

**PREPONDERANCE OF THE EVIDENCE
VERSUS *INTIME CONVICTION*:
A BEHAVIORAL PERSPECTIVE
ON A CONFLICT BETWEEN AMERICAN
AND CONTINENTAL EUROPEAN LAW**

Christoph Engel^{*†}

INTRODUCTION

Continental law is irrational. American law is irresponsible. These beliefs are the essence of one of the few true conflicts between American common law and Continental Civil Law. At the surface, the conflict is confined to an apparently technical issue in the law of evidence. On the European continent, for the court to hold against the defendant, the judge must be convinced that the facts brought forward by the plaintiff in support of the claim are indeed true. In principle, Continental law does not differentiate between civil law and criminal law. The standard of proof is *intime conviction*¹ throughout.² By contrast, American law has three different standards of proof.³ In criminal law, the charge must be established “beyond a reasonable doubt.”⁴ In civil law, the plaintiff prevails only if “the preponderance of the evidence” is in the plaintiff’s favor.⁵ Only in a limited number of civil law matters, of particular gravity for the defendant, must the intermediate standard of “clear and convincing evidence” be met.⁶

* Professor of Law, Bonn University; Director, Max Planck Institute for Research on Collective Goods, Bonn. Habilitation 1992, University of Hamburg; Dr. jur. 1988, University of Tübingen.

† Helpful comments by Kevin Clermont, Stephan Dickert, and Martin Hellwig are gratefully acknowledged.

1. “*Intime Conviction*” is a French term that means “reasonable conviction; reasonable certainty; state of being satisfied beyond reasonable doubt (personally convinced); personal conviction of the court (after considering all the evidence).” F.H.S. BRIDGE, THE COUNCIL OF EUROPE FRENCH-ENGLISH LEGAL DICTIONARY 173 (2002).

2. E.g., CODE DE PROCÉDURE PÉNALE [C. PR. PÉN.] art. 353 (Fr.), translated in John Rason Spencer, Code of Criminal Procedure 100 (Jan. 1, 2006), http://195.83.177.9/upl/pdf/code_34.pdf; Zivilprozeßordnung [ZPO] [civil procedure statute] Dec. 5, 2005, Bundesgesetzblatt, Teil I [BGBl. I] 3202, § 286(1) (F.R.G.), translated in STEFAN RÜTZEL, GERHARD WEGEN, & STEPHAN WILSKE, COMMERCIAL DISPUTE RESOLUTION IN GERMANY 224 (2005); Strafprozeßordnung [StPO] [Code of Criminal Procedure] Apr. 7, 1987, Bundesgesetzblatt, Teil I [BGBl. I] 1074, last amended by Gesetz, Oct. 31, 2008, BGBl. I at 2149, § 261 (F.R.G.), available at <http://bundesrecht.juris.de/bundesrecht/stpo/gesamt.pdf>.

3. See, e.g., *Addington v. Texas*, 441 U.S. 418, 422–25 (1979) (discussing the three commonly applied standards of proof in American law).

4. *Id.* at 424.

5. *Id.* at 423.

6. *Id.* at 424.

On closer inspection, the conflict roots deeper. American law not only differentiates; it also conceptualizes proof differently. On the Continent, proof is understood as a strictly subjective impression in the judge's mind. By contrast, in American law, proof is an objective concept. The conflict in the law of evidence is closely related to a fundamental divide in epistemology. In line with the prevalent position in the sciences, the American law of evidence aims at objectivity. The Continental law of evidence is closer to the opposite view that, in epistemology and in the history of sciences, is characterized as a quest for truth.⁷ Partisans of the latter view stress that even science, if understood as practice, has an artistic component.⁸ Ultimately, the position holds that strict intersubjectivity is prohibitively costly.⁹

As usual, at closer sight, the conflict is more nuanced. There are partisans of probabilistic approaches on the Continent.¹⁰ And, particularly in criminal law, the strict objectivist position is hotly debated in the United States.¹¹ Yet the underlying difference is net. It is both normative and positive. Which is the goal of judicial procedure? And by which means is this goal best attained? All legal orders agree that courts have to decide under conditions of uncertainty.¹² It is less clear, however, whether the appropriate way of dealing with this uncertainty is some form of utilitarian calculus.¹³ Alternatively, and more modestly, judicial procedure might only strive to make error sufficiently unlikely.¹⁴

The deepest divide, however, is psychological. Those who believe in objectivity point to the fact that individual judgment is error prone. It is the purpose of procedural law to impose as much rationality as possible as a

7. See LORRAINE DASTON & PETER GALISON, OBJECTIVITY 58 (2007) (discussing "the quest for truth-to-nature," a smaller component of "the ur-epistemic virtue" of "seeking truth").

8. *Id.* at 58–59.

9. *Id.*

10. A prominent voice in German law is Gerhard Kegel. See, e.g., Gerhard Kegel, *Der Individualanscheinsbeweis und die Verteilung der Beweislast nach Überwiegender Wahrscheinlichkeit*, in *DAS UNTERNEHMEN IN DER RECHTSORDNUNG* 321 (Kurt H. Biedenkopf et al. eds., 1967).

11. See generally James Franklin, *Case Comment—United States v. Copeland*, 369 *F. Supp. 2d* 275 (E.D.N.Y. 2005): *Quantification of the "Proof Beyond Reasonable Doubt" Standard*, 5 *LAW PROBABILITY & RISK* 159 (2006) (arguing that proof beyond a reasonable doubt should mean well above eighty percent certainty).

12. James Brook, *Inevitable Errors: The Preponderance of the Evidence Standard in Civil Litigation*, 18 *TULSA L. REV.* 79, 79 (1982).

13. See, e.g., Peter Tillers & Jonathan Gottfried, *Case Comment—United States v. Copeland*, 369 *F. Supp. 2d* 275 (E.D.N.Y. 2005): *A Collateral Attack on the Legal Maxim That Proof Beyond A Reasonable Doubt Is Unquantifiable?*, 5 *LAW PROBABILITY & RISK* 135 (2006) (clarifying the advantages and disadvantages of quantifying the reasonable-doubt standard).

14. See, e.g., Jon O. Newman, *Beyond "Reasonable Doubt"*, 68 *N.Y.U. L. REV.* 979 (1993) (suggesting ways in which appellate judges can better enforce the reasonable-doubt standard).

check.¹⁵ Those believing in truth counter that rational calculus is not what empirical judges and jurors do. They believe a norm for decision quality is pointless if judges and jury members have no chance of living up to it. Empirically, from the evidence presented to them, judges and jury members construct stories. They decide for the plaintiff, or for the prosecution for that matter, if it has been able to induce a coherent story, with no striking gaps, that resonates with the world knowledge of the juror.¹⁶ While having deliberate components, the bulk of this mental activity is performed by the automatic system.¹⁷ Only the result is propelled back to consciousness as an intuition.¹⁸ Specifically, consciousness is not only informed that the automatic system has come to a resolution. This information is combined with a level of confidence.¹⁹

Actually, at this point emotions come into play. It is quite plausible that the level of confidence is itself tagged by a somatic marker. While this hypothesis still awaits the empirical test, a related effect is well established. If an individual is aware of a serious risk, this prospect is emotionally highlighted by a somatic marker. It induces the individual to steer clear from the risk, even if the ensuing benefit seems appealing.²⁰ Judges and jurors decide on other people's lives. In so doing, they exercise sovereign powers. In this perspective, asking them for their *intime conviction* is a way of making personal accountability salient.²¹ The jury instruction is a way to trigger the somatic marker, and thereby to influence how the automatic system assesses the evidence.

The automatic system is very powerful. It handles vast amounts of information in almost no time. It elegantly combines current information with the traces of previous experiences stored in memory. Most importantly, it enables human decision-makers to make meaningful

15. See Jonathan J. Koehler, *Train our Jurors*, in HEURISTICS AND THE LAW 303, 303 (Gerd Gigerenzer & Christoph Engel eds., 2006) (arguing that jury verdicts reflect systematic biases, which procedural rules attempt to minimize and balance against other policy concerns).

16. See Reid Hastie, *Introduction to INSIDE THE JUROR: THE PSYCHOLOGY OF JUROR DECISION MAKING* 3, 3–4 (Reid Hastie ed., 1993) (discussing the juror decision-making process and accompanying theories generally).

17. Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511, 512–13 (2004).

18. *Id.* at 533.

19. *Id.* at 517.

20. ANTONIO R. DAMASIO, *DESCARTES' ERROR: EMOTION, REASON, AND THE HUMAN BRAIN* 173 (1994).

21. See Philip E. Tetlock, *Accountability and Complexity of Thought*, 45 J. PERSONALITY & SOC. PSYCHOL. 74, 74 (1983) (studying the impact of accountability on individual thought processes and suggesting that increased accountability leads to more complex reasoning).

decisions on an incomplete factual basis.²² The automatic system is not foolproof, however. It can be betrayed by spurious analogies.²³ It is influenced by stereotypes and predilections,²⁴ and it can be manipulated by outsiders.²⁵ The heart of the conflict between American and Continental law is therefore their different assessments of cost and benefit. Whereas Continental law is attracted by the benefit, American law is scared of the cost. At the end of the day, this is a normative issue. Once one becomes aware of what the conflict is really about, one is in a better position to decide. And, hopefully, clever institutional arrangements might be able to overcome the apparent yes-or-no decision.

The remainder of this Article is organized as follows: Part I takes a closer look at doctrine on both sides of the Atlantic. Part II discusses the norm. Part III establishes the psychological underpinnings. The Article ends with normative conclusions.

I. DOCTRINE

Comparative law tends to be a disinterested affair. Not so in this case. Two American observers—Kevin Clermont and Emily Sherwin—wonder “how civilians can be so wrong.”²⁶ To them, the lack of sophistication in the Continental law of evidence “is downright bizarre.”²⁷ They put it down to the “disruption in legal procedure” resulting from the French Revolution,²⁸ and to a lack of interest in later years,²⁹ which induced civilian legal systems “to adhere to the unrealistic and potentially unfair and inefficient standard of *intime conviction* in civil cases.”³⁰ They call on civilian lawyers to improve their system in light of the superior American solution.³¹

An Italian commentator, Michele Taruffo, did not mince his words either, asserting that Clermont and Sherwin have fallen prey to the “reductivist fallacy.”³² “They adduce some scattered statements made by

22. Christoph Engel & Wolf Singer, *The Brain, the Psyche, Behavior, and Institutions, in* BETTER THAN CONSCIOUS? 12–13 (Christoph Engel & Wolf Singer eds., 2008).

23. Simon, *supra* note 17, at 538.

24. *Id.* at 536–37.

25. *Id.* at 539–40.

26. Kevin M. Clermont & Emily Sherwin, *A Comparative View of Standards of Proof*, 50 AM. J. COMP. L. 243, 244 (2002).

27. *Id.* at 254.

28. *Id.* at 258.

29. *Id.* at 259.

30. *Id.*

31. *See id.* at 274 (“[W]e shed enough light to justify calling on civilians, immersed in their own system but traditionally inattentive to standard of proof, to study and explain this bizarre puzzle.”).

32. Michele Taruffo, *Rethinking the Standards of Proof*, 51 AM. J. COMP. L. 659, 659 (2003).

some European scholars (mainly French), taking them out of their theoretical context and without checking to which extent they correctly represent the practice of courts in [C]ivil [L]aw countries.”³³ Moreover, Taruffo asserts that Clermont and Sherwin fail to understand that *intime conviction* is not the same as “beyond [a] reasonable doubt.”³⁴ “They rely upon a rather naïve idea of probability, roughly corresponding to the popular concept of statistical or quantitative probability, and upon a naïve and unqualified idea of truth.”³⁵

Acrimony is not the recipe for improving scientific understanding. Let us take a disinterested view of doctrine on both sides of the Atlantic. There is a clear difference in the handling of civil-law disputes. In the United States, the plaintiff prevails if only “the preponderance of the evidence” speaks in her favor. “The litigants thus share the risk of error in roughly equal fashion.”³⁶ The Pennsylvania Suggested Standard Civil Jury Instructions put the standard thus:

Preponderance of the evidence means the claim is more likely true than not.

. . . .

Think about an ordinary balance scale with a pan on each side to hold objects. Imagine using the scale as you deliberate in the jury room. Place all the evidence favorable to the plaintiff, in one pan. Place all the evidence favorable to the defendant in the other. If the scales tip, even slightly, to the plaintiff’s side, then, you must find for the plaintiff.³⁷

There is some criticism of the standard in American legal literature, but this criticism is chiefly concerned with extending the standard into matters that are close to criminal law, with a focus on sentencing under the Federal Sentencing Guidelines.³⁸ Another line of criticism is concerned with the

33. *Id.* at 664–65.

34. *Id.* at 666.

35. *Id.* at 669.

36. *Addington v. Texas*, 441 U.S. 418, 423 (1979).

37. PENNSYLVANIA SUGGESTED STANDARD CIVIL JURY INSTRUCTIONS § 1.42 (3d ed. 2005).

38. *See, e.g.*, Steven M. Salky & Blair G. Brown, *The Preponderance of Evidence Standard at Sentencing*, 29 AM. CRIM. L. REV. 907, 908, 918 (1992) (arguing that the Federal Sentencing Guidelines should employ a higher standard of proof); Lauren Greenwald, Note, *Relevant Conduct and the Impact of the Preponderance Standard of Proof Under the Federal Sentencing Guidelines: A Denial of Due Process*, 18 VT. L. REV. 529, 530, 553 (1994) (questioning the constitutionality of the standard of proof applied under the Federal Sentencing Guidelines); Daniel J. Lyons, Note, *Federal Sentencing Guidelines: Retaining the Preponderance Standard of Proof*, 67 ST. JOHN’S L. REV. 639, 641–42, 653 (1993) (discussing the constitutional sufficiency of the preponderance standard for convicted criminals); Joseph P. Sargent, Comment, *The Standard of Proof Under the Federal Sentencing Guidelines: Raising*

robustness of judicial outcomes to small perturbations, and recommends an alternative standard that weighs probability with the normative desirability of the respective outcome.³⁹ American legal scholars thus seem to feel comfortable with the objectivist approach, and with the probabilistic interpretation of the standard of proof.

By contrast, the standard of proof in Civil Law countries is predominantly interpreted as being nonprobabilistic.⁴⁰ The classic formulation is in article 353 of the French Code of Criminal Procedure:

The law does not ask the judges to account for the means by which they convinced themselves; it does not charge them with any rule from which they shall specifically derive the fullness and adequacy of evidence. It requires them to question themselves in silence and reflection and to seek in the sincerity of their conscience what impression has been made on their reason by the evidence brought against the accused and the arguments of his defence. The law asks them but this single question, which encloses the full scope of their duties: are you inwardly convinced?⁴¹

This standard is also applied in civil-law disputes,⁴² although there is no explicit provision to that effect in French law.⁴³

In German law, the situation is even clearer. Section 286(1) of the

the Standard to Beyond a Reasonable Doubt, 28 WAKE FOREST L. REV. 463, 484–85 (1993) (suggesting that the Supreme Court should raise the standard of proof under the Federal Sentencing Guidelines); Stephanie C. Slatkin, Note, *The Standard of Proof at Sentencing Hearings Under the Federal Sentencing Guidelines: Why the Preponderance of the Evidence Standard is Constitutionally Inadequate*, 1997 U. ILL. L. REV. 583, 609 (1997) (claiming that “the preponderance standard does not adequately protect the liberty interest in freedom from excessive punishment created by the Sentencing Guidelines”); Michael D. Wysocki, Comment, *Beyond a Reasonable Doubt: The Effects of Blakely v. Washington, United States v. Booker, and the Future of the Federal Sentencing Guidelines*, 38 TEX. TECH L. REV. 495, 529–30 (2006) (contending that the standard of proof under the Federal Sentencing Guidelines needs to conform with constitutional principles).

39. See Franklin, *supra* note 11, at 161–62, 165 (identifying the necessity of deciding close cases on “probabilities of low weight” as “alarming,” and advocating abandoning the preponderance-of-the-evidence standard in favor of a higher burden of proof); Neil Orloff & Jerry Stedinger, *A Framework for Evaluating the Preponderance of the Evidence Standard*, 131 U. PA. L. REV. 1159, 1159–60 (1983) (criticizing the differences in outcome to defendants in uncertain cases as “unduly harsh” under the preponderance-of-the-evidence standard, and proposing the “expected value rule” as an alternative that would award damages in proportion to the probability of the defendant’s culpability).

40. Clermont & Sherwin, *supra* note 26, at 265.

41. C. PR. PÉN. art. 353 (Fr.), translated in John Rason Spencer, Code of Criminal Procedure 100 (Jan. 1, 2006), http://195.83.177.9/upl/pdf/code_34.pdf.

42. Jean-Denis Bredin, *Le doute et l’intime conviction [Doubt and Intime Conviction]*, 23 DROITS: REVUE FRANÇAISE DE THÉORIE, DE PHILOSOPHIE ET DE CULTURE JURIDIQUES 21, 23 (1996) (Fr.).

43. Taruffo, *supra* note 32, at 667.

Code of Civil Procedure states that: “The court shall decide at its discretion, by taking into account the entire substance of the hearings and the result of any evidence taken, whether an allegation regarding the facts should be regarded as true or untrue.”⁴⁴ In the leading case, the German Supreme Court has made it clear that the judge may not content herself with a mere assessment of probabilities.⁴⁵ Even a very high probability would not be enough. Initial doubt is acceptable, but the judge must have overcome this doubt.⁴⁶ This is not meant to defer to judicial discretion, but to judicial intuition.⁴⁷ The standard is an empirical one.⁴⁸ The crucial feature is “the psychological state of taking a fact for true.”⁴⁹ The test is predominantly built on “ethos, experience and intuition.”⁵⁰

Interestingly, in criminal procedure the conflict is less pronounced. In the United States, many instructions for criminal law juries have a subjectivist flavor. For instance, the Pennsylvania Suggested Standard Criminal Jury Instructions read:

To find the defendant guilty beyond a reasonable doubt, you must be convinced of [his] [her] guilt to the same degree you would be convinced about a matter of importance in your own life in which you would act with confidence and without restraint or hesitation.⁵¹

....

... A reasonable doubt is a doubt that would cause a reasonably careful and sensible person to hesitate before acting upon a matter of importance in his or her own affairs.⁵²

In the explanatory part, the document says:

In essence, to prove facts beyond a reasonable doubt is simply to bring a jury to a particular state of certainty. It is akin to raising the temperature in the room to 75 degrees; no one may

44. Zivilprozeßordnung [ZPO] [civil procedure statute] Dec. 5, 2005, BGBl. I 3202, § 286(1) (F.R.G.), *translated in* STEFAN RÜTZEL, GERHARD WEGEN, & STEPHAN WILSKE, *COMMERCIAL DISPUTE RESOLUTION IN GERMANY* 224 (2005).

45. Bundesgerichtshof [BGH] [Federal Court of Justice] Feb. 17, 1970, 53 Entscheidungen des Bundesgerichtshofes in Zivilsachen [BGHZ] 245 (F.R.G.).

46. *Id.* at 255–56.

47. HANS-JOACHIM MUSIELAK, *KOMMENTAR ZUR ZIVILPROZESSORDNUNG* 872 (2007).

48. JOACHIM SCHULZ, *SACHVERHALTSFESTSTELLUNG UND BEWEISTHEORIE: ELEMENTE EINER THEORIE STRAFPROZESSUALER SACHVERHALTSFESTSTELLUNG* 42–43 (1992).

49. *Id.* at 43.

50. *Id.* at 168.

51. PENNSYLVANIA SUGGESTED STANDARD CRIMINAL JURY INSTRUCTIONS § 2.01 (2d ed. 2005).

52. *Id.* § 7.01.

hold in their hand or find the 75th degree in the room, but one may sense the presence of that state nonetheless. Were the courts ever to develop a meter that worked with thermometer-like precision and ease of determination, judges could advise jurors on when the evidence in a case rose to a level that guilt would be shown beyond a reasonable doubt, much as a scientist could tell his or her class at what temperature they might expect the water to boil.⁵³

The Massachusetts Supreme Court once put it:

Then, what is reasonable doubt? . . . It is that state of the case, which, after the entire comparison and consideration of all the evidence, leaves the minds of jurors in that condition that they cannot say they feel an abiding conviction, to a moral certainty, of the truth of the charge. . . . For it is not sufficient to establish a probability, though a strong one arising from the doctrine of chances, that the fact charged is more likely to be true than the contrary; but the evidence must establish the truth of the fact to a reasonable and moral certainty. . . .⁵⁴

This remains the law,⁵⁵ despite voices in the literature calling for a quantification of the standard.⁵⁶

II. NORMS

Whether the American or the Continental approach is superior depends on the goals that the legal order strives to achieve by defining the standard of proof. In legal literature, a more modest and a more ambitious approach compete.⁵⁷ At the most basic level, a consequentialist approach looks at the implications for the decision taken in this one case. At a more sophisticated level, one might also be concerned with the implications for the functioning of jurisdiction and for the place of jurisdiction in the social fabric,

53. *Id.* § 7.01 cmt. (citation omitted) (citing Commonwealth v. Jones, 2004 PA Super. 301 ¶¶ 4–9).

54. Commonwealth v. Webster, 59 Mass. (5 Cush.) 295, 320 (1850).

55. Jack B. Weinstein & Ian Dewsbury, *Comment on the Meaning of “Proof Beyond a Reasonable Doubt,”* 5 LAW, PROBABILITY & RISK 167, 170 (2006).

56. Tillers & Gottfried, *supra* note 13, at 156–57; Weinstein & Dewsbury, *supra* note 55, at 170–73; Harry D. Saunders, *Quantifying Reasonable Doubt: A Proposed Solution to an Equal Protection Problem 2* (bepress Legal Series, Working Paper No. 881, 2005), available at <http://law.bepress.com/expresso/eps/881>. *But see* Franklin, *supra* note 11, at 165 (“A first step towards justice would be to rule out numerical probabilities that are clearly unreasonable.”).

57. Elisabeth Stoffelmayr & Shari Seidman Diamond, *The Conflict Between Precision and Flexibility in Explaining “Beyond a Reasonable Doubt,”* 6 PSYCHOL. PUB. POL’Y & L. 769, 770 (2000).

particularly for perceptions of its legitimacy. In a consequentialist approach, one is concerned with both the probabilities of different types of errors and their consequences. A utilitarian version of this approach would postulate a social welfare function and would assess the different norms in terms of their welfare implications. A cruder approach would postulate rule-of-thumb values for the cost of false convictions versus false acquittals and, on this basis, develop norms for assessing the trade-off between the probabilities of the different types of errors.

All agree that the problem originates in the social cost of error. The modest approach confines itself to making the probability of certain errors sufficiently unlikely. The more ambitious approach is utilitarian. From the latter perspective, it is easy to incorporate additional concerns that have been voiced in the complementary discussion in economics. The standard of proof will be anticipated by private parties in their pretrial decision-making. They, for instance, might decide to increase their level of care. Moreover, society might also care about the transaction costs incurred by the parties and the courts. These concerns can be incorporated as additional arguments into society's welfare function.

A. Acceptable Probability of Error

Frequently, the courts are not in a position to establish the facts with near certainty. Even if they subjectively believe a stated fact to be true, there is always at least some small objective probability that the statement is actually false.⁵⁸ The most basic normative question to be addressed by the standard of proof is defining the maximum tolerable probability of this happening.⁵⁹ The United States Supreme Court has put the question thus:

In a criminal case . . . the interests of the defendant are of such magnitude that . . . they have been protected by standards of proof designed to exclude as nearly as possible the likelihood of an erroneous judgment. In the administration of criminal justice, our society imposes almost the entire risk of error upon itself.⁶⁰

58. See Brook, *supra* note 12, at 79 ("In a system in which disputed facts are judged by fallible humans, as opposed to one such as trial by ordeal which calls upon divine guidance, it is acknowledged that, even under the best circumstances, errors will occur."); Stoffelmayr & Diamond, *supra* note 57, at 770.

59. Newman, *supra* note 14, at 981.

60. *Addington v. Texas*, 441 U.S. 418, 423–24 (1979) (citations omitted).

The legal order is in the same position as a statistician.⁶¹ Both are in the business of drawing inferences from evidence they are aware might be biased. Statisticians are happy to make the inference if the probability of an unwarranted statement is sufficiently small.⁶² To do so, they compare the tested hypothesis, or null hypothesis, to some alternative hypothesis.⁶³ In criminal law, the analogue to the alternative hypothesis is the charge. The null hypothesis is the presumption of innocence.⁶⁴

A statistician will only accept the alternative hypothesis if she is able to refute the competing null hypothesis with very high probability.⁶⁵ In the social sciences, the convention is this: the probability that the null hypothesis is false must be below five percent.⁶⁶ That way, statisticians make sure that alpha errors are very unlikely. An alpha error occurs if the tested hypothesis is taken to be false, but in reality is true.⁶⁷ Even if this standard has not been met by the available data, this is not to say that the opposite is proven to be true. In statistical jargon, the finding is simply not significant.⁶⁸ This does not exclude the possibility that, with better data, the statement could be proven to the requisite standard.⁶⁹ Likewise, in criminal law, guilt must be proven “beyond a reasonable doubt.” The law focuses on alpha errors and requires them to be very rare. By implication, “preponderance of the evidence” requires much less certainty. The legal order tolerates a substantially higher error rate.⁷⁰

One may wonder why society dislikes false convictions. Obviously, the constitutional presumption of innocence matters.⁷¹ Moreover, society might want to make sure that the defendant’s effort to convince the court of her innocence is more productive if the defendant is actually innocent.⁷² Last,

61. David Kaye, *The Limits of the Preponderance of the Evidence Standard: Justifiably Naked Statistical Evidence and Multiple Causation*, 7 LAW & SOC. INQUIRY 487, 501 (1982); Thomas J. Miceli, *Optimal Prosecution of Defendants Whose Guilt is Uncertain*, 6 J.L. ECON. & ORG. 189, 189, 191 (1990).

62. WILLIAM L. HAYS, STATISTICS 282 (5th ed. 1994).

63. *Id.* at 269.

64. Michael L. Davis, *The Value of Truth and the Optimal Standard of Proof in Legal Disputes*, 10 J.L. ECON. & ORG. 343, 346 (1994).

65. HAYS, *supra* note 62, at 281–82.

66. *Id.* at 283.

67. *Id.* at 282.

68. *Id.* at 297.

69. *See id.* at 281 (noting that social scientists may deem it best to suspend their judgment due to inconclusive results, especially when the harm in waiting for better evidence is less than reaching an incorrect conclusion).

70. Brook, *supra* note 12, at 85.

71. Davis, *supra* note 64, at 346.

72. *See* Daniel L. Rubinfeld & David E. M. Sappington, *Efficient Awards and Standards of Proof in Judicial Proceedings*, 18 RAND J. ECON. 308, 310 (1987) (“The basic goals of the adversarial

but not least, the general acceptance of the legal system might suffer if the courts make patently false decisions.⁷³

Every empirical researcher has had this experience: she is subjectively convinced that she has indeed found the effect, yet she is unable to refute the null hypothesis. The finding is not significant, and therefore not publishable. A promising line of research has been a waste of time and resources. The legal order is in the same situation. The stricter one is with alpha errors, the more likely it becomes that one commits the opposite error. In statistical jargon it is called the beta error.⁷⁴ A beta error occurs if one wrongly refutes the alternative hypothesis.⁷⁵ Empirical researchers are trained to respond by calculating test power before they engage in the generation of data. The smaller the expected effect, the larger the sample must be to meet the significance level.⁷⁶ The analogue in judicial proceedings is the amount of energy that goes into producing evidence. The stricter the standard of proof, the more effort is required by the plaintiff, or by the prosecution for that matter.

Ultimately, this effort may become prohibitive. Moreover, the legal order might consider it unfair to impose such a heavy burden on the plaintiff. This concern drives the harsh criticism of the Continental standard of proof cited at the beginning of this Article.⁷⁷ Lowering the standard of proof is, however, not the only technology for helping plaintiffs where the legal order deems this appropriate. The strongest intervention is shifting the burden of proof to the defendant.⁷⁸ Making prima facie evidence acceptable

process suggest that litigation effort by an innocent defendant generally should be more effective than an equal expenditure of effort by a guilty defendant.”).

73. See Clermont & Sherwin, *supra* note 26, at 271 (explaining that “the high civil-law standard of proof . . . perhaps exists in hope of enhancing the perceived legitimacy of judicial decisions”); Dominique Demougin & Claude Fluet, *Deterrence vs. Judicial Error: A Comparative View of Standards of Proof*, 161 J. INSTITUTIONAL & THEORETICAL ECON. 193, 193–96 (2005) [hereinafter Demougin & Fluet, *Deterrence*] (assuming that judicial error, in addition to impacting a legal system’s power to deter unlawful conduct, is undesirable *per se*); Franklin, *supra* note 11, at 159 (noting that “[t]he majesty of the law and its powers of deterrence would be ill-served, if the law were forced to admit the truth about the number of false convictions it allows and the number of criminals it allows to go free.”); Eric Lillquist, *Recasting Reasonable Doubt: Decision Theory and the Virtues of Variability*, 36 U.C. DAVIS L. REV. 85, 176–77, 184–85 (2002) (explaining the view that society’s concern for fairness in its legal system justifies the reasonable-doubt standard, “irrespective of any costs and benefits from accurate and inaccurate verdicts”).

74. HAYS, *supra* note 62, at 282.

75. *Id.*

76. *Id.* at 291–92.

77. *E.g.*, Clermont & Sherwin, *supra* note 26, at 243 (opining that the Continental standard of proof “seems strange”).

78. Taruffo, *supra* note 32, at 672. *But cf.* Weinstein & Dewsbury, *supra* note 55, at 167 (suggesting that the burden of proof in criminal law is meant to favor the defendant).

is a more cautious intervention.⁷⁹ The plaintiff may content herself with proving the preconditions of an established factual regularity. It then is for the defendant to cast sufficient doubt on the claim that the case at hand is in line with this regularity. Finally, instead of lowering the standard of proof across the board, a more lenient standard may exclusively be applied to facts that are particularly hard to prove. An example is the amount of damage once liability has been established.⁸⁰

B. Utilitarian Norm

In the American literature, there is mounting unease with the exclusive focus on error rates.⁸¹ In essence, this criticism is utilitarian.⁸² Authors advocate defining a social welfare function.⁸³ The main advantage of this approach is conceptual. It becomes possible to address alpha and beta errors within the same framework.⁸⁴ Quite naturally, the welfare function features two more arguments: the conviction of guilty defendants and the acquittal of innocent defendants.⁸⁵ Each of these four events can be weighted. The social optimum requires choosing the standard of proof such that this function is maximized.

Once one starts from a social welfare function, it is easy to add further arguments. Economists have claimed that the standard of proof should also be chosen such that innocent defendants stand a chance to credibly signal

79. See, e.g., Rhea Gertken, Note, *Causation in Retaliation Claims: Conflict Between the Prima Facie Case and the Plaintiff's Ultimate Burden of Pretext*, 81 WASH. U. L.Q. 151, 172–73 (2003) (arguing that in retaliation cases, “the standard for causation in the prima facie case ought to be heavier than just showing temporal proximity and should include the overall pattern of antagonism”).

80. As stipulated by Zivilprozeßordnung [ZPO] [civil procedure statute] Dec. 5, 2005, BGBI. I 3202, § 287(1) (F.R.G.), translated in STEFAN RÜTZEL, GERHARD WEGEN, & STEPHAN WILSKE, COMMERCIAL DISPUTE RESOLUTION IN GERMANY 224–25 (2005).

81. See, e.g., Kaye, *supra* note 61, at 493, 501 (suggesting that “the error equalizing characteristic” of the expected-value rule, which provides the injured party damages multiplied by the probability that the costs are attributable to the defendant, does not “provide a rationale for its unrestricted use”); Lillquist, *supra* note 73, at 195 (noting that jury instructions in the United States do not need to be altered to eliminate error because they are inevitably flexible and allow for different amounts of proof in different cases to convict criminals); Orloff & Stedinger, *supra* note 39, at 1167 (demonstrating that “the measure of large errors under the preponderance-of-the-evidence standard will always be greater than the measure of large errors under the expected value standard”).

82. Brook, *supra* note 12, at 86; see also Davis, *supra* note 64, at 348 (“It would seem reasonable to suppose that the total cost to society of each type of error will be an increasing function of the total number of errors.”).

83. Tillers & Gottfried, *supra* note 13, at 152.

84. *Id.*

85. See Miceli, *supra* note 61, at 189 (discussing two types of legal errors: “incorrect conviction of innocent defendants (type I errors), and incorrect acquittal of guilty defendants (type II errors”).

their innocence or guilt.⁸⁶ Others have suggested that the standard of proof might be used to calibrate would-be defendants' incentives to exert care.⁸⁷ Moreover, in maximizing welfare, society might also take into account the transaction cost of judicial proceedings that must be borne by defendants,⁸⁸ plaintiffs,⁸⁹ and the courts.⁹⁰

Although appealing from a theoretical perspective, a utilitarian norm has a number of undesirable properties. The positive utility from convicting guilty defendants, and from not committing beta errors, is bound to differ between classes of cases. Likewise, the negative utility from committing an alpha error will not in all cases be the same.⁹¹ A variable standard of proof may be difficult to administer, and it may meet with public resistance. Actually, if the utilitarian approach is taken seriously, it is not enough to differentiate between types of offenses. One would also need knowledge about the composition of the respective population.⁹² It will be next to impossible to make this knowledge available in court. And it might not be normatively acceptable to have a different standard of proof for each and every case.⁹³ Quite a few concerns that would matter from a strict utilitarian perspective might also not be acceptable in court, like the criminal record of the defendant.⁹⁴ Both with respect to practicality and to the legitimacy of the judicial system, the more robust definition of the maximum acceptable alpha error might therefore be

86. See Rubinfeld & Sappington, *supra* note 72, at 314 (“[T]he socially efficient magnitudes of [two policy instruments, the standard of proof and the penalty for conviction,] are very sensitive to the exact relationship between the litigation efforts of defendants and the judge’s ultimate assessment of a defendant’s guilt (i.e., the signalling technology).”).

87. See Dominique Demougine & Claude Fluet, *Preponderance of Evidence*, 50 EUR. ECON. REV. 963, 965 (2006) (finding that “incentives to exert care are maximized if negligence is established on the basis of a more-likely-than-not” decision rule); Demougine & Fluet, *Deterrence*, *supra* note 73, at 203 (suggesting that “providing incentives [for defendants] better captures the characteristics of evidentiary rules and standard of proof in the common law”).

88. See Rubinfeld & Sappington, *supra* note 72, at 314 (suggesting that “an increase in court expenditure will [possibly] increase the ratio of expenditures by innocent defendants to expenditures by guilty defendants” and that “a decrease in court expenditure . . . will [possibly] increase spending by innocent defendants relative to that by guilty defendants”).

89. See Miceli, *supra* note 61, at 190 (expressing concern that Rubinfeld and Sappington did not consider litigation expenditures by prosecutors).

90. See Chris William Sanchirico, *The Burden of Proof in Civil Litigation: A Simple Model of Mechanism Design*, 17 INT’L REV. L. & ECON. 431, 440 (1997) (suggesting that courts can, to a certain degree, get around the disadvantage of not knowing “whether a case is worth ‘paying’ to hear until after it has paid to hear it”).

91. Lillquist, *supra* note 73, at 184–85.

92. Davis, *supra* note 64, at 347.

93. See *id.* at 350 (“[S]ome care should be taken in extending the [optimal standard of proof] model into areas where there is not a clear and [objective] notion of the legal standard.”).

94. Franklin, *supra* note 11, at 164.

preferable. At any rate, court practice on both sides of the Atlantic seems to be in line with this more robust definition.

III. A BEHAVIORALLY INFORMED PERSPECTIVE

The previous Part on utilitarian norms informs those who design the law of evidence. In the perspective advocated here, the standard of proof has to make sure that the probability of alpha errors is kept below society's level of tolerance. Achieving this goal is an empirical matter. Ultimately, whether courts make materially wrong decisions depends on what judges and jury members do. A behaviorally informed perspective is paramount. There are two competing views of this perspective.

One line of research shows how often courts have indeed erred. If measured against the standards of rational-choice decision theory, judges and jurors perform poorly. The competing line of research starts from a different definition of the task. Judicial decision-making is not applied science. In limited time and, more importantly, on a patently incomplete factual basis, courts have to decide on overly complex issues. Empirically, jurors and judges decide by constructing stories. They decide in favor of the plaintiff, or the prosecution for that matter, if it has been able to induce a coherent story that resonates with the jury members' world knowledge. The underlying psychological mechanism strives at making sense of the evidence. It aims at generating coherence. This is an unconscious activity. From this perspective, a second function of the jury instructions, and of professional legal training for the judges, becomes visible. By stressing the undesirability of alpha errors and by holding jury members accountable, the unconscious generation of coherence is altered. The intervention has clout since it musters emotions. In so doing it creates a somatic marker.

A. *Violations of the Rational-Choice Benchmark*

“Statistics, Not Experts”!⁹⁵ “Train Our Jurors”!⁹⁶ Among behavioral researchers there is not much trust in either legal experts, i.e., judges, nor in laymen on jury duty. If one compares judicial performance to the rational-choice benchmark, there is indeed reason for skepticism. Statistically, untrained subjects have been demonstrated to get nearly everything wrong

95. See William Meadow & Cass R. Sunstein, *Statistics, Not Experts*, 51 DUKE L.J. 629, 631 (2001) (proposing that the legal system rely on statistics rather than on expert testimony).

96. See Koehler, *supra* note 15, at 304 (suggesting that jurors are “legally and logically unprepared for trial”).

that should matter according to statistical theory.⁹⁷ Since statistics is the most sophisticated approach to decision-making under uncertainty, this is no minor concern. Moreover, even though professional judges only marginally outperform laypeople on such tasks as assessing the credibility of an eyewitness, they are massively more confident in their judgment.⁹⁸ These concerns resonate with colorful remarks by Benjamin Cardozo:

The great tides and currents which engulfed the rest of men do not turn aside in their course and pass the judges by.⁹⁹

. . . .
 . . . The spirit of the age, as it is revealed to each of us, is too often only the spirit of the group in which the accidents of birth or education or occupation or fellowship have given us a place. No effort or revolution of the mind will overthrow utterly and at all times the empire of these subconscious loyalties.¹⁰⁰

. . . .
 . . . Deep below consciousness are other forces, the likes and the dislikes, the predilections and the prejudices, the complex of instincts and emotions and habits and convictions, which make the man, whether he be litigant or judge.¹⁰¹

Specifically, jury instructions have been shown to do a rather poor job. To many jurors, “[j]ury instructions are like foreign movies without subtitles.”¹⁰² Verbal instructions have misled subjects in mock juries.¹⁰³ Estimates of what “beyond a reasonable doubt” meant ranged from 51% to 92% of certainty.¹⁰⁴ A recent study even produced a range between 30% and 100%.¹⁰⁵ Further studies show high variance in these estimates.¹⁰⁶

97. John Conlisk, *Why Bounded Rationality?*, 34 J. ECON. LITERATURE 669, 670 (1996).

98. Siegfried Ludwig Sporer, *Evaluating Eyewitness Testimony: The Fallacies of Intuition*, in THE IMPACT OF COURT PROCEDURE ON THE PSYCHOLOGY OF JUDICIAL DECISION MAKING 111, 136 (Christoph Engel & Fritz Strack eds., 2007).

99. BENJAMIN N. CARDOZO, THE NATURE OF THE JUDICIAL PROCESS 168 (1921).

100. *Id.* at 174–75.

101. *Id.* at 167.

102. Patricia M. Wald, *Guilt Beyond a Reasonable Doubt: A Norm Gives Way to Numbers*, 3 U. CHI. LEGAL F. 101, 111 (1993) (quoting *Symposium Generates Ideas for Jury Communication*, 6 CRIM. PRAC. MANUAL (BNA) 207, 208 (Apr. 29, 1992)).

103. See, e.g., Dorothy K. Kagehiro & W. Clark Stanton, *Legal vs. Quantified Definitions of Standard of Proof*, 9 LAW & HUMAN BEHAV. 159, 162, 164 (1985) (explaining results of an experiment with a mock jury in which “subjects perceived more evidence to favor the plaintiffs when they received quantified definitions . . . than when they received legal definitions”).

104. Reid Hastie, *Algebraic Models of Juror Decision Processes*, in INSIDE THE JUROR: THE PSYCHOLOGY OF JUROR DECISION MAKING 84, 106 (Reid Hastie ed., 1993).

105. Saunders, *supra* note 56, at 5.

106. See, e.g., Irwin A. Horowitz, *Reasonable Doubt Instructions: Commonsense Justice and*

Other studies have demonstrated that subjects also were confused on other elements of the instructions.¹⁰⁷

B. A Proper Definition of the Task

This is certainly troubling news. It is understandable why the literary voices for quantifying the standard of proof are so strong in the United States.¹⁰⁸ The critics of the Continental approach claim that the European quest for truth is futile and obnoxious because it opens the door to mistake, randomness, prejudice, and outright abuse.¹⁰⁹ They counsel turning judicial procedure into an exercise in objectivity.¹¹⁰ Legal decisions should be made as science-like as possible.

Standard of Proof, 3 PSYCHOL. PUB. POL'Y & L. 285, 293 (1997) (discussing a study that notes a variance from 79% to 92%); Irwin A. Horowitz & Laird C. Kirkpatrick, *A Concept in Search of a Definition: The Effects of Reasonable Doubt Instructions on Certainty of Guilt Standards and Jury Verdicts*, 20 LAW & HUM. BEHAV. 655, 664 (1996) (demonstrating that mock jurors' reasonable-doubt ratings ranged from 49.75% to 81.87%, depending upon the strength of the evidence and reasonable doubt definitions pre- and post-deliberations); Stoffelmayr & Diamond, *supra* note 57, at 777 (noting that although "jurors understand that the threshold for reasonable doubt is high," they may vary in their interpretations, resulting in "substantial inter-juror variability"); Tillers & Gottfried, *supra* note 13, at 156 (quoting *United States v. Hall*, 854 F.2d 1036, 1044 (7th Cir. 1988) (Posner, J., concurring)) (citing figures as low as 50% and 76%).

107. See Bradley Saxton, *How Well Do Jurors Understand Jury Instructions? A Field Test Using Real Juries and Real Trials in Wyoming*, 33 LAND & WATER L. REV. 59, 86 (1998) ("[T]he jurors' responses to the substantive questions revealed that many jurors had not understood the instructions as well as they thought they had."); Richard L. Wiener, Christine C. Pritchard & Minda Weston, *Comprehensibility of Approved Jury Instructions in Capital Murder Cases*, 80 J. APPLIED PSYCHOL. 455, 458-59, 463 (1995) (showing high rates of juror miscomprehension in the areas of jury responsibility, sentence process, mitigation agreement, and mitigation content).

108. Saunders, *supra* note 56, at 1; Rita James Simon, *"Beyond a Reasonable Doubt"—An Experimental Attempt at Quantification*, 9 J. APPLIED BEHAV. SCI. 203, 206-07 (1970); Rita James Simon & Linda Mahan, *Quantifying Burdens of Proof: A View from the Bench, the Jury, and the Classroom*, 5 LAW & SOC'Y REV. 319, 319 (1971); Tillers & Gottfried, *supra* note 13, at 135; Weinstein & Dewsbury, *supra* note 55, at 173.

109. See Clermont & Sherwin, *supra* note 26, at 263, 267, 271 (purporting that the Civil Law's high standard of proof leads to unfavorable procedural consequences, uncertainty in knowledge of disputed past events, mischaracterization of evidence, a failure to facilitate dispute resolution, and "makes it difficult for plaintiffs to succeed"). Interestingly, one of these authors, in an earlier paper, and based on psychological grounds, had argued that standards of proof should be expressed on the "customary seven categories of uncertainty," i.e., on the scale of "(1) slightest possibility, (2) reasonable possibility, (3) substantial possibility, (4) equipoise, (5) probability, (6) high probability, and (7) almost certainty." Kevin M. Clermont, *Procedure's Magical Number Three: Psychological Bases for Standards of Decision*, 72 CORNELL L. REV. 1115, 1123, 1147 (1987).

110. See, e.g., Burkhard Schafer & Olav K. Wiegand, *Incompetent, Prejudiced and Lawless? A Gestalt-Psychological Perspective on Fact Finding in Law as Learning*, 3 LAW, PROBABILITY & RISK 93, 94 (2004) (discussing objectivity in the law of evidence).

Critics object that judicial decision-making is an extremely complex task.¹¹¹ Judges and jurors are supposed to handle vast amounts of information in limited time.¹¹² They are only allowed to pass judgment once they have heard all of the evidence. Pieces of evidence are not additively separable.¹¹³ In the language of probability theory, conditional probabilities matter.¹¹⁴ The evidence is presented in a scrambled order. It normally remains incomplete.¹¹⁵ Courts are aware of strategic attempts by the parties to mislead them.

C. Sense-Making

The most powerful criticism of the objectivist definition of the task is, however, empirical. Judges and jury members simply do not act like miniature scientists. This is not a defect, but the best they can do, given the patently incomplete evidence that is characteristic for almost all court cases. Empirically, judges and jury members are engaged in “sense-making.”¹¹⁶ Decision-making is explanation based.¹¹⁷ It is interpretative.¹¹⁸ It relies on reasoning about the evidence, rather than an algebra-like process.¹¹⁹ Jurors attempt to create a narrative story from the pieces of evidence they have heard.¹²⁰ Decision-making is based on the construction of “mental models.”¹²¹

111. Nancy Pennington & Reid Hastie, *Explaining the Evidence: Tests of the Story Model for Juror Decision Making*, 62 J. PERSONALITY & SOC. PSYCHOL. 189, 189 (1992) [hereinafter Pennington & Hastie, *Explaining the Evidence*].

112. Nancy Pennington & Reid Hastie, *Evidence Evaluation in Complex Decision Making*, 51 J. PERSONALITY & SOC. PSYCHOL. 242, 242 (1986) [hereinafter Pennington & Hastie, *Evidence Evaluation*]; Tillers & Gottfried, *supra* note 13, at 155.

113. Pennington & Hastie, *Explaining the Evidence*, *supra* note 111, at 190.

114. See Miceli, *supra* note 61, at 189–90 (“By the appropriate choice of the variables of the legal process—namely, prosecutorial effort in collecting evidence, the standard of proof employed by courts, and the punishment for convicted criminals—the likelihood and impact of each of these outcomes can be affected.”).

115. See Pennington & Hastie, *Evidence Evaluation*, *supra* note 112, at 243 (stating that “inferences about missing or unstated intermediate conclusions are frequent and determinative”).

116. See KARL E. WEICK, *SENSEMAKING IN ORGANIZATIONS* 4 (1995) (describing “sensemaking” as the ability to “structure the unknown”); Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 519, 519 (1991) (concluding that jurors are “sense-making information processor[s]”).

117. Pennington & Hastie, *Explaining the Evidence*, *supra* note 111, at 189; Nancy Pennington & Reid Hastie, *Explanation-Based Decision Making: Effects of Memory Structure on Judgment*, 14 J. EXPERIMENTAL PSYCHOL.: LEARNING MEMORY & COGNITION 521, 521 (1988) [hereinafter Pennington & Hastie, *Explanation-Based Decision Making*]; Pennington & Hastie, *Evidence Evaluation*, *supra* note 112, at 242.

118. Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 524.

119. *Id.* at 531.

120. See Nancy Pennington & Reid Hastie, *Reasoning in Explanation-Based Decision Making*, 49 COGNITION 123, 136 (1993) [hereinafter Pennington & Hastie, *Reasoning*] (“[J]urors construct

Decision-making is an inherently constructive activity.¹²² The constructive element is particularly important, since the evidence is not presented in story-like form.¹²³ In this mental activity, jurors not only build on the evidence, they also combine it with their world knowledge¹²⁴ and with their “expectations about what constitutes an adequate explanation” in the respective area of life.¹²⁵ In so doing, they look out for causal and intentional relations.¹²⁶ Due to this memory component, by necessity, “[d]ifferent jurors will construct different stories.”¹²⁷

If the case is not utterly simple, the ultimate narrative will be composed of episodes that themselves have story form.¹²⁸ Each episode consists of initiating events, goals, actions, consequences, and accompanying states, in a particular causal configuration.¹²⁹ Typically, the result is a hierarchy of embedded episodes.¹³⁰ It is possible that jurors try out competing stories.¹³¹

Story construction is instantaneous. It starts early on, when the jurors hear the first pieces of evidence.¹³² The fact that jurors have to defer explicit judgment until they have heard the entire evidence fosters story

narrative representations of evidence spontaneously in the course of reaching a decision”); Nancy Pennington & Reid Hastie, *The Story Model for Juror Decision Making*, in *INSIDE THE JUROR: THE PSYCHOLOGY OF JUROR DECISION MAKING* 192, 194 (Reid Hastie ed., 1993) [hereinafter Pennington & Hastie, *The Story Model*] (“The story model is based on the hypothesis that jurors *impose* a narrative story organization on trial information.”); Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 521 (discussing how “the decision maker constructs a causal explanation of the evidence by combining evidence with related world knowledge and expectations about what constitutes an adequate explanation in the decision domain”); Pennington & Hastie, *Evidence Evaluation*, *supra* note 112, at 243 (describing the “Story Model” of evidence evaluation, which “is based on the hypothesis that jurors impose a narrative story organization on trial information”); *see also* WEICK, *supra* note 116, at 10–11 (suggesting that jurors create a narrative story to “fit” a predetermined outcome).

121. Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 521; *see also* PHILIP N. JOHNSON-LAIRD, *MENTAL MODELS: TOWARDS A COGNITIVE SCIENCE OF LANGUAGE, INFERENCE, AND CONSCIOUSNESS* 1–2 (1983) (explaining that people use mental models to understand and to draw inferences).

122. Pennington & Hastie, *The Story Model*, *supra* note 120, at 194.

123. *Id.* at 195; Pennington & Hastie, *Evidence Evaluation*, *supra* note 112, at 243.

124. Pennington & Hastie, *The Story Model*, *supra* note 120, at 195.

125. Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 521.

126. Pennington & Hastie, *Evidence Evaluation*, *supra* note 112, at 243.

127. Pennington & Hastie, *The Story Model*, *supra* note 120, at 196.

128. *Id.* at 197.

129. Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 522.

130. Pennington & Hastie, *Evidence Evaluation*, *supra* note 112, at 243–44.

131. Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 522.

132. *See* REID HASTIE, STEVEN D. PENROD & NANCY PENNINGTON, *INSIDE THE JURY* 18 (1983) (suggesting that a juror’s first task in making a decision involves “encoding trial contents”); *see also* Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 521, 524 (explaining how “when evidence is presented, the subject constructs a verbatim representation of the surface structure of the evidence . . . , a semantic representation of the evidence in the form of a propositional textbase, and a situation model that represents an interpretation of what the evidence is about”).

construction.¹³³ The order in which the evidence is presented matters. Stories that fit better to the pieces of evidence heard initially are more likely to persist in the jurors' minds.¹³⁴

Jurors decide by matching stories to the representation of the verdict categories given to them in the judge's instructions on the law.¹³⁵ Whether a story is accepted, or whether it is selected, depends on its "goodness-of-fit" with the verdict category.¹³⁶ In this assessment, jurors look out for coverage, coherence, and uniqueness.¹³⁷ They check "the extent to which the story accounts for evidence presented at trial."¹³⁸ They require the story to be coherent, meaning that it must be consistent, plausible, and complete.¹³⁹ Consistency is an internal criterion, applying the laws of logic to the elements of the story.¹⁴⁰ Plausibility is an attempt at matching the story with the juror's world knowledge.¹⁴¹ A story is complete if it has all the elements a story of this kind should have.¹⁴² It lacks completeness if it has conspicuous gaps.¹⁴³ Finally, jurors check for uniqueness. They are more likely to accept a story if they are unable to come up with a plausible alternative story based on the same evidence.¹⁴⁴ If there are competing stories, jurors prefer the one that is more coherent.¹⁴⁵

Coverage, coherence, and uniqueness also determine the level of confidence.¹⁴⁶ Subjects are more confident the more they feel they have a clear view of the case. In line with this correlation, confidence is higher if both the prosecution and the defense have presented their evidence in story order.¹⁴⁷

133. Pennington & Hastie, *Explaining the Evidence*, *supra* note 111, at 189.

134. See Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 522 ("More than one story may be constructed by the juror, but one story will usually be accepted as more coherent than the others.").

135. See Pennington & Hastie, *Evidence Evaluation*, *supra* note 112, at 243 (explaining how jurors reason about verdict categories).

136. Pennington & Hastie, *The Story Model*, *supra* note 120, at 201.

137. Pennington & Hastie, *Explaining the Evidence*, *supra* note 111, at 190.

138. *Id.*

139. Pennington & Hastie, *The Story Model*, *supra* note 120, at 198–99.

140. *Id.* at 199.

141. *Id.*

142. *Id.* Stories have multiple elements: "In stories and in episodes, events considered to be *initiating events* cause characters to have psychological *responses* and to form *goals* that motivate subsequent *actions* which cause certain *consequences* and accompanying *states*." *Id.* at 197.

143. D.H. Kaye, Commentary, *Do We Need a Calculus of Weight to Understand Proof Beyond a Reasonable Doubt?*, 66 B.U. L. REV. 657, 663–72 (1986); Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 522.

144. Pennington & Hastie, *Explaining the Evidence*, *supra* note 111, at 190–91.

145. Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 522.

146. Pennington & Hastie, *Explaining the Evidence*, *supra* note 111, at 190–91; Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 521, 528.

147. Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 530.

D. Consistency Maximization

In the legal literature, many have noted that a trier of fact has to rely on her intuition.¹⁴⁸ Some have even linked this to Gestalt psychology.¹⁴⁹ It has been said that, to convict the defendant, the fact-finder must be in the “subjective state of near certitude.”¹⁵⁰

The underlying mental process is relatively well understood. In line with the basic claim from Gestalt psychology,¹⁵¹ the assessment of the evidence is holistic.¹⁵² Decision-makers aim at forging coherence.¹⁵³ They strive at parallel constraint satisfaction.¹⁵⁴ They treat the evidence somehow like the pieces of a jigsaw puzzle, and try to fit as many pieces as possible, with the twist that they are happy to go with a less-than-perfect fit of the elements, as long as the resulting picture looks good. Based on a connectionist cognitive architecture, decision-making progresses bidirectionally.¹⁵⁵ Not only do facts determine conclusions, potential conclusions also affect the perception of the evidence. The mental model reconfigures itself until maximal coherence is achieved.¹⁵⁶

The process has been modeled mathematically. Based on the model, and on assumptions about the parameters, simulations have been run. The results of these simulations have been compared to experimental findings, with a fairly good fit.

The basic components of the process are these: First, the evidence activates cues.¹⁵⁷ These cues are positively or negatively related to decision options.¹⁵⁸ If options exclude each other, they are negatively related.¹⁵⁹ As a

148. Zivilprozeßordnung [ZPO] [civil procedure statute] Dec. 5, 2005, BGBl. I 3202, § 286(1) (F.R.G.), translated in STEFAN RÜTZEL, GERHARD WEGEN, & STEPHAN WILSKE, COMMERCIAL DISPUTE RESOLUTION IN GERMANY 224 (2005) (giving the court discretion in regarding the facts on the record as true or false); MUSIELAK, *supra* note 47, at 872; SCHULZ, *supra* note 48, at 42, 168.

149. Schafer & Wiegand, *supra* note 110, at 95.

150. Jackson v. Virginia, 443 U.S. 307, 315 (1979).

151. Hazel Markus & R.B. Zajonc, *The Cognitive Perspective in Social Psychology*, in 1 THE HANDBOOK OF SOCIAL PSYCHOLOGY 137, 139 (Gardner Lindzey & Elliot Aronson eds., 3d ed. 1985).

152. Simon, *supra* note 17, at 560, 562.

153. Dan Simon, Lien B. Pham, Quang A. Le & Keith J. Holyoak, *The Emergence of Coherence Over the Course of Decision Making*, 27 J. EXPERIMENTAL PSYCHOL.: LEARNING MEMORY & COGNITION 1250, 1257 (2001) [hereinafter Simon et al., *The Emergence of Coherence*].

154. Dan Simon & Keith J. Holyoak, *Structural Dynamics of Cognition: From Consistency Theories to Constraint Satisfaction*, 6 PERSONALITY & SOC. PSYCHOL. REV. 283, 285 (2002).

155. See Keith J. Holyoak & Dan Simon, *Bidirectional Reasoning in Decision Making by Constraint Satisfaction*, 128 J. EXPERIMENTAL PSYCHOL.: GENERAL 3, 3 (1999) (“Bidirectional inferences are inherent in the operation of models of thinking that are based on parallel constraint satisfaction.”).

156. Simon, *supra* note 17, at 522.

157. *Id.*

158. *Id.*

159. *Id.*

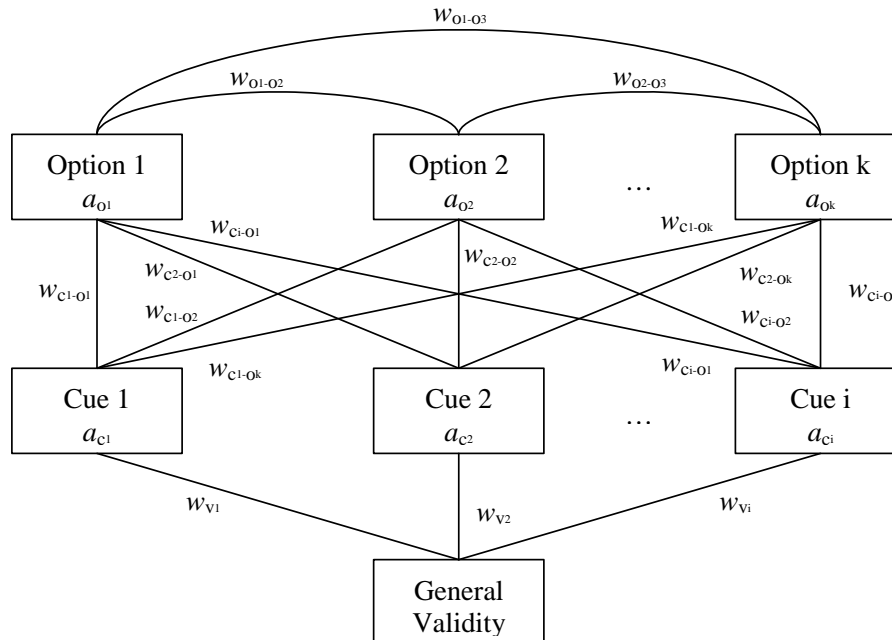


Figure 1
Model of Consistency Maximization¹⁶⁰

first step, the evidence provides support for at least some of the options.¹⁶¹ But the process does not stop here. Depending on how strongly an option is activated initially, it propels positive activation back to the supporting evidence, and negative activation to the conflicting evidence and to the competing options.¹⁶² Based on the resulting reassessment of the evidence, the process enters the next iteration.¹⁶³ It is repeated as long as the marginal changes in the activation of the options are substantial.¹⁶⁴ The most simple decision criterion is the difference in the final activation of options.¹⁶⁵ The option with the highest positive activation is chosen,

160. Andreas Glöckner & Tilmann Betsch, *Modeling Option and Strategy Choices with Connectionist Networks: Towards an Integrative Model of Automatic and Deliberate Decision Making*, 3 JUDGMENT & DECISION MAKING 215, 218 fig.1 (2008).

161. *Id.*

162. *Id.* at 218–19.

163. *Id.* at 219.

164. Andreas Glöckner, *How Evolution Outwits Bounded Rationality: The Efficient Interaction of Automatic and Deliberate Processes in Decision Making and Implications for Institutions*, in BETTER THAN CONSCIOUS? 259, 265 (Christoph Engel & Wolf Singer eds., 2008).

165. Glöckner & Betsch, *supra* note 160, at 219.

provided all competing options are sufficiently less activated.¹⁶⁶ A more elaborate decision criterion also takes the structure of the network into account that supports the decision.¹⁶⁷ It also checks for the overall consistency of this network.¹⁶⁸

For the decision-maker, this mechanism has a number of advantages. First and foremost, the mechanism is able to handle vast amounts of information.¹⁶⁹ The severe limitations for the deliberate handling of information are not present. As is well known, the memory span for deliberate processing is limited to seven items plus or minus two in most people.¹⁷⁰ Even more importantly, parallel constraint satisfaction enables the individual to decide although the problem is overly complicated,¹⁷¹ or known to be ill-defined.¹⁷² The process makes the best out of partial evidence. In gradually accentuating the differences between options, the individual is much more often in a position to decide, although not always with full certainty. The mechanism is therefore best interpreted as an enabling device.¹⁷³ It empowers the individual to make relatively good decisions although there is no chance to fully understand the problem.¹⁷⁴

These advantages come at a price.¹⁷⁵ Intuition is not a calculator. The mechanism is programmed to transform the information input. Over the consecutive iterations, information supporting the final decision is overvalued.¹⁷⁶ Conflicting evidence is undervalued.¹⁷⁷ Information is thus polarized.¹⁷⁸ This process has been dubbed a coherence shift.¹⁷⁹ It has been

166. *Id.* at 218.

167. *Id.* at 217–18.

168. *Id.* at 217.

169. See PAUL W. GLIMCHER, DECISIONS, UNCERTAINTY, AND THE BRAIN: THE SCIENCE OF NEUROECONOMICS 256 (2003) (“[A]ny rational decision-making system must encode the likelihoods of all possible outcomes . . .”); Christoph Engel, *Institutions for Intuitive Man*, in BETTER THAN CONSCIOUS 391, 399 (Christoph Engel & Wolf Singer eds., 2008) [hereinafter Engel, *Intuitive Man*] (“[T]he nonconscious mental apparatus is able to handle a huge amount of information in very little time . . .”); Simon, *supra* note 17, at 513 (explaining how the brain organizes vast amounts of information into difficult and easy tasks).

170. ALAN D. BADDELEY, WORKING MEMORY 4–5 (1986).

171. Simon, *supra* note 17, at 517.

172. Engel, *Intuitive Man*, *supra* note 169, at 401. For a description of these two categories of tasks, see GERD GIGERENZER, PETER M. TODD & ABC RESEARCH GROUP, SIMPLE HEURISTICS THAT MAKE US SMART 136–37, 139–40, 211 (2003).

173. *But cf.* Simon, *supra* note 17, at 547 (“At least in some conditions, then, the coherence can be understood as an ad hoc state that is constructed only to solve the task at hand.”).

174. Engel, *Intuitive Man*, *supra* note 169, at 401.

175. *Id.*

176. Simon, *supra* note 17, at 522–23.

177. *Id.* at 523.

178. *Id.* at 522–23.

179. Simon et al., *The Emergence of Coherence*, *supra* note 153, 1257–58.

shown that these shifts are pronounced.¹⁸⁰ The higher the confidence of the decision-maker, the stronger the reevaluation of evidence.¹⁸¹ Since the intuition is formed in the automatic system, decision-makers are not aware of the coherence shift.¹⁸² Ultimately, coherence shifts fade away.¹⁸³ There is, however, the possibility that a shift resulting from one decision spills over to another decision in the near future.¹⁸⁴

These general findings have been shown to hold in the assessment of evidence by the members of mock juries.¹⁸⁵ Individual predilections are able to influence the jurors' judgment via the automatic system.¹⁸⁶ The effects are even present if jury members are admonished to withhold judgment until they have heard the entire evidence.¹⁸⁷

E. Rescued by Emotions?

Continental observers have said that the subjectivist model of proof is built on "ethos, experience, and intuition."¹⁸⁸ Apparently, the civilian legal orders do not blindly trust intuition. They are (at least intuitively) sensitive to the risk that judicial intuition is misled by clever parties, by prejudice, or simply by heedless decision-making. Yet the response to this risk is not an attempt at turning judicial procedure into a quasi-scientific affair. Rather, these legal orders have recourse to yet another determinant of human behavior that is not under direct conscious control. They appeal to emotions. Specifically they call on judges' professional ethos and on jurors' sense of republican duty. If this intervention is to have the desired effect, two conditions must be met: it must be possible to calibrate the subconscious mechanism of constraint maximization, and the institutional interventions that make judicial responsibility salient must be sufficiently powerful to bring about the normatively expected level of scrutiny.

180. Simon, *supra* note 17, at 532.

181. Dan Simon, Chadwick J. Snow & Stephen J. Read, *The Redux of Cognitive Consistency Theories: Evidence Judgments by Constraint Satisfaction*, 86 J. PERSONALITY & SOC. PSYCHOL. 814, 819 (2004) [hereinafter Simon et al., *Constraint Satisfaction*].

182. Simon, *supra* note 17, at 545.

183. *See id.* at 542–43 (explaining a finding that, in some instances, "coherence effects are transient").

184. *Id.* at 546.

185. Pennington & Hastie, *Explanation-Based Decision Making*, *supra* note 117, at 521–22 (using empirical evidence to explain mock jurors' assessment of evidence in criminal trials); *see also* Pennington & Hastie, *Reasoning*, *supra* note 120, at 126–27 (reiterating the story model used by mock juries to assess evidence).

186. *See* Simon, *supra* note 17, at 550 (noting that research shows that jury "verdicts tend to correspond closely to predeliberation individual preferences").

187. *See id.* at 542, 553 (explaining that research has shown that people with pro-death-penalty attitudes are more likely to convict a defendant than people with anti-death-penalty attitudes).

188. SCHULZ, *supra* note 48, at 168; *see* Taruffo, *supra* note 32, at 667 (describing a "subjective 'intimate' persuasion" based on individual experience and emotional beliefs).

For the purposes of this Article, the most relevant subquestion is this: Are standards of proof moot if triers of fact rely on their automatic system? One might be afraid that this would happen. After all, it is the function of parallel constraint satisfaction to give individuals more scope for decision-making under uncertainty. In the extreme, this could mean that the mind always comes down on one side. This would be at variance with the two cornerstones of the law of evidence—the standard of proof and the burden of proof. The law wants two things: the court shall only take a fact for true if the requisite standard of proof has been met; if the evidence remains inconclusive, the court shall decide as defined by the burden of proof. Behaviorally, this doctrinal model presupposes that constraint satisfaction must be able to fail. In the definition of success versus failure, individuals must be able to apply different, exogenously imposed standards of strictness.

The psychological correlate of the standard of proof is confidence. It has been shown that reported confidence levels do not mirror the degree to which the evidence has remained unaltered in the process of forming an intuition.¹⁸⁹ On the contrary, subjects that exhibit particularly strong coherence shifts are also most confident of their final judgment.¹⁹⁰ The most troubling piece of evidence is illustrated through an experiment with two stages. In the first stage, subjects were asked to rate the credibility of unrelated vignettes.¹⁹¹ In the second stage, material from these vignettes was presented as evidence in a mock jury trial.¹⁹² The trial was framed as a criminal case.¹⁹³ Subjects were reminded that the respective standard of proof was beyond a reasonable doubt.¹⁹⁴ In line with the general parallel-constraint-satisfaction model, subjects inflated the relevance of evidence that supported their final decision, and they deflated conflicting evidence.¹⁹⁵ However, those who ultimately convicted the defendant deflated exculpatory evidence much more strongly than acquitters deflated inculpatory evidence.¹⁹⁶ One interpretation of the result is this: the only effect of the stricter standard of proof is to deflate conflicting evidence more strongly.

189. Simon, *supra* note 17, at 532.

190. *Id.*

191. *Id.* at 524.

192. *Id.* at 525.

193. *Id.* at 528–29.

194. *See id.* at 529 (noting that participants “were presented with a task that would be demanded of actual jurors”).

195. *Id.* at 531. “In other words, by the point of decision, the mental models of evidence are skewed toward strong support of the respective verdicts.” *Id.*

196. *Id.* The difference was, however, not statistically significant. Letter from Dan Simon to author (on file with author).

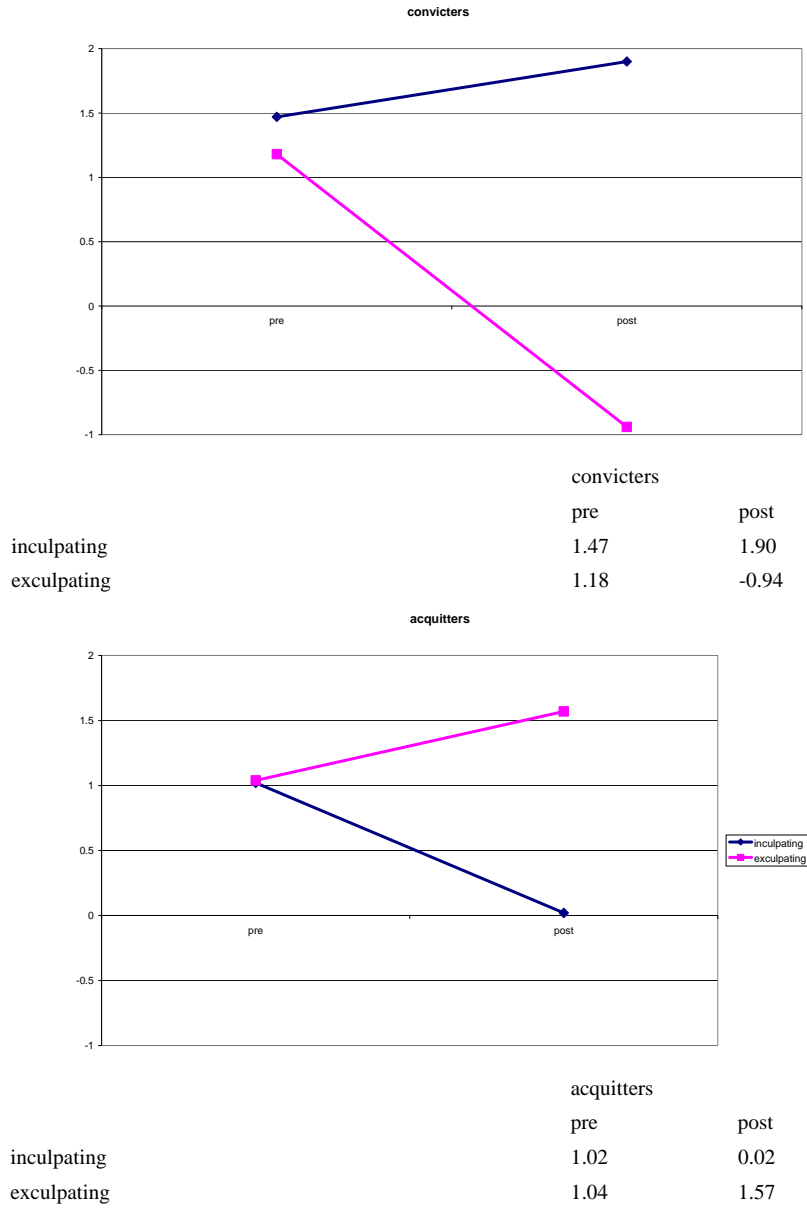


Figure 2
 Dan Simon's Experiment¹⁹⁷

197. Simon, *supra* note 17, at 531 n.58, fig.4. Evidence was rated on a scale from -5 to +5. *Id.* at 531 n.59.

Happily, this interpretation seems to be wrong.¹⁹⁸ In our lab, we have conducted three studies with a total of 238 subjects. In two of the studies, we gave two groups of subjects the same evidence. One group was instructed to convict the defendant only if her guilt was beyond a reasonable doubt. The other group was instructed to decide to the detriment of the defendant if the preponderance of the evidence spoke against her. For these experiments, we used translated jury instructions from the United States. The results are depicted in Figure 3.

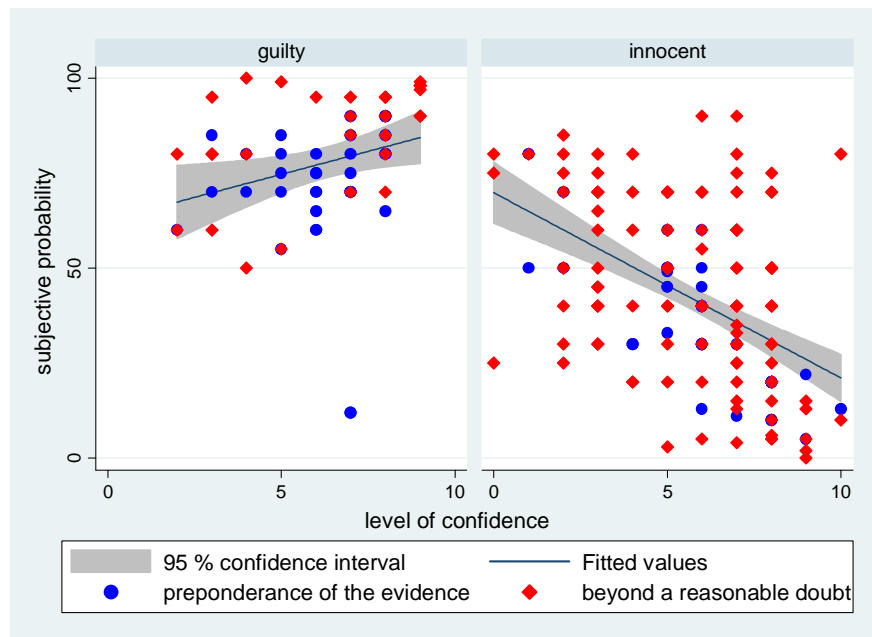


Figure 3
Erfurt Experiment¹⁹⁹

Subjects only convict if their estimate of guilt is above fifty percent.²⁰⁰ Although some convict under the beyond-a-reasonable-doubt standard at a disturbingly low subjective probability of guilt, the majority of diamonds

198. See Norbert L. Kerr et al., *Guilt Beyond a Reasonable Doubt: Effects of Concept Definition and Assigned Decision Rule on the Judgments of Mock Jurors*, 34 J. PERSONALITY & SOC. PSYCHOL. 282, 293 (1976) (showing that standards of proof empirically make a difference); SAUL M. KASSIN & LAWRENCE S. WRIGHTSMAN, *THE AMERICAN JURY ON TRIAL* 156 (1988) (explaining the effects of various standards of proof).

199. Andreas Glöckner & Christoph Engel, *Can We Trust Intuitive Jurors? An Experimental Analysis* 22 fig.6 (Max Planck Inst. for Research on Collective Goods, Preprint No. 2008/36, 2008), available at http://www.coll.mpg.de/pdf_dat/2008_36online.pdf.

200. The one outlier notwithstanding.

are in the normatively expected places—subjects require a higher subjective probability under the stricter standard. If they acquit under the preponderance-of-the-evidence standard, their subjective probability of guilt is not above fifty percent. If they acquit although the subjective probability is above fifty percent, this is always under the beyond-a-reasonable-doubt instruction. Subjects are most confident in their decision if they believe the case to be clear, i.e., if they acquit and subjective probability of guilt is low, or if they convict and subjective probability of guilt is high.

In our experiments, total coherence shifts are indeed more pronounced when subjects convict. However, on closer inspection this effect does not have a normatively troublesome cause. As Figure 4 shows, it is caused by the fact that, under the beyond a reasonable doubt instruction, subjects do not devalue inculcating evidence as strongly as under the preponderance of the evidence instruction.²⁰¹ Or more legally, many subjects acquit “for want of evidence,” and are not induced to devalue conflicting evidence to come up with a decision “for cause.”²⁰² This is exactly what the legal order wants to achieve with the stricter standard of proof.²⁰³

It is relatively easy to explain this finding in light of the storytelling model. When instructed to convict only if guilt is beyond a reasonable doubt, subjects apply a stricter standard of coverage, coherence, and uniqueness. They acquit if minor pieces of evidence do not find a place in the story told by the prosecution. They do not tolerate a shade of doubt regarding the logical consistency of the story. In light of their world knowledge, they require high plausibility for this story. They apply a strict standard of scrutiny when testing the completeness of the prosecution’s story. Finally, they impose a large minimum gap in plausibility if there is more than one story that can be told based on the evidence.

The easiest way to translate these techniques of calibrating judgment into mental mechanisms is by manipulating the threshold. Irrespective of the standard of proof, the automatic system always processes the information as long as this results in more than minimal changes in the activation of options. However, depending on the standard of proof, the mind asks for a larger minimum difference in the activation of the preferred alternative, compared to all competing alternatives. This explanation fits best for the minimum gap in plausibility between competing stories. It also

201. Glöckner & Engel, *supra* note 199, at 18.

202. *Id.* at 23.

203. *Id.* at 22–23. In a regression with the rerating of conflicting evidence as the dependent variable, frame and probability manipulations as controls, and robust standard errors, the regressor for standard of proof is (weakly) significant if subjects convict, $\beta = .18$, $p = .09$. *Id.* at 20.

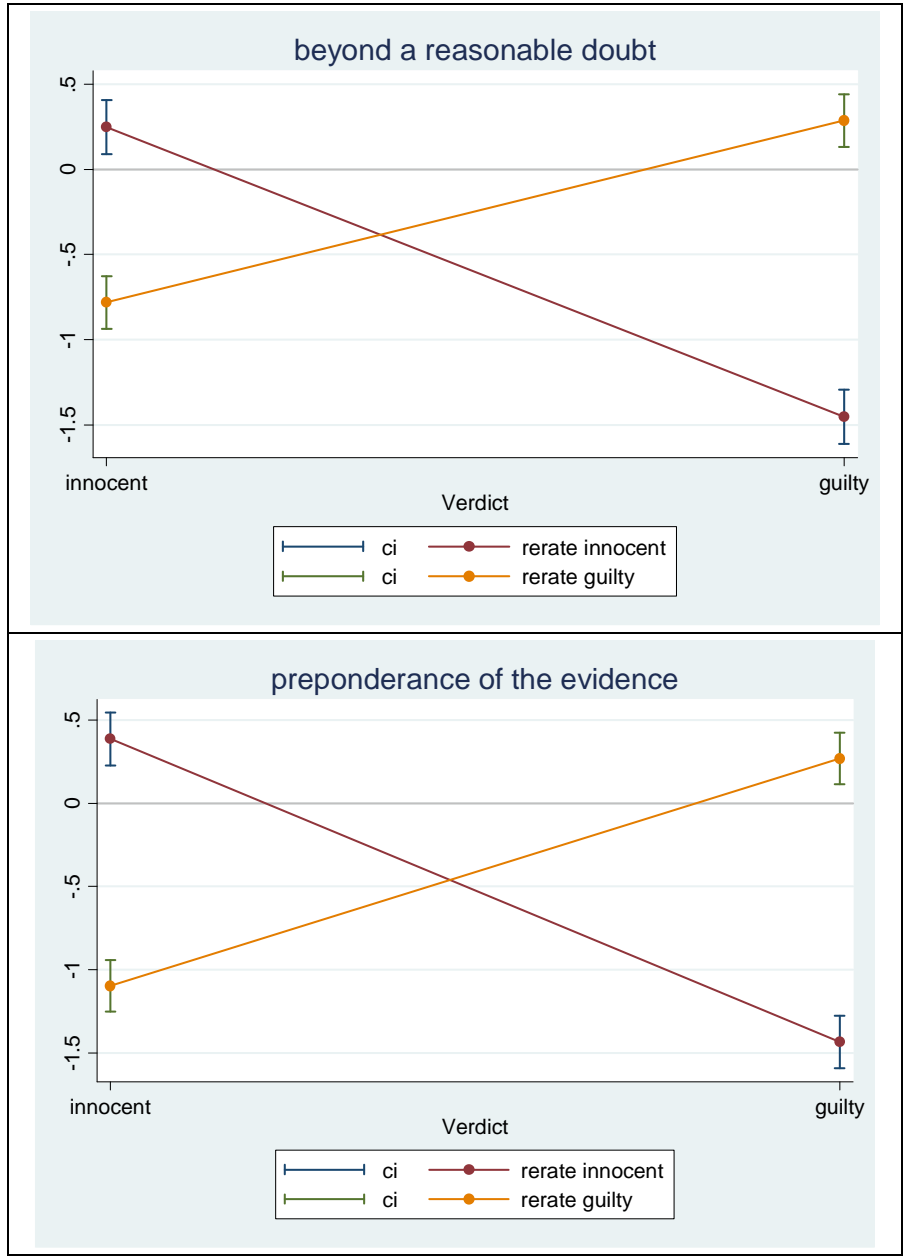


Figure 4
Rerating of Evidence by Standard of Proof²⁰⁴

204. *Id.* at 21 fig.5.

works for the level of plausibility, given world knowledge. For the remaining elements of the story model, one needs a richer decision criterion. This criterion must also take into account the structure of the network that supports the decision, as proposed by the parallel-constraint-satisfaction model.²⁰⁵

Theoretically, decision-makers might split up their mental mechanism. First, they would strive to intuitively make sense of the evidence. Second, they would deliberately assess the result against a benchmark level of confidence.²⁰⁶ To fully disprove this claim, one would need data on decision time. Deliberate reasoning requires exponentially more time than intuition. Since legal cases are inherently complex, measuring decision time is tricky. Yet even without this additional data, such a split mental mechanism is very unlikely. The strongest piece of counterevidence is precisely our finding on the relationship between coherence shifts and standards of proof. If decision-making came later, coherence shifts should be symmetric under the beyond-a-reasonable-doubt instruction. The fact that they are not implies that the standard of proof already influences the intuitive assessment of the facts.

Thus far, it has been shown that different standards of proof are psychologically feasible. But how is the legal order able to induce these standards? What does “ethos” mean in psychological terms? By which mental mechanism are the thresholds for discriminating options, and for assessing the overall consistency of the supporting network, calibrated? What is the psychological effect of the many institutional interventions that make judicial responsibility salient? Only a very small number of lawyers are chosen to be judges. By the profession, being a judge is regarded as an honor. Judges are appointed in a ceremonial way. Throughout judicial procedure, all are reminded repeatedly that the judge holds an office of the people and speaks in the people’s name. Similarly, what is the psychological effect of being selected as a jury member, and of the jury instructions in particular?

Psychologically speaking, these interventions make accountability salient.²⁰⁷ Accountability induces subjects to invest more cognitive effort²⁰⁸

205. Simon & Holyoak, *supra* note 154, at 285.

206. Kevin Clermont has suggested this interpretation in private communication.

207. Christoph Engel, *The Psychological Case for Obliging Judges to Write Reasons*, in *THE IMPACT OF COURT PROCEDURE ON THE PSYCHOLOGY OF JUDICIAL DECISION MAKING* 73, 88–89, 92 (Christoph Engel & Fritz Strack eds., 2007) [hereinafter Engel, *The Psychological Case*].

208. Daniel W. McAllister, Terence R. Mitchell & Lee R. Beach, *The Contingency Model for the Selection of Decision Strategies: An Empirical Test of the Effects of Significance, Accountability, and Reversibility*, 24 *ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE* 228, 243 (1979); Philip E. Tetlock, Linda Skitka & Richard Boettger, *Social and Cognitive Strategies for Coping with*

and to become open to greater complexity of thought.²⁰⁹ Specifically, the subjectivist standard of proof in Continental law directly appeals to personal responsibility. “Only convict if you are sure you can take on responsibility for this decision!” The wording of jury instructions for the reasonable-doubt standard in the United States sounds pretty similar to this.²¹⁰ It seems that, in judicial practice, the objectivist norm that wants to minimize the alpha error is translated into a requirement with a pronounced subjectivist flavor. By contrast, the preponderance-of-the-evidence instructions can be interpreted as a tool for exonerating jury members from personal responsibility. Society is happy with quite a number of materially wrong judgments. Accountability is reduced to avoiding gross errors.

Ultimately, jury instructions are just words. Society has no chance to check, in the individual instance, whether these words have had their desired effects. There is no *ex post facto* control. There are no sanctions. Why is one nonetheless justified in believing that the instructions do their job? Arguably, at this point emotions come into play. The instructions, together with the described institutional arrangement, set a somatic marker.²¹¹ Somatic markers are particularly instrumental if decisions have to be taken under conditions of high complexity and uncertainty.²¹² They reduce options and focus attention.²¹³ They work as a biasing device.²¹⁴ They induce the individual to take a risk very seriously.²¹⁵ They are particularly instrumental if low-probability events are paired with highly affective reactions.²¹⁶ Much the same way, the jury instructions set a

Accountability: Conformity, Complexity, and Bolstering, 57 J. PERSONALITY & SOC. PSYCHOL. 632, 633 (1989); Elizabeth Weldon & Gina M. Gargano, *Cognitive Loafing: The Effects of Accountability and Shared Responsibility on Cognitive Effort*, 14 PERSONALITY & SOC. PSYCHOL. BULL. 159, 160 (1988). See also Robert H. Ashton, *Effects of Justification and a Mechanical Aid on Judgment Performance*, 52 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 292, 301 (1992) (noting that requiring justifications of one’s actions is “associated with a significant improvement in accuracy that can be traced in part to a significant improvement in consistency”).

209. Philip E. Tetlock & Jae Il Kim, *Accountability and Judgment Processes in a Personality Prediction Task*, 52 J. PERSONALITY & SOC. PSYCHOL. 700, 701 (1987).

210. See *infra* Part I (noting the “subjectivist flavor” of many criminal law jury instructions).

211. DAMASIO, *supra* note 20, at 173; see also Antoine Bechara & Antonio R. Damasio, *The Somatic Marker Hypothesis: A Neural Theory of Economic Decisions*, 52 GAMES & ECON. BEHAV. 336, 339 (2005) (describing the somatic marker hypothesis as rooted in the influence of emotions on the human decision-making process).

212. DAMASIO, *supra* note 20, at 174–75.

213. *Id.* at 173.

214. *Id.* at 174, 198.

215. See *id.* at 173 (“[The somatic marker] forces the attention on the negative outcome to which a given action may lead, and functions as an automated alarm signal . . .”).

216. George F. Loewenstein, Elke U. Weber, Christopher K. Hsee & Ned Welch, *Risk as Feelings*, 127 PSYCHOL. BULL. 267, 273, 276 (2001); Paul Slovic, Melissa L. Finucane, Ellen Peters & Donald G. MacGregor, *Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason,*

somatic marker that biases judgment away from committing alpha errors. For the argument to go through, the somatic marker must be attached to one of the potential outcomes: conviction.

The mental mechanism has been studied in the Iowa gambling task.²¹⁷ Experimental subjects have been given a choice between two classes of lotteries.²¹⁸ In the first class, gains were higher than in the second.²¹⁹ However, in the first class, losses were dramatic, whereas they were mild in the second class.²²⁰ Subjects did not get this information in advance.²²¹ They had to learn from experience.²²² Ordinary subjects did so quite rapidly, and quite reliably.²²³ However, subjects with brain damage in the ventromedial prefrontal cortex, the region of the brain responsible for integrating emotional information with decision-making, failed miserably on this task.²²⁴ They were overly attracted by the prospect of immediate gain, and quickly went bankrupt.²²⁵

CONCLUSION

The difference between American and Continental standards of proof is not just doctrinal. It also is not just a matter of degree. These legal orders conceptualize proof differently. Some American authors even advocate a different norm. They want the standard of proof to result from maximizing a social welfare function. The same approach is prevalent in the discussion of standards of proof in the economic literature. If one accepts the utilitarian norm, however, one is forced to have different standards of proof for different classes of offenses. Ultimately, one even needs information about the respective composition of the population of defendants. This is not what legal orders do, on either side of the Atlantic. Judicial practice is captured by a more modest norm. It defines the maximum tolerable number of materially wrong decisions. In statistical jargon, it is exclusively concerned with the alpha error.

Risk, and Rationality, 24 RISK ANALYSIS 311, 318 (2002).

217. Antoine Bechara, Antonio R. Damasio, Hanna Damasio & Steven W. Anderson, *Insensitivity to Future Consequences Following Damage to Human Prefrontal Cortex*, 50 COGNITION 7 (1994).

218. *Id.* at 8–10.

219. *Id.* at 8.

220. *Id.* at 8–10.

221. *Id.* at 10–12.

222. *Id.*

223. *Id.*

224. *Id.*

225. *Id.* at 10, 13–14.

Judicial proceedings are characterized by a high degree of complexity. Normally, decisions have to be taken on an incomplete factual basis. Human decision-makers are unable to take such decisions in a quasi-scientific manner. By contrast, the unconscious automatic system is made for such tasks. The parallel-constraint-satisfaction mechanism calculates the relatively best way to make sense of the available evidence. In so doing it polarizes judgment. Evidence supporting the final decision is inflated. Conflicting evidence is discounted.

While the automatic system polarizes, it nonetheless is sensitive to exogenous intervention. A high standard of proof increases the minimum distance in activation between the preferred and all alternative decisions, and it increases the minimum level of overall consistency necessary for taking action. The stricter standard is imposed on the automatic system by a somatic marker. The jury instructions appeal to jurors' ethos. They make accountability towards the defendant, and towards society at large, salient.

The subjectivist standard of proof in Continental legal orders therefore mirrors the mental activity of real judges and jurors. It is descriptively correct, and it in principle is able to reach the stated normative goal. Some American observers do not deny the descriptive statement, but they are opposed to the prescriptive part.²²⁶ Judicial intuition is indeed not foolproof. Since it partly relies on idiosyncratic memory,²²⁷ the outcome is not fully predictable.²²⁸ For instance, the members of mock juries differ in the strictness of standards.²²⁹ Inadmissible evidence²³⁰ and stereotypes influence judgment.²³¹ If jurors believe in an outcome, this (preexisting) belief biases how they assess the evidence.²³² Coherence shifts continuously, even if jurors are exhorted to withhold judgment until they have heard all of the evidence.²³³ This can

226. See, e.g., Clermont & Sherwin, *supra* note 26, at 271 (admitting that "a standard of virtual certainty is better designed to produce true answers to disputed questions of fact than the preponderance standard," but concluding that *intime conviction* does not facilitate dispute resolution because "no court has unlimited time and resources" and Civil Law courts have become adversarial).

227. See Simon, *supra* note 17, at 536 ("All mental processing draws closely from one's background knowledge.").

228. Cf. CHRISTOPH ENGEL, GENERATING PREDICTABILITY: INSTITUTIONAL ANALYSIS AND INSTITUTIONAL DESIGN 20–21 (2005) (discussing how certain kinds of behavior can be predicted by using various theories such as rational-choice theory or behavior economics); Engel, *Intuitive Man*, *supra* note 169, at 397–98 (noting that intuitive decision-making is by definition "not under conscious control" and is thus unpredictable).

229. Franklin, *supra* note 11, at 163, 165; Saunders, *supra* note 56, at 18.

230. See Simon, *supra* note 17, at 538 (including personality traits as inadmissible evidence).

231. *Id.* at 542.

232. Simon et al., *Constraint Satisfaction*, *supra* note 181, at 827.

233. Pennington & Hastie, *Explaining the Evidence*, *supra* note 111, at 190; Simon, *supra* note 17, at 542.

be exploited by the parties. They can manipulate the order in which they present the evidence, in the interests of tilting juror judgment.

Continental law therefore has no reason to be haughty. The subjectivist standard of proof is far from perfect. Judicial performance might well be improved by adding objectivist elements. Actually, this is already the case. For instance, Continental courts are happy to rely on DNA evidence, and to assess its quality with formal statistical tools. The most important and furthest-reaching objectivist element is the obligation to write down reasons. This is not meant to be a description of the mental process. But it also is not just lip service. The need to later give an explicit justification is anticipated when the decision is taken. It influences the mental representation of the case.²³⁴

Conversely, American law, in light of the knowledge about juror psychology, might become more skeptical about academic attempts to erase subjectivist traces from the law of evidence. It also might want to explore the psychological feasibility of the preponderance of the evidence standard. How good are jurors at finding that a claim is “probably true”? Do they perform more reliably if they have a chance to compare the plausibility of two competing stories? Does the procedure of civil law disputes pay due regard to the conditions for the formation of juror intuition?

234. Engel, *The Psychological Case*, *supra* note 207, at 94.