e., to determine the probability of an hypothesis H , given evidence E), they tend not to take sufficient account of either the *base rate* of H (how likely H is before E is considered) or the *weight* of E (how reliably E correlates with H). To see these concepts at work, imagine a gumball guru who can predict with eighty percent accuracy what color gumball will be dispensed by a gumball machine. If the gumball machine has seventy-five blue gumballs and twenty-five red gumballs, then the *base rate* of red gumballs is twenty-five percent, while the *weight* or reliability of the guru's prediction is eighty percent. By using the Bayes Theorem, one can calculate the probability that the gumball will be red when the guru predicts that it will be red.¹⁷⁶

174 See Mitchell, *supra* note 15, at 1066 n.1 (citing sources and noting that evidence scholars were the first to recognize the potential significance of the heuristics and biases research program for the legal system); see also sources cited *supra* note 15.

175 See, e.g., Dale Griffin & Roger Beuhler, *Frequency Probability and Prediction: Easy Solutions to Cognitive Illusions?*, 38 COGNITIVE PSYCHOL. 48, 49 (1999) (describing the psychological phenomenon of looking to particular cases rather than sets of cases as the "power of the particular"); Daniel Kahneman & Dan Lovallo, *Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking*, 39 MGMT. SCI. 17, 22–23 (1993) (claiming that decisionmakers are excessively prone to treat problems as unique and to neglect background statistical information); Daniel Kahneman & Amos Tversky, *Intuitive Prediction: Biases and Corrective Procedures*, in JUDGMENT UNDER UNCERTAINTY 414, 415 (Daniel Kahneman et al. eds., 1982) (finding that in making predictions people rely not enough on general "distributional data" and too much on evidence about particular cases); Gideon Keren, *On the Ability of Monitoring Non-Veridical Perceptions and Uncertain Knowledge: Some Calibration Studies*, 67 ACTA PSYCHOLOGICA 95, 115 (1988) (finding that subjects were more likely to be overconfident about intellectual tasks than perceptual ones); Richard E. Nisbett et al., *Popular Induction: Information Is Not Necessarily Informative*, in JUDGMENT UNDER UNCERTAINTY, *supra*, at 101, 111–12 (speculating that a variety of psychological findings in heuristics research can be explained by the fact that people tend to ignore information that is "remote, pallid and abstract" and pay attention to that which is "vivid, salient and concrete").

176 First, we must assume that the machine dispenses gumballs randomly, that each ball is replaced after it is dispensed, and that the guru does not know how many of each color gumball are in the machine (perhaps because the machine is opaque). Let $P(R)$ refer to the *base rate* of red balls, i.e., the probability of a red gumball being dispensed at random, which in this case is 0.25. And let G_R refer to the guru's prediction that a gumball is red. So $P(G_R/R)$ represents the probability that the guru will correctly predict a gumball to be red when the ball is in fact red. This measures the *weight* or reliability of the guru's predictions, and here that value is 0.8. Let us also

This example helps distinguish between two senses of the term “weight.” In a broad sense, weight simply refers to the probative value or predictive validity of *any* relevant class or piece of evidence in making a probabilistic judgment. In that sense, one could speak intelligibly of the need to give appropriate weight to base rates in making certain predictions or according proper weight to circumstantial evidence in a crime trial. But weight in the narrower sense refers to the *reliability* of direct evidence of something. That is, it measures how reliably something that purports to indicate when X is true or when X occurs is in fact probative of X being true or X occurring. Weight in this sense describes the relationship between a witness’ testimony and the facts to which that witness is testifying. In the example above, it measures the gumball guru’s predictive accuracy.

That people pay insufficient attention to base rates in making predictions is now widely known and its consequences for forensic factfinding much discussed. But human errors in accounting for the weight or reliability of evidence are no less common. In one study, for instance, subjects were asked to predict a hypothetical student’s choice of graduate school based on a description of him that had been itself based on psychological test he took in high school.¹⁷⁷ When told that the student had chosen a different field than the one they had predicted and asked to explain why that might be, most subjects either looked to other aspects of the description to explain his choice in graduate work or hypothesized alternative psychological motivations for his choice of school other than intellectual fit.¹⁷⁸ Very few subjects considered (1) the overall distribution of people in the various graduate schools (i.e., the base rate) or (2) the predictive

assume that the probability that he will guess it to be red when it is really blue, or $P(G_R/B)$, is 0.2. What we are looking for is the probability that the ball is red, given the new evidence that the guru has predicted it to be red, or $P(R/G_R)$. According to the Bayes formula, $P(R/G_R) = [P(R) \times P(G_R/R)] / ([P(R) \times P(G_R/R)] + [P(B) \times P(G_R/B)])$. Plugging those numbers in, we get: $(0.25 \times 0.8) / [(0.25 \times 0.8) + (0.75 \times 0.2)] = 0.571$. So, if the gumball guru predicts that the gumball will be red, there is about a 57% chance that it will in fact be red. See generally G.A. Barnard, *Thomas Bayes’s Essay Towards Solving a Problem in the Doctrine of Chances*, 45 BIOMETRIKA 293 (1958) (reproducing and explaining the Bayes Theorem).

177 The findings of this study were discussed in two separate papers. See Daniel Kahneman & Amos Tversky, *On the Psychology of Prediction*, 80 PSYCHOL. REV. 237, 238–39 (1973) [hereinafter Kahneman & Tversky, *Prediction*]; Amos Tversky & Daniel Kahneman, *Causal Schemas in Judgment Under Uncertainty*, in JUDGMENT UNDER UNCERTAINTY, *supra* note 175, at 117, 117–28 [hereinafter Tversky & Kahneman, *Causal Schemas*].

178 See Tversky & Kahneman, *Causal Schemas*, *supra* note 177, at 127–28.

validity (i.e., weight) of such personality tests as possible explanations for the discrepancy.¹⁷⁹

In another study, researchers measured whether and to what extent subjects took the size of a sample into account in making predictive judgments.¹⁸⁰ They asked the subjects to make a probabilistic judgment as to the likelihood that a given coin, when spun, was biased in favor of heads or tails. In this context, the “weight” of the evidence is the sample size (i.e., total number of times the coin is flipped), and the “strength” of the evidence is the proportion of heads to tails.¹⁸¹ The researchers provided the subjects with sample tosses of a variety of different weights and strengths and, relative to what Bayesian analysis requires, they found that (1) the predictions were relatively inelastic to the weight or sample size, with the result that (2) subjects systematically overestimated the likelihood of bias when the strength was high and the weight was low.¹⁸² This and comparable studies led the researchers to conclude that “people assess their confidence in one of two competing hypotheses on the basis of their balance of arguments for and against this hypothesis, with insufficient regard for the quality of the data.”¹⁸³ They thus speculated that when people evaluate a letter of recommendation they tend to pay too much attention to how glowing or damning the recommendation is and insufficiently account for the writer’s limited knowledge of the subject of it.¹⁸⁴

These findings are also consistent with the numerous studies comparing clinical and actuarial forms of predictive and diagnostic judgments.¹⁸⁵ These studies suggest that when experts in such fields

179 Kahneman & Tversky, *Prediction*, *supra* note 177, at 239. Of course, it would not be fair to assume that the subjects knew as an empirical matter what such a distribution looked like, but the point is that they did not even consider their own subjective view of graduate school distribution. A control group in the experiment suggests that most students believed there to be many more students in the humanities and education than in computer science and engineering. *See id.* at 238.

180 *See* Dale Griffin & Amos Tversky, *The Weighing of Evidence and the Determinants of Confidence*, 24 *COGNITIVE PSYCHOL.* 411, 414 (1992).

181 *Id.*

182 *See id.* at 416.

183 *Id.* at 425.

184 *See id.* at 413; *see also* Griffin & Beuhler, *supra* note 175, at 75 (stating that “in making real life predictions and judgments, the problem is . . . in choosing what information to use and how to weigh and combine the information”).

185 *See, e.g.*, Sarah Lichtenstein & Baruch Fischhoff, *Do Those Who Know More Also Know More About How Much They Know?*, 20 *ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE* 159, 180–81 (1977); Stuart Oskamp, *Overconfidence in Case-Study Judgments*, 29 *J. CONSULTING PSYCHOL.* 261, 264–65 (1965).

as psychology or medicine are required to make difficult diagnoses or predictions, they are better off in the long run relying on statistical information about a few key variables than on their own all-things-considered clinical judgments about the particular case.¹⁸⁶ One study, for instance, showed that the application of a statistically derived personality test better distinguished between patients diagnosed as neurotic and psychotic, respectively, than did twenty-nine subjects, including many Ph.Ds in psychology, who were provided with all the same relevant data.¹⁸⁷ Such results appear to stem from both experts' tendency to pay insufficient attention to base rates when making predictive or diagnostic judgments, as well as their inability to accord proper weight to the variables they consider.¹⁸⁸ In fact, some researchers have suggested that experts are usually quite adept at identifying which variables are *relevant* to making accurate predictions or diagnoses; the problem is that they tend to misallocate the relevant *weights* of such variables.¹⁸⁹ And they do so to such an extent that they would be even better off combining them in quite crude ways, such as by assigning each variable the same weight.¹⁹⁰

From this body of empirical research, three generalizations emerge. First, people tend not to appreciate sufficiently the significance of an indicator's relative weight or reliability when reconciling conflicting evidence. Second, where the weight of a piece of evidence is relatively low, people place undue confidence in its probative value, and this is true even when they recognize that the weight is low. Finally, accuracy in making predictive or diagnostic judgments may often be improved by applying fixed weights to statistically relevant evidence, rather than by attempting to determine individually the appropriate weight in each case.

186 The classic work in this area is PAUL E. MEEHL, *CLINICAL VERSUS STATISTICAL PREDICTION* (1954).

187 See Robyn M. Dawes et al., *Clinical Versus Actuarial Judgment*, 243 *SCIENCE* 1668, 1669 (1989).

188 See *id.* at 1672; Paul E. Meehl, *Causes and Effects of My Disturbing Little Book*, 50 *J. PERSONALITY ASSESSMENT* 370, 372 (1986).

189 See Dawes et al., *supra* note 187, at 1672; see also Robyn M. Dawes, *The Robust Beauty of Improper Linear Models in Decision Making*, 34 *AM. PSYCHOLOGIST* 571 (1979), reprinted in *JUDGMENT UNDER UNCERTAINTY*, *supra* note 175, at 391, 395 ("People are good at picking out the right predictor variables and at coding them in such a way that they have a conditionally monotone relationship with the criterion. People are bad at integrating information from diverse and incomparable sources.").

190 See Dawes et al., *supra* note 187, at 1672.

b. Application to Evidence Law

Whether such generalizations may validly be applied to judges and juries is a difficult and controversial question. Recently, scholars have criticized not only the normative and methodological premises on which the heuristics and biases research program is based, but also the assumption that its findings apply uniformly across social and institutional contexts and thus unproblematically justify proposals for legal reform.¹⁹¹ Such caution is well taken. Still, even if these biases are not as strong or as pervasive as the research suggests, they at least suggest that a type of rule that attempts to suppress the probative value of testimony might be a useful complement to our factfinding methods. As noted above, such rules vary both in the extent to which they bind finders of fact and in their scope of application.¹⁹² How broad and how binding any particular rule should be depends on the context, but there is reason to think that using some form of such a rule could increase the rate of accurate verdicts in the long run.

They would do so in two ways. First, rules of weight might improve accuracy by decreasing the weight ascribed to already admissible evidence. The clearest and most obvious example here is eyewitness identifications. It is now widely understood by courts and scholars that eyewitness identifications based on police lineups are far less reliable than our commonsense intuitions might otherwise consider them.¹⁹³ Whether allowing psychological experts to testify on

191 See generally Gerd Gigerenzer & Peter M. Todd, *Fast and Frugal Heuristics: The Adaptive Toolbox*, in SIMPLE HEURISTICS THAT MAKE US SMART 3, 5 (Gerd Gigerenzer et al. eds., 1999) (applying the concept of first and frugal heuristics to decisionmaking limited by time and knowledge); Jonathan J. Koehler, *The Base Rate Fallacy Reconsidered: Normative, Descriptive and Methodological Challenges*, 19 BEHAV. & BRAIN SCI. J. 1, 1 (1996) (criticizing the current approach for “its failure to consider how the ambiguous, unreliable, and unstable base rules of the real world are and should be used”); Mitchell, *supra* note 15, at 143–44 (arguing that the heuristics and biases research is not as broadly applicable to factfinding as some have claimed).

192 See *supra* Part II.B.

193 The problems with eyewitness testimony have long been recognized. Nearly one hundred years ago, the Harvard psychologist Hugo Munsterberg conducted a number of studies on human perceptual abilities and published a book highly critical of the law’s apparent complacent attitude to the defects of eyewitness testimony. See HUGO MUNSTERBERG, ON THE WITNESS STAND 10–11 (1908). The classic modern account is ELIZABETH F. LOFTUS, EYEWITNESS TESTIMONY (1979). “The problem can be stated rather simply: on the one hand, eyewitness testimony is very believable and can wield considerable influence over the decisions reached by a jury; on the other hand, eyewitness testimony is not always reliable.” *Id.* at 6–7. For a useful collection of some of the most important psychological research on the topic since then, see 2 THE HANDBOOK OF EYEWITNESS PSYCHOLOGY (R.C.L. Lindsay et al. eds., 2007) and ELIZABETH F. LOFTUS ET AL., EYEWITNESS TESTIMONY: CIVIL AND CRIMINAL (4th ed. 2007).

the unreliability of eyewitness identifications is an effective and appropriate cure for this problem is still a subject of some debate.¹⁹⁴ Regardless, another approach would be to admit such identifications, or a particular subclass of them that have proven to be particularly unreliable, as a lower grade of evidence or one that cannot sustain conviction without further corroboration.

Second, if we used rules of weight, we could allow more relevant evidence in. Consider hearsay testimony. Most judges and scholars recognize that hearsay evidence is relevant and thus has some probative value; however, it is excluded on the ground that juries might accord it more weight than it deserves.¹⁹⁵ If, for instance, all hearsay evidence were admitted as a *low grade* of evidence—or one that is less reliable than any direct evidence that directly contradicted it—then it may be that the true probative value of hearsay evidence would be better accounted for in the long run than it is in the current system of exclusion and exceptions.¹⁹⁶ A jury could be instructed accordingly: “X’s testimony about Y is hearsay. The law considers such testimony to be generally of questionable reliability. As such, you ought not accord that statement much weight in your deliberations.”

Of course, one might immediately register skepticism as to the capacity of juries to apply such rules. Much research conducted over the last thirty years, for instance, suggests that jury instructions on the substantive law, the burden of proof, and the rules of evidence are rarely effective at guiding jury decisionmaking.¹⁹⁷ There may be little reason to think that jurors would be much better at applying rules of weight. Indeed, the research on capital punishment instructions, which ask juries to “weigh” mitigating factors against aggravating factors, offers a bleak assessment of the jury’s capacity to follow what

194 See, e.g., Steven D. Penrod & Brian Cutler, *Preventing Mistaken Convictions in Eyewitness Identification Trials: The Case Against Traditional Safeguards*, in *PSYCHOLOGY AND LAW* 89, 112–15 (Ronald Ruesch et al. eds., 1999) (endorsing the use of such experts in trials).

195 See Kenneth Culp Davis, *Hearsay in Administrative Hearings*, 32 *GEO. WASH. L. REV.* 689, 689–90, 695–96 (1964) (noting that just like nonhearsay evidence, “the reliability of hearsay ranges from the least to the most reliable” and that “[a]ll evidence scholars, without exception, express dissatisfaction with the hearsay rule”).

196 Of course, admission of hearsay testimony in the criminal context might violate the Confrontation Clause. See *Crawford v. Washington*, 541 U.S. 36, 68 (2004) (holding that prior opportunity to cross-examine a witness is required by the Sixth Amendment’s Confrontation Clause).

197 For a useful overview of this literature, see Joel D. Lieberman & Bruce D. Sales, *What Social Science Teaches Us About the Jury Instruction Process*, 3 *PSYCHOL. PUB. POL’Y & L.* 589 (1997).

appear to be a quite similar type of instruction to that required for rules of weight.¹⁹⁸

But the research on jury instructions does not prove that juries are incapable of applying rules of weight, because that research also suggests that the problem with instructions is not that the concepts employed are too conceptually complex for the average juror to handle, but rather that the language used is too legalistic and arcane.¹⁹⁹ One study on death penalty instructions, for instance, found that comprehension of instructions was extremely low not because jurors did not know how to “weigh” factors but because they did not understand what the terms “mitigating” and “aggravating” meant in the first place.²⁰⁰ And this finding is consistent with the experience of many courts in which juries have asked for explanations as to the meaning of such terms.²⁰¹ Perhaps not surprisingly, researchers have found that they can improve comprehension of all sorts of jury instructions by replacing legal jargon with simpler language and more straightforward syntax.²⁰² It thus seems reasonable to infer that so long as instructions for rules of weight are written in straightforward, plain language such as that suggested above, jurors will be able to understand them and therefore apply them.

Now one might question whether this last inference is always warranted. Some tasks asked of jurors may simply be too psychologically difficult for them to follow, even if the instructions are clearly written

198 See, e.g., Theodore Eisenberg & Martin T. Wells, *Deadly Confusion: Juror Instructions in Capital Cases*, 79 CORNELL L. REV. 1, 9–12 (1993) (suggesting that jurors have difficulty understanding the burden of proof and weighing evidence during the penalty phase of capital trials); Craig Haney & Mona Lynch, *Comprehending Life and Death Matters: A Preliminary Study of California’s Capital Penalty Instructions*, 18 LAW & HUM. BEHAV. 411, 425 (1994) (same).

199 See Nancy S. Marder, *Bringing Jury Instructions into the Twenty-First Century*, 81 NOTRE DAME L. REV. 449, 454 (2006) (reviewing three decades of empirical literature on jury instructions and concluding that “[r]esearchers have found that jurors are confused because the instructions use legal jargon or ambiguous language, awkward grammatical constructions, and an organization that is difficult to discern”).

200 Haney & Lynch, *supra* note 198, at 420–21.

201 See Peter Meijes Tiersma, *Dictionaries and Death: Do Capital Jurors Understand Mitigation?*, 1995 UTAH L. REV. 1, 11–23 (discussing several cases where juries have asked courts for clarifying definitions of “mitigating” or “aggravating” circumstances).

202 See Robert P. Charrow & Veda R. Charrow, *Making Legal Language Understandable: A Psycholinguistic Study of Jury Instructions*, 79 COLUM. L. REV. 1306, 1331 (1979) (finding that juror comprehension improved thirty-five to forty percent after rewriting instructions in more simple and direct language); see also AMIRAM ELWORK ET AL., MAKING JURY INSTRUCTIONS UNDERSTANDABLE 145–81 (1982) (giving guidelines for rewriting jury instructions).

and the jurors well intentioned.²⁰³ Such may be the case with so-called “limiting instructions,” which tell the jury to consider certain pieces of evidence for one purpose but not another.²⁰⁴ For some research suggests that judges are not much better at ignoring such evidence than are jurors.²⁰⁵ Perhaps, then, asking jurors to apply rules of weight, like asking them to ignore evidence for certain purposes, violates the principle of “ought implies can,” according to which “one can be obliged to do A only if one has an effective choice as to whether to do A.”²⁰⁶

203 See Shari Seidman Diamond, *Beyond Fantasy and Nightmare: A Portrait of the Jury*, 54 *BUFF. L. REV.* 717, 751 (2006) (noting that certain limiting instructions, such as those allowing that a defendant’s criminal record be considered for impeachment purposes but not as evidence of bad character, “may ask the jurors to engage in mental gymnastics that are not easy to perform” and that “jurors may be unwilling or find it impossible to perform the required cognitive adjustments”).

204 There has been much research on the effectiveness (or lack thereof) of limiting instructions. See, e.g., Joel D. Lieberman & Jamie Arndt, *Understanding the Limits of Limiting Instructions: Social Psychological Explanations for the Failures of Instructions to Disregard Pretrial Publicity and Other Inadmissible Evidence*, 6 *PSYCHOL. PUB. POL’Y & L.* 677, 686 (2000) (noting several studies that show that jurors are influenced by evidence of a defendant’s prior conviction even when told to disregard it for purposes other than credibility); Roselle L. Wissler & Michael J. Saks, *On the Inefficacy of Limiting Instructions: When Jurors Use Prior Conviction Evidence to Decide on Guilt*, 9 *LAW & HUM. BEHAV.* 37, 47 (1985) (concluding that “the presentation of the defendant’s criminal record does not affect the defendant’s credibility, but does increase the likelihood of conviction, and that the judge’s limiting instructions do not appear to correct that error”). In fact, some studies have suggested that such instructions produce a “backfire effect,” whereby the instruction calls attention to the relevant piece of evidence and makes the jury more likely to consider it. See Lieberman & Arndt, *supra*, at 689 (noting the backfire effect and citing a 1959 study showing that “participants awarded higher damages to a plaintiff after being instructed by a judge to disregard a statement that a defendant was covered by insurance than when given no instructions to that effect” (citing Dale W. Broeder, *The University of Chicago Jury Project*, 38 *NEB. L. REV.* 744, 744–60 (1959))). But see Mitchell, *supra* note 15, at 1121 (noting that recent studies suggest that jury deliberations may increase the capacity of jurors to ignore inadmissible evidence).

205 See Wistrich et al., *supra* note 15, at 1259 (concluding that their study showed that “some types of highly relevant, but inadmissible, evidence influenced the judges’ decisions”).

206 William P. Alston, *The Deontological Conception of Epistemic Justification*, in 2 *PHILOSOPHICAL PERSPECTIVES: EPISTEMOLOGY* 257, 259 (1988). The application of this principle to the law of evidence is hardly new. Two centuries ago, James Wilson explained why rules of admissibility are preferable to rules of weight on precisely this ground:

The law will not order that which is unnecessary: it will not attempt that which is impracticable. In no case, therefore, does it order a witness to be believed; for jurors are triers of the credibility of witnesses, as well as of the truth of facts In no case, likewise, does the law order a witness not to be believed; for belief might be the unavoidable result of his testimony.

The question is ultimately an empirical one, but it seems plausible to think that applying a defeasible rule of weight does not present nearly as formidable a psychological challenge as does applying a conventional limiting instruction. This is so because, as an intuitive matter, it seems far more difficult to ignore a factor completely than merely to accord it reduced significance in making a decision. To see why this might be, consider the following scenario: you have just been offered a great job at a start-up company, e-widgets.com, that makes e-widgets. You are excited about the opportunity, because your offer includes a lot of stock options at an attractive (but hopefully not backdated) strike price, and there is much industry buzz about the company. Still, you are undecided because you recognize that there is some risk that the company will not survive. So you discuss the matter over lunch with your friend Alfred (whom you know) and Alfred's new friend Bob (whom you've never met). Bob advises you against taking the job because he has heard horrible things about the e-widgets.com management team—that they are incompetent, deceitful, and egomaniacal. As Bob then recounts various stories of perfidy and incompetence, you begin to think you better stay with your current job. But the day after your lunch, Alfred calls you and tells you that Bob works for a direct competitor of e-widgets.com and that although he does not know Bob himself well, he would take what *anyone* at his company says about e-widgets.com with a grain of salt.

Now consider two questions: (1) Are you able to do as Alfred suggests and discount the information Bob gave you accordingly? (2) If Alfred had instead told you to ignore what Bob said completely for the purposes of your job offer but to consider it for the purpose of, say, making an investment in the company, would you be able to do that? If the answer to (1) is yes, then you can successfully apply a rule of weight. And if the answer to (2) is no, then applying rules of weight seems to present a quite different psychological task than following limiting instructions.

2. Fairness and Risk Allocation

That the law should treat people equally is a bedrock legal principle, and the capacity of rules to ensure such equal treatment is considered to be one of their chief virtues.²⁰⁷ Rules deprive the

James Wilson, *Of the Nature and Philosophy of Evidence*, in 1 *THE WORKS OF JAMES WILSON* 369, 383 (Robert Green McCloskey ed., 1967).

²⁰⁷ See, e.g., Antonin Scalia, *The Rule of Law as a Law of Rules*, 56 *U. CHI. L. REV.* 1175, 1178 (1989) ("The Equal Protection Clause epitomizes justice more than any other provision of the Constitution. And the trouble with the discretion-conferring

decisionmaker of a type of decisional jurisdiction and, in so doing, make decisions more consistent and predictable.²⁰⁸ This feature of rules remains a virtue in procedural and evidentiary matters.²⁰⁹ Indeed, Professor Alex Stein has recently argued that allocating the risk of factfinding error is one of the central goals of evidence law.²¹⁰ According to Stein, evidence rules not only serve to ensure that civil parties bear the risk of error equally but also apportion risk according to substantive moral and political principles.²¹¹ The most obvious example, of course, is the heightened burden of proof in criminal trials, but Stein also argues that rules excluding hearsay and character evidence can be justified on comparable grounds.²¹²

Rules of weight, however, are even more effective than rules of admissibility at ensuring consistency in factfinding. By specifying in advance the amount of weight certain types of testimony or evidence deserve, rules of weight prevent judges or juries from willfully or mistakenly disregarding or overvaluing evidence. Indeed, this has long been considered a chief virtue of such rules. In his 1838 treatise on circumstantial evidence, William Wills explained that the Roman law system of “half-proofs” served as an “important limitation[] upon the tyranny and inconstancy of judicial discretion,” particularly given that in those times “laws were administered by a single judge, without the salutary restraints of publicity and popular observation.”²¹³ For the same reason, scholars have criticized Bentham’s “natural system” of free proof precisely on the ground that it entails placing undue faith in the fairness and competence of officials.²¹⁴

Rules of weight serve the interests of fairness in another, related way as well. Because substantive rights are always at issue in any adjudication and because no factfinding procedure can be guaranteed to

approach to judicial law making is that it does not satisfy this sense of justice very well.”).

208 See SCHAUER, *supra* note 133, at 158.

209 See STEIN, *supra* note 9, at 183–208 (arguing that the use of exclusionary rules is necessary to allocate the risk of factfinding error between the parties in legitimate ways); L. Jonathan Cohen, *Freedom of Proof*, in *FACTS IN LAW* 1, 2–5 (William Twining ed., 1983) (presenting, though ultimately rejecting, a rule-of-law argument in favor of regulating proof).

210 See STEIN, *supra* note 9, at 1.

211 See *id.* at 172–78, 214–19.

212 See *id.* at 183–98. Stein’s justification for excluding such evidence requires more explanation than can be provided here.

213 WILLS, *supra* note 23, at 34–35; see also DAMAŠKA, *supra* note 9, at 20 (noting that most of the rules of weight and sufficiency used in criminal trials under the Roman system served to protect criminal defendants from unwarranted conviction).

214 See TWINING, *supra* note 38, at 70.

be free from error, any method of factfinding has the effect of allocating the risk of such error between the parties.²¹⁵ And proper allocation of such risk depends on moral or political criteria, not epistemic ones.²¹⁶ A system of factfinding, for instance, might allocate risk in such a way as to minimize overall error costs without regard for who pays such costs (i.e., bears the risk of error). Or it could allocate risk according to moral criteria, such as a reluctance to expose innocent individuals to the risk of unjust criminal punishment.²¹⁷

Rules of weight can serve well this risk allocation function. The best example of existent rules are corroboration rules (of the absolute sort), which preclude a verdict against a criminal defendant based on certain types of uncorroborated testimony.²¹⁸ Of course, such rules are partly motivated by accuracy concerns since, if there were no particular doubts about the reliability of the covered testimony, there would be no need for a rule. But they also rest on a substantive moral judgment that criminal defendants should not be forced to bear the risk that the testimony may turn out to be false.²¹⁹ And this is not only true of absolute corroboration rules, which directly result in judgments as a matter of law. Defeasible rules of weight do not foreclose the possibility of a verdict against the defendant based on certain evidence, but they may lower the likelihood that a verdict will be based on such evidence by calling the factfinder's attention to its weakness. In other words, like corroboration rules, they are capable of allocating the risk of factfinding errors in such a way as to reflect principles of fairness.

One might object, however, that such rules of weight are not meaningfully enforceable. Given that even simple cases may involve a mass of conflicting and corroborating evidence, how can one possibly tell whether a judge or jury has accorded "low weight" to a given piece of evidence pursuant to a rule or instruction? The answer to this objection differs as between jury trials, on the one hand, and bench trials or nontrial dispositions on the other. As used in jury trials, it is true that whether the jury applied a rule of weight will sometimes be difficult to tell, but two points deflate the force of this objection. First, in those cases in which the rule matters the most it will probably be the most clear. That is, if there is a great deal of evidence adduced in support of a particular fact, whether or not the jury properly dis-

215 See STEIN, *supra* note 9, at 12–13.

216 See *id.*

217 See *id.*

218 See *supra* Part II.B.1.

219 See STEIN, *supra* note 9, at 208.

counted one piece of evidence will be difficult to determine but not terribly significant. On the other hand, if a factual conclusion rests virtually entirely on a piece of evidence deemed unreliable or suspect, a possible violation will likely be both more important and more apparent. Furthermore, any evaluation of rules of weight must be made relative to the status quo. Under current practices, the jury's reasoning processes are already treated as a "black box," so it is virtually impossible to tell whether it even understood the applicable law correctly.

One might respond that regardless of what the jurors actually do, given the frequency with which objections are already made to a court's jury instructions on the law, attempting to instruct them on questions of fact will open up a procedural can of worms. Wigmore made this argument, suggesting that rules of weight would introduce "ten thousand more quibbles" into the system.²²⁰ But while incorporating dozens of reliability rules may become unworkably complex, there is no reason to think that the use of a few, simple ones would. Indeed, this point is really about the *complexity* of jury instructions, which likely applies with equal or greater force to instructions on law.²²¹

As applied to nonjury dispositions, rules of weight may actually make it *easier* to enforce fair and accurate judicial findings of fact. Insofar as rules of weight serve as sufficiency standards that determine whether a case may go to trial or whether a verdict is supported by sufficient evidence, they offer appellate courts substantive standards by which to determine whether a court properly interpreted the evidence before it. Rules of admissibility, on the other hand, offer no comparable guidance.²²² One result may be that appellate courts

220 Wigmore, *supra* note 7, at 478. This argument may find some empirical support in the findings of the Law Commission of the English High Chancellor, which recommended to Parliament in 1991 that it abolish a corroboration rule that applied to the testimony of accomplices. Its report noted that "[t]he *complexity* of the corroboration rules is notorious, as are the very great difficulties that that complexity causes both for the judges who have to expound the rules and for the juries who have to try to understand and apply them." CORROBORATION OF EVIDENCE, *supra* note 27, at 4.

221 See, e.g., JEROME FRANK, LAW AND THE MODERN MIND 184 (1936) ("These instructions are like exorcising phrases intended to drive out evil spirits [T]he more unintelligible and technical instructions on the law may be considered as part of this mechanism of exorcism, resembling the 'tremendous words' from Hebrew and Greek . . . which the medieval exorcists employed to scare away the minions of Satan.").

222 See Mnookin, *supra* note 9, at 138 (noting that appellate review of a judge's application of admissibility rules in bench trials is nearly impossible because "we almost entirely lack formal rules about weight or probative value").

would accord trial courts less deference on questions of fact than they currently do. But while doing so may raise its own concerns, a lack of enforceability would not likely be one of them.

3. Efficiency and Agency Adjudication

That rules reduce decision costs is also one of the classic justifications for a regime of rule-based decisionmaking.²²³ Professor Kaplow usefully distinguishes among several different types of costs associated with any legal regime that will vary according to whether the legal directives take the form of rules or standards.²²⁴ The costs of promulgating a rule will be higher than those of promulgating a standard because more information is required to specify the content of the rule.²²⁵ But it costs less for individuals to comply with the law, and for judges to enforce it, in rule-like regimes than it does in regimes governed by standards, all other things being equal.²²⁶ Furthermore, the more frequent the conduct that the law governs is, the greater the efficiency gains of a regime of rules will be.²²⁷

One need not engage in an in-depth economic analysis of evidence law to see how rules may serve efficiency goals in factfinding.²²⁸ Adjudicative factfinding is expensive for society, but inaccurate verdicts in civil or criminal litigation result in real individual and social costs. A plausible goal of an evidentiary regime is thus to minimize

223 See SCHAUER, *supra* note 133, at 146 (“[R]ules allocate the limited decisional resources of individual decisionmakers, focusing their concentration on the presence or absence of some facts and allowing them to ‘relax’ with respect to others.”); Kathleen M. Sullivan, *Foreword: The Justices of Rules and Standards*, 106 HARV. L. REV. 22, 73 (1992) (“[A]dherence to precedent increases judicial efficiency by eliminating the duplicative work and the risk of error from incompetence or bias that would result from starting each case anew from first principles.”); Cass R. Sunstein, *Problems with Rules*, 83 CAL. L. REV. 953, 972 (1995) (“Rules make it unnecessary for each of us to examine fundamental issues in every instance; in this way rules create a convergence on particular outcomes by people who disagree on basic matters. Rules can, in short, be the most efficient way to proceed, by saving time and effort, and by reducing the risk of error in particular cases.”).

224 See Kaplow, *supra* note 135, at 577.

225 See *id.*

226 See *id.* at 562–64, 577.

227 See *id.* at 586–96, 621 (“If conduct will be frequent, the additional costs of designing rules—which are borne once—are likely to be exceeded by the savings realized each time the rule is applied.”).

228 For such a treatment arguing that American evidence law is more efficient and cost effective than other evidentiary systems, see generally Richard A. Posner, *An Economic Approach to the Law of Evidence*, 51 STAN. L. REV. 1477 (1999). For further discussion of this topic examining a cost efficiency approach to evidence, see STEIN, *supra* note 9, at 141–72. In this section, I draw primarily from these sources.

the sum of these two costs, which requires searching for the truth about disputed factual issues up to the point where the cost of securing a piece of evidence is equal to the reduction in error costs that the production of such evidence entails.²²⁹ Evidence rules may serve this goal either directly or instrumentally. So, for instance, the hearsay rule may represent a straightforward judgment that the reduction in error costs that hearsay evidence may yield does not, on balance, outweigh the costs of evaluating such evidence in every case (and that the opposite is true of evidence falling under hearsay exceptions).²³⁰ Some rules may also reduce error costs indirectly by forcing parties to produce the best evidence within their control.²³¹ Many burden-shifting legal presumptions serve this function, as did the original documents rule, which did not admit copies of documents unless the original was unavailable.²³²

Rules of weight also enable factfinding to be more efficient. Even though, as stated, some such rules are more accurately characterized as guidelines or standards, they still offer a more rule-like approach than the current state of unconstrained discretion. Like admissibility rules and other evidentiary devices, they could improve efficiency in two ways, either directly or instrumentally. Most straightforwardly, they often facilitate resolution of cases without the expense of a jury. But even within a jury trial, rules of weight could reduce overall costs. A rule that admits hearsay as a lower grade of evidence, for instance, reduces error costs if such a rule more accurately describes the amount of probative value the factfinder ought to attach to such evidence. And while it would not reduce decision costs as much as an exclusionary rule would, since the evidence would still need to be considered by the factfinder, it would reduce them more than no rule at all because, given its second-class status, the opposing party would devote fewer resources to disputing it and the factfinder would likely spend less time considering its relative reliability.

Rules of weight are also instrumentally effective at reducing errors and are thus consistent with the “best evidence principle,” according to which the function of rules of evidence is to ensure that parties produce the best evidence in their control.²³³ Although

229 See Posner, *supra* note 228, at 1481.

230 See *id.* at 1530.

231 See STEIN, *supra* note 9, at 154; see also Nance, *supra* note 30, at 230–31 (providing the justification for a best evidence rule).

232 The original documents rule has been replaced in federal courts by FED. R. EVID. 1002–1004. For the use of presumptions as burden-shifting devices, see generally Allen, *supra* note 108, at 332–38.

233 See Nance, *supra* note 30, at 271.

excluding evidence is clearly a more severe sanction than simply assigning it low weight, the latter would still likely have some effect and, once again, any increase in decision costs may be outweighed by the reduction in error costs.

Of course, trying to calculate in advance how the social costs of any procedural regime will ultimately net out is tricky business, not least because one needs to account not only for the costs of the trial and to the parties, but also for the incentive effects on future actors. Professor Stein, for instance, has suggested that admitting hearsay evidence might trigger ancillary litigation, which would produce additional costs.²³⁴ If that is so, a rule of weight that merely depresses the probative value of such evidence may not be sufficient to forestall such costly behavior. But even if it is not possible to evaluate the relative efficiency of a system of rules of weights in the abstract, it certainly is possible to draw some conclusions about the procedural contexts in which they would be most effective. Recall that the efficiency gains will be greatest in a rule-based regime where the costs of promulgating the rule are low and the frequency of its application is likely to be high.²³⁵ Both of these factors suggest that rules of weight will be particularly appropriate when used by specialized tribunals that can leverage their expertise to develop and implement such rules to deal with frequently recurring factual circumstances.

No surprise, then, that currently rules of weight are most commonly used by administrative agencies. As discussed in Part I, many agencies have employed rules of weight that deem certain types of testimony to be unreliable or less reliable than other sources of testimony. The United States International Trade Commission, to cite just one example, has held that the testimony of the employees of an owner of an alleged trademark is not accorded much weight.²³⁶ Not only do agencies derive clear efficiency gains given the relatively low cost of promulgating such rules and the high frequency of their application, but the other benefits of rules of weight also apply with particularly strong force. The accuracy of the rules is likely heightened by an agency's relative expertise over matters within its jurisdiction. Equally important, because Administrative Law Judges (ALJs) do not enjoy the salary protections that federal judges do, their discretion may be that much more in need of checking to ensure that parties

234 See STEIN, *supra* note 9, at 156.

235 See Kaplow, *supra* note 135, at 564, 577.

236 See, e.g., *Certain Braiding Machs.*, USITC Pub. 1435, Inv. No. 337-TA-130 (Oct. 1983) ("Generally speaking, little weight is accorded to the testimony of an employee of the owner of the alleged trademark because of the potential risk that the perceptions conveyed are colored by bias."). For other examples, see *supra* Part I.D.2.

receive fair and equal adjudication of their claims.²³⁷ The wildly disparate treatment of asylum applicants by immigration judges that has recently received national attention is just one particularly egregious example.²³⁸ Indeed, if bureaucrats display systematic prejudices against certain types of claims, better that such prejudgments be formulated as rules so that they can be reviewed by agencies and appellate courts.²³⁹

Indeed, for all these reasons, the use of rules of weight by agencies may teach a lesson applicable to judicial proceedings, for they demonstrate how such rules may preserve a system of relatively free proof by offering a less extreme form of regulating factfinding than through the adoption of exclusionary rules. Because agencies typically do not make extensive use of rules of admissibility, agency adjudication has often been held up as a model of a freer system of proof.²⁴⁰ But it may be that rules of weight are necessary to maintain such a system without granting judges unfettered discretion over factfinding.

At this point one might object that the administrative experience is the exception that proves the rule. Since administrative agencies have never been held to the same procedural standards as have civil and criminal courts of law, a defense of rules of weight in that context hardly advances the cause. However, the Supreme Court has suggested that, even when used by agencies, rules of weight may be not just ineffective, but unconstitutional.

237 Compare 5 U.S.C. § 5372 (2000) (detailing the pay scale of administrative law judges statutorily determined by the executive office of personnel management), with U.S. CONST. art. III, § 1 (“The [Federal] Judges . . . shall, at stated times, receive for their Services a Compensation which shall not be diminished during their Continuance in Office.”).

238 See Julia Preston, *Big Disparities Found in Judging of Asylum Cases*, N.Y. TIMES, May 31, 2007, at A1. For the massive, several-year study on which this Article is based, documenting the enormous disparities in the rates with which immigration judges grant asylum to applicants, see Jaya Ramji-Nogales et al., *Refugee Roulette: Disparities in Asylum Adjudication*, 60 STAN. L. REV. 295, 310–49 (2007).

239 Criticism of immigration judges has been severe recently in this regard. See, e.g., *Benslimane v. Gonzales*, 430 F.3d 828, 830 (7th Cir. 2005) (collecting cases in which circuit courts have criticized immigration judges and the Board of Immigration Appeals in asylum cases and noting that “adjudication of these cases at the administrative level has fallen below the minimum standards of legal justice”).

240 See, e.g., *Davis*, *supra* note 13, at 586–91 (outlining the contours and advantages of evidentiary procedures in administrative courts).

III. A DEEPER CHALLENGE: *ALLENTOWN MACK v. NLRB*

In *Allentown Mack*, the Supreme Court condemned the use of rules of weight as inconsistent with fundamental notions of due process.²⁴¹ The Court's opinion is not entirely clear and lends itself to a couple of different interpretations, suggesting both *epistemic* objections to rules of weight, as well as *moral* ones. But neither sort of objection is sufficient to justify the Court's blanket denunciation of rules of weight.

At issue in *Allentown Mack* were two closely related rules that the National Labor Relations Board used to rely on when adjudicating collective bargaining disputes. Under federal law, companies have a duty to bargain collectively with the union representing their employees.²⁴² Whenever a company's ownership changes, the law presumes that the union continues to enjoy majority support of the employees, but the company can overcome that presumption if it has a "good faith reasonable doubt" about such majority support.²⁴³ In evaluating the evidence as to whether an employer does in fact have such good faith reasonable doubt, the NLRB used to cite cases in support of the propositions that (1) one employee's testimony conveying the views of other employees ought not be accorded much weight;²⁴⁴ and (2) that *any* statement related to union support made by an employee in the context of an interview with the new employer was not particularly reliable.²⁴⁵ For years, federal courts reviewing such decisions

241 *Allentown Mack Sales & Serv., Inc. v. NLRB*, 522 U.S. 359, 368–71 (1998).

242 29 U.S.C. § 158(a)(5) (2000).

243 *E.g.*, *Pioneer Inn Assocs. v. NLRB*, 578 F.2d 835, 839 (1978).

244 *See Deutsch*, 293 N.L.R.B. 1000, 1001 n.11 (1989) ("The Board has stated that 'testimony concerning conversation directly with the employees involved . . . is much more reliable than testimony concerning merely a few employees ostensibly conveying the sentiments of their fellows.'" (alteration in original) (quoting *Sofco, Inc.*, 268 N.L.R.B. 159, 160 n.10 (1983))); *La.-Pac. Corp.*, 283 N.L.R.B. 1079, 1080 n.6 (1987) ("With respect to the statement of employee Harker purporting to represent the views of other employees, we find that little weight can be accorded to such a statement."); *Sofco*, 268 N.L.R.B. at 160 n.10 ("[T]estimony concerning conversations directly with the employees involved, as here, is much more reliable than testimony concerning merely a few employees ostensibly conveying the sentiments of their fellows.").

245 *See Middleboro Fire Apparatus, Inc.*, 234 N.L.R.B. 888, 894 (1978) ("Statements made by employees during the course of an interview with a prospective employer that they approve of his unqualified position that he has no contract and is not obligated to bargain with the Union claiming to represent them are not voluntary, uncoerced expressions of employee sentiment upon which their employer can rely in asserting a good-faith doubt of an incumbent union's majority status.").

endorsed such rules or deferred to the NLRB's judgment as to their utility.²⁴⁶

In *Allentown Mack*, however, the Supreme Court stripped these rules of any independent legal force and, possibly, found them unconstitutional.²⁴⁷ The Court vacated the NLRB's judgment in which the Board had found that Allentown Mack did not have a reasonable good faith doubt that the union representing its employees did not enjoy majority support.²⁴⁸ Writing for the Court, Justice Scalia noted that several pieces of evidence supported Allentown Mack's good faith doubt about the union's support, but he found particularly significant one employee's statement that "'the entire night shift did not want the Union'" and another's that "'it was his feeling that the employees did not want a Union.'"²⁴⁹ The Court thus concluded that the NLRB's decision was not supported by substantial evidence and so reversed the D.C. Circuit's judgment in its favor.²⁵⁰

Administrative law scholars tend to see *Allentown Mack* as significant insofar as it reflects the Court's more rigorous review of agency factual findings and thus its departure from previous understandings of "substantial evidence" review.²⁵¹ But the Court seemed as troubled by *how* the NLRB found the facts in the case as it did by the fact *that* the Board got them wrong. Specifically, the Court objected to the

246 See, e.g., *Bryan Mem'l Hosp. v. NLRB*, 814 F.2d 1259, 1262 (8th Cir. 1987) ("Unverified claims by employees that they speak for others is not a sufficient basis for an employer's reasonable good faith doubt about union support where there are no other reliable indicia of employee attitudes."); *NLRB v. Middleboro Fire Apparatus, Inc.*, 590 F.2d 4, 9 (1st Cir. 1978) (deferring to the NLRB's "seasoned feel for the meaning of events in a labor-management setting" including its judgment on such factual issues as whether "the factors of a few employees dealing with a prospective employer in a close relationship and informal surroundings were likely to mean more or less coercion"); *NLRB v. Cornell of Cal., Inc.*, 577 F.2d 513, 516-17 (9th Cir. 1978) (approving of the NLRB's rule according little weight to employee assertions about the views of other employees on the ground that it "avoids allowing a few employees to undermine their union merely by making an assertion that is not easily verified by the employer" and concluding that "an employer cannot satisfy its burden of proof that the doubts were reasonable by resting exclusively on such normally unreliable assertions" (citation omitted)).

247 *Allentown Mack Sales & Serv., Inc. v. NLRB*, 522 U.S. 359, 370 (1998).

248 See *id.* at 380.

249 *Id.* at 369-70 (quoting employee statements).

250 See *id.* at 380.

251 See, e.g., 2 RICHARD J. PIERCE, JR., *ADMINISTRATIVE LAW TREATISE* § 11.2, at 775 (4th ed. 2002) ("The majority opinion is extraordinary in terms of its degree of departure from the Court's traditional approach to the substantial evidence test."); M. Elizabeth Magill, *Agency Choice of Policymaking Form*, 71 U. CHI. L. REV. 1383, 1429 n.152 (2004) ("Some scholars point to the recent case of [*Allentown Mack*] as illustrative of a new era of intense examination of agency factfinding." (citation omitted)).

NLRB's reliance on the rules relating to testimonial weight mentioned above.²⁵² It noted, for instance, that it was error for the ALJ to discount one employee's statements on the basis of "the Board's historical treatment of unverified assertions by an employee about another employee's sentiments."²⁵³ Whether or not the Board treated such statements in this way, they "provide no justification for the Board's factual inferences here."²⁵⁴ According to the Court, such inferences depend exclusively on "logic and sound inference from all the circumstances, not an arbitrary rule of disregard to be extracted from prior Board decisions."²⁵⁵ While recognizing that the Board could adopt exclusionary rules of evidence or even "forthrightly and explicitly adopt counterfactual evidentiary presumptions (which are in effect substantive rules of law) as a way of furthering particular legal or policy goals," it considered such rules very different from the rules of weight used by the NLRB:

That is not the sort of Board action at issue here, however, but rather the Board's allegedly *systematic undervaluation of certain evidence* When the Board purports to be engaged in simple factfinding, unconstrained by substantive presumptions or evidentiary rules of exclusion, it is not free to prescribe what inferences from the evidence it will accept and reject, but must draw all those inferences that the evidence fairly demands. "Substantial evidence" review exists precisely to ensure that the Board achieves minimal compliance with this obligation, which is the *foundation of all honest and legitimate adjudication*.²⁵⁶

In other words, the Court seemed to be saying, fair adjudication of disputes requires drawing factual inferences entirely unencumbered

252 *Allentown Mack*, 522 U.S. at 378 (examining "the Board's allegedly systematic undervaluation of certain evidence").

253 *Id.* at 379 (quoting *Allentown Mack Sales & Serv., Inc.*, 316 N.L.R.B. 1199, 1208 (1995)).

254 *Id.*

255 *Id.* at 378.

256 *Id.* at 378-79 (emphases added). Justice Breyer, writing in dissent, found nothing wrong with the NLRB "drawing upon both reason and experience," to interpret the evidence and in effect saying that "it will 'view with suspicion and caution' one employee's statements 'purporting to represent the views of other employees.'" *Id.* at 393 (Breyer, J., dissenting) (quoting *Wallkill Valley Gen. Hosp.*, 288 N.L.R.B. 103, 109 (1988)). Breyer thus asked rhetorically, "How is it unreasonable for the Board to provide this kind of guidance, about what kinds of evidence are more likely, and what kinds are less likely, to support an 'objective reasonable doubt' . . . ?" *Id.* at 394. In his view, it was perfectly proper for the NLRB to "develop rules of thumb about the likely weight assigned to different kinds of evidence." *Id.* at 393.

by rules of law and thus precludes the use of rules of weight.²⁵⁷ But there are two different rationales that could justify the Court's position. Let us consider each in turn.

A. *Epistemic Rationale*

Some of the Court's language suggests that its objection is primarily an epistemic one; that is, a concern with verdict accuracy. It notes, for instance, that whatever validity the NLRB's rules may have as a general matter, "they . . . provide no justification for the Board's factual inferences here."²⁵⁸ Instead, the ALJ must use "logic and sound inference from all the circumstances" to reach a conclusion.²⁵⁹ This view finds support in a theory of factfinding that has been dubbed "evidentiary holism,"²⁶⁰ "relative plausibility theory,"²⁶¹ or "coherence" theory,²⁶² according to which factfinders do not and should not assign each piece of evidence a precise probative value reducible to a Bayesian conditional probability estimate.²⁶³ Instead, they ought to

257 It is possible to read the Court as demanding only the *publicity* of procedural rules—a traditional rule-of-law virtue. See Jack Goldsmith & Eric A. Posner, *Response: The New International Law Scholarship*, 34 GA. J. INT'L & COMP. L. 463, 480 (2006) (listing publicity as one of the "standard elements of the domestic rule of law" along with "generality, stability, impartiality, . . . [and] equality before the law"). On this view, the evil of the NLRB's policy was not that it prescribed a rule of weight but that it did so secretly or without sufficient notice. But that reading seems impossible to square with the NLRB's frequent and explicit statements of its policy that it would view with skepticism certain statements made by employees, which it would support with case citations. Such precedent surely suffices as notice under any plausible understanding of what the rule of law demands. See *Payne v. Tennessee*, 501 U.S. 808, 848 (1991) (Marshall, J., dissenting) (stressing that overruling the Court's precedent "ought to be a matter of great moment [because] fidelity to precedent is fundamental to 'a society governed by the rule of law'" (citation omitted) (quoting *Akron v. Akron Center for Reprod. Health, Inc.*, 462 U.S. 416, 420 (1983))).

258 *Allentown Mack*, 522 U.S. at 379.

259 *Id.*

260 Michael S. Pardo, Comment, *Juridical Proof, Evidence, and Pragmatic Meaning: Toward Evidentiary Holism*, 95 NW. U. L. REV. 399, 399 (2000).

261 Ronald J. Allen & Brian Leiter, *Naturalized Epistemology and the Law of Evidence*, 87 VA. L. REV. 1491, 1527–37 (2001).

262 Amalia Amaya, Reasoning About Facts in Law: Essays in Coherence, Evidence, and Proof 66–70 (May 2007) (unpublished S.J.D. dissertation, Harvard Law School) (on file with Harvard Law School Library).

263 See Pardo, *supra* note 260, at 401 ("A holistic theory of evidence posits that the meaning or value of any particular atom of evidence depends on the role it plays in relation to all other evidence available to an interpreter."). This theory has its philosophical analogue in the view known as coherentism. See, e.g., LAURENCE BONJOUR & ERNEST SOSA, *EPISTEMIC JUSTIFICATION* 42 (2003) (defining a coherence theory of justification as "a view according to which (1) there are no basic or foundational beliefs

assess it by reference to the overall plausibility of the evidence presented at trial.²⁶⁴ If evidentiary holism is correct, it may be argued, then rules of weight such as the NLRB's would disrupt artificially the otherwise natural process of holistically drawing inferences among and between the various other items of evidence at trial in which factfinders can and ought to engage. Inaccurate verdicts will result.

This objection is not persuasive against rules of weight. First, even if jurors (and perhaps judges) do, as an empirical matter, tend to reason holistically about evidence, there is some reason to doubt that doing so is good for accurate factfinding.²⁶⁵ More important, even if evidentiary holism is true as a normative matter, it does not undermine rules of weight because such rules would simply be incorporated into the holistic analysis.

An analogy to statutory interpretation makes this point clear. Even those who agree that the goal of interpreting statutes is to discern legislative intent may disagree about the most effective way to meet that goal. Some believe that judges ought to look to all the relevant evidence particular to the statute at hand as well as the structure of the statute as a whole, while others remain skeptical of judges' capacity to conduct such an inquiry and think the use of canons of construction will better capture legislative intent in the long run.²⁶⁶

and (2) at least the primary basis for empirical justification is the fact that such beliefs fit together and support each other in a variety of complicated ways, thus forming a coherent system of beliefs—or perhaps more than one such system”).

264 Supporters of the view that jurors ought to weigh evidence holistically have found empirical support for the view that they do in fact reason in this way in the work of Nancy Pennington and Reid Hastie. See, e.g., Allen & Leiter, *supra* note 261, at 1528 n.110 (citing Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 519 (1991)).

265 In one study, for instance, Professor Dan Simon found that when subjects were presented with a complicated and evenly balanced legal case involving a variety of independent legal and factual issues, if a key fact was adjusted so that one party's argument became much stronger than the other's, subjects' confidence in their own initial judgments increased not only with respect to the relevant legal issue, but also with respect to other legal issues on which the adjusted fact could have had no conceivable relevance. See Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511, 538 (2004) (noting that new evidence about a party's previous bad conduct appeared to affect subjects' confidence in their judgments as to whether an Internet website was analogous to a newspaper for the purposes of free speech doctrine).

266 See ADRIAN VERMEULE, JUDGING UNDER UNCERTAINTY 4 (2006) (“[J]udges should sharply limit their interpretive ambitions, in part by limiting themselves to a small set of interpretive sources and a restricted range of relatively wooden decision-rules.”). Compare Eskridge & Frickey, *supra* note 2, at 351–52 (endorsing a model of statutory interpretation captured by the metaphor of a “hermeneutical circle” according to which, “[a] part can only be understood in the context of the whole, and the

These two approaches are not in *psychological* tension with each other, because even if words can only be understood contextually or holistically,²⁶⁷ this fact does not preclude the use of descriptive canons, because such interpretive rules may themselves simply be included in the holistic reading of the statute.

Of course, there *is* a tension between these two interpretive approaches. But it is a substantive interpretive tension as to how Congress operates. The less Congress considers canons of construction or any general interpretive rules when drafting statutes and the more uniquely drafted each legislative act is, the less likely it is that applying such canons will yield accurate results and the better the holistic interpreter may be. But whether that is true depends on facts about Congress, not about how people think or reason. To use the example in Part II, Alfred's advice only presents a problem for the holistic reasoner to the extent that his generalization about the credibility of the people who work at Bob's company (and, by logical implication, about Bob's credibility) is mistaken.

Furthermore, a rule that suppresses the probative value of direct evidence may very well aid the factfinder in reaching the truth of the matter. As we have seen, some studies suggest that in drawing an "inference to the best explanation" of a set of evidence,²⁶⁸ people tend not to consider adequately the unreliability of a source as a possible explanation for apparent evidentiary conflicts and inconsistencies.²⁶⁹ Regardless, though, whether using rules to find facts increases accuracy depends on facts about the world, not whether people reason atomistically or holistically.

But perhaps that is just the point. Another version of the epistemic objection asserts that the world is simply too complicated, diverse

whole cannot be understood without analyzing its various parts," so that a judge interpreting a statute ought to "consider it in light of the whole enterprise, including the history, purpose, and current values" such that "none of the interpretive threads can be viewed in isolation, and each will be evaluated in its relation to the other threads"), with Caleb Nelson, *What Is Textualism?*, 91 VA. L. REV. 347, 349 (2005) (noting that someone who agrees that the goal of statutory interpretation is to divine legislative intent "could still be a textualist on the ground that judges are likely to make more accurate assessments of legislative intent if they use a relatively rule-like approach").

267 PAUL GRICE, *STUDIES IN THE WAY OF WORDS* 222 (1989).

268 See Amaya, *supra* note 262, at 136 ("Coherence may be construed in the course of legal decision-making through an 'inference to the best explanation.' That is, coherence-based legal inference is at bottom an explanatory kind of inference.").

269 Cf. Heller, *supra* note 15, at 244–45 (suggesting that the disproportionate weight that jurors assign direct evidence relative to circumstantial evidence may be due to the fact that jurors' unwillingness to convict depends on their ability to imagine a scenario in which the defendant could be not guilty).

and messy to ever allow us to fashion rules of weight adequately tailored for their purpose.²⁷⁰ The proper response to this objection is to acknowledge its force but note its limits. Certainly our ability to generalize accurately about the reliability of types of evidence is constrained by the limited reach of empirical methods. It may be difficult, for instance, to determine with precision how reliable the average person's memory of his or her childhood is, because it is not always possible to confirm or deny alleged memories. But, as we saw above, recent work in cognitive psychology suggests that even fairly crude generalizations may nonetheless serve as reasonably reliable predictors.²⁷¹ Thus, for instance, it may be that a rule of weight that ascribes relatively low weight to cross-racial eyewitness identifications or those based on certain types of police lineups, though over- and underinclusive as any rule would be, might nevertheless more accurately take account of the reliability of such evidence. In any case, all such arguments depend on the particular rule of weight at hand. It is not an objection to the very *concept* of rules of weight that they will be difficult to fashion. Nor does it appear to be the problem in *Allentown Mack*, for the Court suggested that the very idea of relying on rules when finding facts warrants suspicion, not merely that the rules at issue there were insufficiently precise.²⁷²

270 Professor Jennifer Mnookin, for instance, notes that a system of rules of weight seems hard to imagine because "evidence evaluation is so particularized, so fact-intensive, and so variable." See Mnookin, *supra* note 9, at 142–43. She thus rhetorically asks, "What could such rules of weight possibly look like?" *Id.* at 143. Mnookin's concern echoes that of both Bentham and Wigmore. See BENTHAM, *supra* note 41, at 180 ("To find infallible rules for evidence, rules which insure a just decision, is, from the nature of things, absolutely impossible."); see also JOHN HENRY WIGMORE, PRINCIPLES OF JUDICIAL PROOF 750 (1913) (noting that logic and psychology have "done nothing practical towards a method for measuring the net effect of a series or mass of mixed data bearing on a single alleged fact"). According to William Twining, for Wigmore, "the main difficulty [of evaluating evidence] relate[d] to the complexity of particular cases rather than to fundamental questions about epistemology or about the kinds of logical processes involved." TWINING, *supra* note 38, at 125. Professor Damaška argues that this view likely motivates the frequent criticisms of the Roman system of proof: given how variable and context-dependent factfinding is, any attempt to fashion rules suited to the purpose will inevitably result in "dangerous overgeneralizations." DAMAŠKA, *supra* note 9, at 21. Thus, under this view, Damaška explains, "To legislate on a subject so deeply contextual is like legislating against a chameleon by reference to its color." *Id.*

271 For instance, one study reported that whether a couple reported to be happily married or not correlated very well to whether or not the formula "rate of love making minus rate of arguments" had a positive value. See Dawes, *supra* note 189, at 393.

272 See *Allentown Mack Sales & Serv., Inc. v. NLRB*, 522 U.S. 359, 367–80 (1998).

B. *Moral Rationale*

The Court may have been more concerned with procedural fairness and political legitimacy than verdict accuracy. In noting, for instance, that when engaged in “simple factfinding,” the Board “is not free to prescribe what inferences from the evidence it will accept and reject,”²⁷³ the Court echoed Thayer’s declaration that “[t]he law has no mandamus to the logical faculty; it orders nobody to draw inferences.”²⁷⁴ This view reflects perhaps the deepest and most common objection to rules of weight, which condemns a factfinding process governed by technical rules as excessively “mechanical.”²⁷⁵ Hence the Court’s disdain for the NLRB’s “arbitrary rule of disregard.”²⁷⁶ But if the worry is not the epistemic one that using overbroad rules will lead to empirically incorrect verdicts, precisely wherein lies the problem?

There are a couple of possibilities. The first is that insofar as the factfinder reaches factual conclusions out of legal obligation rather than on the basis of her own reasoned judgments, such conclusions are politically illegitimate. Like presumptions, rules of weight may entail that the burden of proof is only reached “artificially” because the factfinder does not truly *believe* the facts found.²⁷⁷ But note, first, that in a deep sense this objection applies equally to the application of substantive law. The rigid application of substantive rules of law may often result in patently unjust consequences disconnected from their apparent purposes.²⁷⁸ Whether such rules of law legitimately command obedience in spite of this disconnect depends on one’s conception of legal authority. But at least under one well-established view, obedience to legal authority is justified on the *epistemic* ground that

273 *Id.* at 378.

274 THAYER, *supra* note 50, at 313 n.1.

275 *See, e.g.,* DAMAŠKA, *supra* note 9, at 19 (describing the criticism of the Roman system of proof as a system of “mechanical regulation” that “turned the fact-finder into a mere automaton compelled to come to a decision independently of his beliefs”).

276 *Allentown Mack*, 522 U.S. at 379.

277 *Cf. United States v. Klein*, 80 U.S. (13 Wall.) 128, 147 (1872) (striking down an act of Congress, which required federal courts to deem presidential pardons to be conclusive evidence of disloyalty to the Union, but barred courts from allowing such pardons to be admitted in support of any claim against the United States, on the ground that the Court was “forbidden to give the effect to evidence which, in its own judgment, such evidence should have, and is directed to give it an effect precisely contrary”).

278 *See, e.g.,* Roscoe Pound, *Mechanical Jurisprudence*, 8 COLUM. L. REV. 605, 606 (1908) (“Law has the practical function of adjusting every-day relations so as to meet current ideas of fair play. It must not become so completely artificial that the public is led to regard it as wholly arbitrary.”).

the authority is better able to balance the various reasons for action its subjects already have.²⁷⁹

The same is true of rules of weight; their justification rests on the premise that the authority promulgating the rule—whether a court or legislature—is better able to secure accurate verdicts by applying its own generalizations about evidentiary reliability, rather than those of the factfinder.²⁸⁰ In other words, the law may legitimately serve as a *theoretical* authority for the factfinder for the same reasons—and to the same extent—as the substantive law may legitimately serve as a *practical* authority for citizens.²⁸¹ Of course, if it becomes clear that their use results in systematically inaccurate verdicts, they would lose their ability to justify compliance with them—just as a body of criminal law would fail to obligate its subjects to comply with it if it systematically resulted in the unjust infliction of punishment. But whether that is true again depends not on the binding force of rules of weight as such but on the empirical adequacy of the generalizations they entrench.

Still, as a practical and doctrinal matter, the objection has some force, at least in the criminal context where procedural rights are most zealously protected. Indeed, for precisely these reasons, the Supreme Court has struck down presumptions that make it easier for prosecutors to convict defendants.²⁸² Even so, the objection only applies to rules of weight that result in affirmatively finding facts

279 See JOSEPH RAZ, *THE MORALITY OF FREEDOM* 38–70 (1988) (articulating and defending a “service” conception of authority); see also SCHAUER, *supra* note 133, at 126 & n.21 (suggesting that one justification for following rules may find support in Raz’s service conception of authority).

280 Cf. Pound, *supra* note 278, at 605 (“Law is scientific in order to eliminate so far as may be the personal equation in judicial administration, to preclude corruption and to eliminate the dangerous possibilities of magisterial ignorance. . . . Being scientific as a means toward an end, it must be judged by the results it achieves, not the niceties of its internal structure . . .”).

281 Indeed, under another view, the two types of authority are virtually indistinguishable. Raz explains that under the “recognitionist conception of authority,” an authority provides “reasons for belief” that there are reasons for acting in a certain way, but does not itself provide the reasons for action: “practical authority is reinterpreted as theoretical authority concerning belief in deontic propositions.” See RAZ, *supra* note 279, at 29–30; cf. Donald H. Regan, *Reasons, Authority, and the Meaning of “Obey”*: *Further Thoughts on Raz and Obedience to Law*, 3 CANADIAN J.L. & JURISPRUDENCE 3, 6 (1990) (arguing that authority provides only “indicative reasons” not “intrinsic reasons” for action).

282 See *supra* note 108; see also *Morissette v. United States*, 342 U.S. 246, 275 (1952) (“A presumption which would permit the jury to make an assumption which all the evidence considered together does not logically establish would give to a proven fact an artificial and fictional effect.”).

against the defendant, and such rules are neither common nor particularly desirable.²⁸³ Both the fairness and accuracy justifications for rules of weight offered in Part II call for rules that *suppress* the probative value of certain types of evidence, not for ones that elevate it. Indeed, rules that affirmatively demand that more weight be assigned to a given piece of evidence than a factfinder might otherwise be inclined to give it were not even common in the much-maligned Roman system of proof.²⁸⁴

A second possibility is that the moral defect of rules of weight lies not in their binding force but rather in their generality. Perhaps the ALJ undermined the “foundation of all honest and legitimate adjudication”²⁸⁵ because he paid too much attention to the NLRB’s institutionally sanctioned background generalizations about testimonial reliability, rather than “draw[ing] all those inferences that the evidence fairly demand[ed].”²⁸⁶ Identifying precisely why relying on generalizations in this way is unfair is surprisingly difficult, but the dilemma lies at the center of the debates over the use of “naked statistical evidence” and is usually revealed by hypothetical scenarios in which a plaintiff’s case depends almost entirely on general statistical evidence, rather than evidence derived from the specific facts of his case.²⁸⁷

283 The only example of such a rule that I have discovered is the “uncontradicted testimony rule,” according to which a jury is not permitted to ascribe zero weight to the testimony of an uncontradicted, unimpeached witness. See sources cited *supra* note 125.

284 See *supra* notes 20–24.

285 Allentown Mack Sales & Serv., Inc. v. NLRB, 522 U.S. 359, 379 (1998).

286 *Id.* at 378.

287 In the famous “Blue Bus” paradox, for instance, the only evidence as to who hit the plaintiff’s car on a dark night was that it was a blue bus and that eighty percent of the buses in the town were operated by the defendant. See Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1340–41 (1971). The hypothetical was based on the real case of *Smith v. Rapid Transit*, 58 N.E.2d 754, 754–55 (Mass. 1945). In the Gatecrasher paradox, the owner of a rodeo knows that 501 of 1000 spectators did not pay for admittance, so he picks one person off the seats and sues him for failure to pay. See L. Jonathan Cohen, *Subjective Probability and the Paradox of the Gatecrasher*, 1981 ARIZ. ST. L.J. 627, 627. The issue in both cases is whether such evidence is sufficient to find in the plaintiff’s favor or even get to a jury. For reviews of the literature describing the debates and noting that they have “not yet run their course,” see TWINING, *supra* note 8, at 74; William Twining & Alex Stein, *Introduction to EVIDENCE AND PROOF* xxi–xxiv (William Twining & Alex Stein eds., 1992) (previewing several essays on the subject). On the connection between these paradoxes and the generality of rules, see FREDERICK SCHAUER, *PROFILES, PROBABILITIES AND STEREOTYPES* 85 (2003), who notes that the Blue Bus and Gatecrasher paradoxes “are best seen as variants on the larger problem of generality in decisionmaking.” *Id.*

Many people's intuition is that finding facts against a defendant by reference to general categories rather than to the particular case-specific facts seems somehow unjust, even if the categorical approach produces more accurate verdicts in the long run.²⁸⁸ Assuming this intuition is sound,²⁸⁹ the difficulty is how to explain or justify it. The most sophisticated effort to do just that can be found in Professor Alex Stein's recent book.²⁹⁰ Stein argues that evidence law is and ought to be as concerned with the fair allocation of the risk of factfinding error as it is with reducing such errors.²⁹¹ To this end, he develops what he calls the "principle of maximal individualization" (PMI), according to which (1) "fact-finders must receive and consider all case-specific evidence pertaining to the case;" and (2) they "must not make any finding against a litigant, unless the argument generating this finding and the evidence upon which this argument rests were exposed to and survived maximal individualized examination."²⁹² Under this principle, because statistical evidence is not susceptible to individualized testing, it cannot provide a legitimate basis for fair adjudicative factfinding.²⁹³

288 See *supra* Part II.C.1.

289 This is by no means an uncontroversial assumption. Some scholars, for instance, argue that the defendants in the naked statistics paradoxes ought properly be held liable. See, e.g., SCHAUER, *supra* note 287, at 79–108; Koehler & Shaviro, *supra* note 15, at 272–77. Advocates of this view tend to consider verdict accuracy to be the primary, if not exclusive, goal of evidentiary and procedural law. See Shaviro, *supra* note 15, at 535 (defending the use of naked statistical evidence on the ground that "[t]rying to minimize the inevitable erroneous verdicts reflects a higher morality than trying to ignore them").

290 See STEIN, *supra* note 9. Several others have criticized the use of naked statistical evidence as well. See L. JONATHAN COHEN, *THE PROBABLE AND THE PROVABLE* 120 (1977) ("[T]he advancement of truth in the long run is not necessarily the same thing as the dispensation of justice in each individual case."); Lea Brilmayer & Lewis Kornhauser, *Review: Quantitative Methods and Legal Decisions*, 46 U. CHI. L. REV. 116, 149–50 (1978) (presenting, but questioning the adequacy of, an objection to the use of statistics generally at trial on the ground that such evidence "denies litigants their rights to be treated as individuals"); Laurence H. Tribe, *An Ounce of Detention: Preventive Justice in the World of John Mitchell*, 56 VA. L. REV. 371, 387 (1970) ("[T]he very enterprise of formulating a tolerable ratio of false convictions to false acquittals puts an explicit price on an innocent man's liberty and defeats the concept of a human person as an entity with claims that cannot be extinguished, however great the payoff to society.").

291 See, e.g., STEIN, *supra* note 9, at x–xi ("[T]he key function of evidence law is to apportion the risk of error in conditions of uncertainty, rather than facilitate the discovery of the truth.").

292 *Id.* at 100 (emphasis omitted).

293 See *id.* at 85.

Doing justice to Stein's impressive account of PMI requires more space than permitted here, but doing so is not necessary because even if we assume both that (1) there is an evidentiary norm that demands individualized attention to case-specific evidence and that (2) rules of weight such as those at issue in *Allentown Mack* potentially stand in some tension with such a norm, such assumptions still do not justify the Court's harsh condemnation of them, because some of the most common exclusionary rules raise the same concerns. The *Allentown Mack* Court emphasized that the Board could legitimately apply "counterfactual evidentiary presumptions" and exclusionary rules for *policy reasons*, so long as such rules were conceived "without reference to [their] inherent probative value."²⁹⁴ But in state and federal courts, evidence is often excluded because of its low probative value. The hearsay rule and its many exceptions, as we have seen, are best understood as largely motivated by concerns about evidentiary reliability.²⁹⁵ So, too, are rules excluding whole classes of evidence, such as polygraph tests. On what basis, then, may a court exclude such evidence entirely but not merely accord it low weight?²⁹⁶ Insofar as admissibility rules exclude evidence based on generalizations about the reliability of evidence, such rules implicate the same fairness concerns as do rules of weight.²⁹⁷

In short, none of the possible rationales for the Court's harsh criticism of rules of weight are sufficient to defeat their use *in toto*. At most, they suggest that certain types of rules may not be appropriate in certain adjudicatory contexts. And to the extent generalizing in factfinding is morally problematic, rules of weight fare no worse than do rules of admissibility. In fact, as suggested above, it may be that rules of weight could result in a system of relatively freer proof than

294 *Allentown Mack Sales & Serv., Inc. v. NLRB*, 522 U.S. 359, 378 (1998).

295 *See supra* note 99 and accompanying text.

296 Stein's answer with respect to hearsay is that such evidence does not satisfy PMI, but his argument is not entirely persuasive because it fails to account convincingly for the many exceptions to the hearsay rule. *See* Mike Redmayne, *The Structure of Evidence Law*, 26 OXFORD J. LEGAL STUD. 805, 819 (2006).

297 For this reason, the Court often views with suspicion rules that exclude evidence categorically on the ground of its unreliability, even though many common rules of admissibility are similarly justified. *Compare* *Rock v. Arkansas*, 483 U.S. 44, 61 (1987) (striking down, as a violation of the Compulsory Process Clause, a state supreme court's *per se* exclusion of post-hypnosis testimony on the ground that a "State's legitimate interest in barring unreliable evidence does not extend to *per se* exclusions that may be reliable in an individual case"), *with id.* at 64 n.* (Rehnquist, C.J., dissenting) (noting that he failed to "discern any meaningful constitutional difference between [admissibility] rules and the one at issue here").

the current system.²⁹⁸ The *Allentown Mack* decision, for instance, could have prompted the NLRB to develop an exclusionary rule that forbid companies from offering testimony about employee comments made during interviews with the company at all for the purpose of showing that it had a good faith reasonable doubt about the union's ongoing support. The suggestion is admittedly speculative, but it indicates how the use of rules of weight could present an attractive alternative to the all-or-nothing approach of exclusionary rules.

CONCLUSION

Of course, all of this may raise more questions than it answers. If we were to develop a system of rules of weight, who should develop them? What precedential effect would they have? What standard of review ought apply to them? Answering such questions requires a more granular analysis, and the answers may differ across adjudicatory contexts. In this Article, my aim has been simply to suggest that such inquiries are worthwhile, and that rules of weight ought not be categorically discarded as anachronistic or alien to the common law.

Finally, and more generally, my hope is that this Article also has made clear that our thinking about facts may profit from our thinking about law—and vice versa. I defined a rule of weight in part as one whose primary purpose was to ensure evidentiary reliability, but perhaps rules of weight may legitimately serve independent policy goals, much as presumptions do. In the administrative context, then, might judicial deference to agency-developed rules such as those at issue in *Allentown Mack* be justified on grounds akin to those underlying the *Chevron* doctrine of statutory construction?²⁹⁹ Alternatively, if rules of weight stand in tension with an evidentiary norm that demands individualized attention to the particular facts of the case, might there be in some contexts a comparable demand on judges when interpreting the law? Whatever the answers to such questions, a recognition that the law treats legal questions quite differently from factual ones ought to begin the inquiry, not end it.

298 See *supra* text accompanying note 240.

299 See *Chevron U.S.A. Inc. v. Natural Res. Def. Council, Inc.*, 467 U.S. 837, 842–43 (1984).

