STANDARDS OF PROOF REVISITED

Kevin M. Clermont*†

ABSTRACT

This Essay focuses not on how fact-finders process evidence but on how they apply the specified standard of proof to their finding. The oddity that prompts speculation is that, in noncriminal cases, the common law asks only that the fact appear more likely than not, while the Civil Law seems to apply the same high standard in these cases as it does in criminal cases. As a psychological explanation of the cognitive processes involved, some theorists posit that the bulk of fact-finding is an unconscious process, powerful but dangerous, which generates a level of confidence against which the fact-finder could apply the standard of proof. But this foggy confidence-based theory fails because standards of proof should, and fact-finders arguably do, concern themselves with probability rather than confidence. Psychology also cannot explain the divide between the common law and the Civil Law because the real explanation likely lies in the different goals that the two procedural systems pursue through their standards of proof.

INTRODUCTION

Magisterially, but opaquely, the law tells its fact-finder to apply the standard of proof. The law will have already chosen the appropriate standard, or required degree of persuasion, from an array of candidates. In the United States, that array includes preponderance of the evidence, clear and convincing proof, and proof beyond a reasonable doubt. At one end of the array is the preponderance standard, which prevails in civil cases and translates into “more likely than not.” This standard acts to minimize the expected cost of error if an error against the plaintiff is just as costly as an error against the defendant. At the other end, proof beyond a reasonable doubt prevails in criminal cases and means a high degree of probability or even almost certainty. This standard would also serve to minimize the expected cost of error because the error of convicting the innocent is especially costly.1

* Ziff Professor of Law, Cornell University.
† I want to thank Ron Allen, Adrienne Clermont, Christoph Engel, Valerie Hans, Reid Hastie, David Kaye, John Leubsdorf, Roald Nashi, Jeff Rachlinski, Emily Sherwin, and Wendy Williams for illuminating conversations and other communications on law and psychology.
The law, most theorists, and I have all proceeded on the assumption that the fact-finder—in a subsequent step, after having somehow processed the evidence in order to test a hypothesis or to find tentatively a fact—dutifully tries, within human limits, to compare the probability of the burdened party’s version of fact to the given standard of proof. This contestable, and contested, leap of faith in assuming a description of behavior is the subject of this Essay.

What this Essay is not about—and this caution is the critical starting point—is how the fact-finder processes the evidence in the first place. Thus, I need not enter the main battlefield of Bayesians versus non-Bayesians over the model for evidence-processing. Although I myself find the psychologists’ theories on holistic and largely nonprobabilistic assessment of evidence very well supported, and convincing based on my experience, I realize that ascertaining the actual approach or even the prescriptive ideal for processing evidence would be an exceedingly difficult endeavor.

In this Essay, to avoid such difficult problems, I simply assume that the fact-finder has somehow evaluated a hypothesized fact. Then I focus on the step when the fact-finder compares the processed proof to the standard of proof—that is, not the evidence-processing step but the standard-of-proof step—even if the two steps might not be entirely distinct in operation or timing. Although the law tries forcibly to make the latter step probabilistic, I shall consider its reality: How does the fact-finder actually decide whether the burdened party has established its version of fact strongly enough?

The reader’s first reaction could very well be to ask what turns on the answer to my question. I think quite a lot. Most obviously, how the law sets and formulates its standards of proof may depend on the answer. The setting and formulating of these standards are certainly not beyond debate. Most notably, the Civil Law system prevailing in much of the world applies a high standard not only to its criminal cases but to all its cases, while our common law does not apply its high standard to noncriminal cases. I begin there.

---

I. THE CIVIL LAW’S PUZZLE

The Civil Law’s position has provoked me to express wonderment in a series of articles over the years. Instead of asking whether some fact $X$ (say, that the defendant executed the promissory note disputed in a noncriminal, or civil, lawsuit) is more likely true than not, the Civil Law asks whether the fact is so probable as to create an inner and deep-seated conviction of its truth. That is, instead of a standard requiring merely a preponderance of the evidence, they apply their intime conviction standard.

What exactly does intime conviction mean? Civilians are usually vague on this point, only occasionally muttering something definitive, which could even sound like proof beyond a shadow of a doubt. The essence, though, is that they want to be very sure before upholding the burdened party’s proof. Alternatively stated, they are much more willing to accept a false negative than a false positive.

The surprising aspect is that civilians say they apply the same or a very similar standard in noncriminal cases as they do in criminal cases. And, in fact, they really do apply the criminal law’s high standard of intime conviction in purely civil cases. The peculiarly French practice of joining civil claims to parallel criminal cases demonstrates this fact. The high criminal standard directly and unavoidably determines the outcome in these French civil cases in which the plaintiff chooses to join with the criminal proceedings by acting as a so-called partie civile. Thus, a loss on the criminal case will dictate a loss on the joined civil case. Most interestingly, French plaintiffs prefer to join their civil cases to parallel criminal proceedings whenever possible in order to save time, effort, and money and to gain leverage. In other words, potential civil plaintiffs seem to calculate that no advantage will be lost by joining with the criminal proceedings, because they realize “that even in a purely civil case there is a presumption of the defendant’s innocence.”

I shall not repeat the ready explanation of how the approaches of Civil and common law diverged on the noncriminal standard of proof around the time of the French Revolution. My articles’ difficulty lay instead in explaining how these two systems could today share general aims and yet preserve their historical difference on such a fundamental


aspect. After all, this difference on standard of proof—intime conviction versus preponderance of the evidence—not only entails all sorts of practical consequences but also implies some basic differences on attitudes toward the nature of truth.

My suspicion as to the persistent difference between the two standards, among all the possibilities developed at length in my previous articles, was that the chosen standard of proof serves a very different function in each of the legal systems. After deeming the expected-error-cost-minimizing approach to be procedurally proper and recognizing that noncriminal cases pose no substantive reason to weight the balance against one side of the dispute, the common law chose the preponderance-of-the-evidence standard. The Civil Law must have chosen to sacrifice any sort of error-minimization in pursuit of a different goal, as to which my best surmise was perceived legitimacy.

The Civil Law’s quest for legitimacy is by no means silly. Public perceptions of what courts do may be just as important as what courts do in actuality. Even judicial myths that enhance legitimacy can sometimes have great utility. Civil Law courts of course are not alone in seeking legitimacy for their decisions. Common-law courts have had their own means of maintaining an appearance of fairness and accuracy, such as delegating fact-finding to an impartial jury of citizens who supposedly embody common sense and community values.

What benefits of legitimacy does the intime conviction standard bestow? Although the probabilistic nature of fact-finding is inevitable in our uncertain world, probabilism does produce discomfort. It also requires sophistication and imposes some destructive side-effects, at least if acknowledged by the legal system. The Civil Law system could be choosing to avoid all those ills by refusing openly to acknowledge probabilism. If so, civilians would be relying on the widespread efficacy of the people’s unexamined intuition, which would yield the legitimating misimpression that requiring virtual certainty comports with finding real “truth.” That is, the high standard would insinuate to the parties and the public that judges will not treat facts as true on less-than-certain evidence. Civil Law courts and their magisterial judges, sensitive to the special need for legitimacy in a social structure historically wary of the judiciary, may have been particularly eager to deal only in “truth.”

Moreover, Civil Law judges would prefer not to appear to have free rein to call close cases or to exercise discretion. That is to say, the high standard helps them to appear to act only when more-or-less forced to act. This rhetoric and form thus sound and appear good, and imply a seriousness of purpose and a caution in action. When the court does render judgment,
the high standard retrospectively implies that the evidence must have been certain and the result necessary. All these implications are admittedly appealing, even if false or misleading.

Relatedly, the high standard of proof gives comfort to judges. A high standard gives the judge a rock to hide behind. The judge will not have to call evenly balanced cases. The party with the burden of proof will win only when clearly entitled to win; if that party loses, the standard of proof, not the judge, is to blame. Likewise, the judge might duck a difficult decision, or otherwise seek to self-protect, by relying on the unrealistically high standard.\(^5\)

Therefore, the current motivation for the difference on standards of proof, in my view, lies in the subtle differences between the two systems’ objectives to which their standards of proof conform. The Civil Law seeks the legitimating benefits of the myth that its courts act only on truly true facts and not on mere probabilities. Courts of the common law seek legitimacy elsewhere, and thus are free to adopt the standard of proof that more efficiently, fairly, and openly captures the facts of the case.\(^6\)

What is important to stress is that the explanatory difficulty in my earlier articles did not regard which system is right or wrong, but instead lay in figuring out why the two legal systems remain apart. Whether either legal system is wrong would not only require a surer determination of the goals underlying the system’s standard of proof, but also turn on a subtle balance of benefits and costs. The civilians have concluded that the cost–benefit balance favors a high standard for their noncriminal cases. As I previously wrote: “Now, I am not arguing that the civil law’s valuing of legitimacy is worse or better than the common law’s view, but rather that it reflects a different weighing of costs and benefits in a different legal context.”\(^7\) I still believe that these very different standards of proof may each be right for the respective legal system.

---


6. See id. at 274–75.


The civil law seeks the legitimating benefits of the myth that their courts act only on true facts and not on mere probabilities. Common-law courts seek legitimacy elsewhere . . . and thus are free to adopt the standard of proof that more fairly and efficiently captures the real truth of the case.

*Id.*
II. OTHERS’ REACTIONS

Such comparative theorizing naturally ruffled some feathers. Some responses expressed mostly irritation. But an interesting responsive theme is the emerging attempt to stretch psychology’s insights on the fact-finder’s processing of proof to the different question of the fact-finder’s handling of the standard of proof. From Germany, Professor Christoph

8. In a remarkably aggressive article on the subject of Clermont & Sherwin, A Comparative View, supra note 3, a European author surprisingly concedes that the common law’s preponderance standard of proof is superior, but nevertheless criticizes our article’s development of arguments that would work in support of that view. Michele Taruffo, Rethinking the Standards of Proof, 51 AM. J. COMP. L. 659, 666 (2003). He makes only one relevant and possibly accurate point: some Civil Law countries, such as Spain and Italy but not the more rigorous Germany, are adamantly nonexpress about their higher civil standard of proof and thereby free up their judges to apply whatever standard the judges think appropriate in the search for “truth” or “moral certainty.” See id. at 666–69 (characterizing intime conviction as “a sort of black box”). He even seems to approve of this practice. Cf. id. at 670–71 (maintaining fashionably that truth is what judges say it is). We in fact had developed at length his observation about judicial freedom to lower the standard in practice. But we did not feel comfortable resting on it, if only because such judicial discretion is so obviously dangerous that we would never have felt competent to attribute it to the Civil Law. Instead, we discounted his explanation on the grounds that some real difference in standards between common law and Civil Law does survive, not only in language but also in practice. Clermont & Sherwin, A Comparative View, supra note 3, at 261–64.

9. An early step in this direction was the clever article by Ronald J. Allen & Sarah A. Jehl, Burdens of Persuasion in Civil Cases: Algorithms v. Explanations, 2003 MICH. ST. L. REV. 893. The article builds on the story model of evidence-processing, see infra note 11 and accompanying text, to produce another theoretical brand of holism, this one called the relative-plausibility theory. See Allen & Jehl, supra, at 929–43 (summarizing Professor Allen’s previous work on the theory). The relative-plausibility theory posits that the fact-finder constructs the story (or stories) that the plaintiff is spinning and another story (or stories) that the defendant is spinning. The fact-finder then compares the two stories (or collections of stories) and gives victory to the plaintiff if the plaintiff’s version is more plausible than the defendant’s. This choice between alternative competing narratives is largely an ordinal process rather than a cardinal one. This theory has the advantage of nicely circumventing the conjunction paradox, which comprises the oddity that applying the standard of proof element-by-element effectively causes the standard to lower as the number of disputed independent elements increases.

But this theory—which couples the holistic evidence-processing of the story model with a new notion of ordinal comparison as the standard of proof—has problems. First, all its work on the conjunction paradox is performed not by relative plausibility’s ordinal comparison, but by the underlying story model’s insistence that the fact-finder ultimately views the claim as a whole rather than going element-by-element. The story model leaves the conjunction paradox standing as a legal aberration, but one of small practical importance. See Richard H. Field, Benjamin Kaplan & Kevin M. Clermont, Materials for a Basic Course in Civil Procedure 1342–45 (9th ed. 2007) (observing also that the story model conflicts not only with the law’s words but also with many of the law’s practices such as nongeneral verdicts and judgments as a matter of law). Second, the most obvious difficulties of ordinal comparison are that ordinality and cardinality are not really so different as explanations for a more-likely-than-not standard and that ordinality cannot easily explain standards of proof higher or lower than preponderance of the evidence. See Friedman, supra note 2, at 2046–47; cf. Clermont, Procedure’s Magical Number, supra note 1, at 1119–20, 1122–26 (discussing not only the standard of clear and convincing evidence, but also procedure’s lower standards including slightest, reasonable, and substantial possibilities). But cf. Allen & Jehl, supra, at 930–37 (stressing the difference between what ordinal and cardinal approaches compare); Ronald J. Allen, The Nature of Juridical
Engel’s new article represents a major step forward in the development of that theme.10

Let me outline the argument presented in Engel’s article. First, he accepts the existence of the difference between Civil Law and common law as to standards of proof. Second, he acknowledges the usual normative position that different societal goals could call for different standards of proof. Third, he shifts to a behavioral account of how the fact-finder actually implements a given standard, which is where things eventually get new and interesting.

In that behavioral account, Engel initially sets forth the broadly accepted “story model” of evidence-processing.11 It holds that the fact-finder, over the process’s course, constructs from the evidence the story that makes maximal sense. The fact-finder then chooses, among the available decisions, the one that fits best with the story it has constructed. Along the way, and by a similarly automatic or unconscious process, the fact-finder generates a level of confidence in the decision by considering the degree of coverage (which means the story accounts for all the evidence), coherence (which means it is internally consistent, plausible with respect to the fact-finder’s world knowledge, and complete without striking gaps in expected components), and uniqueness (which means the absence of plausible alternative stories). The clearer the view of the case

---


that the chosen story delivers to the fact-finder, the more confident the fact-finder will be.

Professor Engel next turns to the literature on “coherence-based reasoning.” He does so primarily to stress the dangers in this unconscious process whereby the fact-finder strives to make sense of the evidence. A major danger is the so-called coherence shift: the fact-finder, in constructing its narrative, will tend to overvalue consistent evidence and devalue conflicting evidence. This cognitive failing will increase the level of confidence as well as the incidence of errors.

He finally reaches the law on standards of proof. Here is his real contribution, because earlier psychological models, including the story model, failed altogether to give attention to the standard-of-proof step in decision-making. He posits that standards can act to set a “somatic marker,” which gets the fact-finder’s emotional juices flowing in a direction that will offset the bias of the coherence shift. In particular, if attached to one of the potential outcomes, the standards can serve as a warning of the undesirability of a false finding in that direction. A high standard of proof makes the fact-finder feel more accountable, inclining it to avoid a falsely positive result. “When instructed to convict only if guilt is beyond a reasonable doubt, subjects apply a stricter standard of coverage, coherence, and uniqueness.” By contrast, a low standard of proof such as the preponderance of the evidence leaves the fact-finder dispassionate, indifferent as between false negatives and false positives.

14. Actually, he writes: “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.” Id. at 464. He indeed believes that the preponderance standard “tolerates a substantially higher error rate” than the criminal standard. Id. at 444 (citing as sole support James Brook, Inevitable Errors: The Preponderance of the Evidence Standard in Civil Litigation, 18 TULSA L.J. 79, 85 (1982), a fine early article whose point was to defend the preponderance standard on the ground that it was error-minimizing). Because he believes that the standard of proof asks only how confident the fact-finder is, he naturally but wrongly concludes that a lower standard, and hence lower confidence, means more errors. Instead, requiring high confidence will greatly increase the number of false negatives, even if that strategy limits false positives; actually, low confidence, as long as the found fact is more likely than not, will minimize the expected number of errors. See supra note 1. In brief, Professor Engel’s view reflects a basic misunderstanding of the nature of the preponderance standard. See D.H. Kaye, Clarifying the Burden of Persuasion: What Bayesian Decision Rules Do and Do Not Do, 3 INT’L J. EVIDENCE & PROOF 1, 27–28 (1999) (establishing that preponderance minimizes expected error costs); Kaye, Equal Error Rates, supra note 1 (refining his own analysis).
To summarize, one new idea is that the bulk of fact-finding is an unconscious process, powerful but dangerous, which generates a level of confidence against which the fact-finder could apply the standard of proof. The other new idea is that the only role of the law on standards of proof is to intrude beneficially on the unconscious process by using a high standard to marshal human emotions against cognitive bias. I find this brace of ideas intriguing. If accurate, I would see either or both as friendly addenda to the viewpoint expressed in my articles. But I hesitate to accept either the breadth of Engel’s psychological explanation or the narrowness of the role he leaves to law.

III. PSYCHOLOGY’S MECHANISM

Psychologically speaking, how does a fact-finder apply a standard of proof to fact X, be it a simple fact or a more complexly formulated narrative? The short answer is that no one yet knows.

Professor Engel says little about exactly how the fact-finder applies the standard. Instead, he talks mainly of evidence-processing. He mentions the well-recognized concept of confidence. He then appears to equate the standard of proof to confidence: “The psychological correlate of the standard of proof is confidence.” 15 But he fails to explain how

Therefore, my text here on the signal sent by the preponderance standard tries to restate his position in sensible terms, if one were properly to view the standard as setting a low probability threshold.

15. Engel, supra note 10, at 458. He further links the standard of proof to confidence by his discussion of statistical significance. Id. at 443–46. He claims that under the criminal standard there must be enough evidence to make the fact-finder more than ninety-five percent confident that the null hypothesis (not guilty) is not true. This view presents a couple of difficulties. First, it is difficult to express communicatively the preponderance standard in like terms, because it would have to involve bidirectionally equalized significance levels; he does not undertake this task, but instead jumps to conclude wrongly that preponderance is very tolerant of error (in fact, it is relatively tolerant only of the falsely positive type of error). See supra note 14. Second, statistical significance does not necessarily imply a high probability, nor does the absence of statistical significance necessarily imply a low probability; thus, unacceptable results will follow unless the criminal standard requires a high estimate of probability. See Simon Blackburn, Think: A Compelling Introduction to Philosophy 224–25 (1999) (criticizing scientific research for inferring causation and correlation from results that statistically could have been a product of chance); cf. infra text accompanying note 23 (treating the psychological version of confidence). In brief, Professor Engel’s view reflects a basic misunderstanding of the nature of the criminal standard. See David H. Kaye, Statistical Significance and the Burden of Persuasion, Law & Contemp. Pros., Autumn 1983, at 13, 17, 23 & n.47 (appreciating “why one cannot identify a unique level of ‘statistical significance’ that would correspond to proof satisfying the burden of persuasion appropriate to a given type of case”); D.H. Kaye, Do We Need a Calculus of Weight to Understand Proof Beyond a Reasonable Doubt?, 66 B.U. L. Rev. 657 (1986) (applying his insight specifically). But see James Q. Whitman, The Origins of Reasonable Doubt: Theological Roots of the Criminal Trial 4 (2008) (arguing that the reasonable-doubt standard originated “to make conviction easier, by assuring jurors that their souls were safe if they voted to condemn the accused”).
confidence could actually work as a standard of proof. He thereby evades three tough explanatory steps.

First, how does the fact-finder generate a usable confidence level, where confidence has the psychological meaning of how sure the fact-finder feels about its output? The task, according to prior research, seems to require considering the coverage, coherence, and uniqueness of the factual account. Formulating a confidence level with sufficient clarity to apply the standard of proof would be no easy mental task. Presumably, humans would be performing this complex task on the unconscious level, and so it defies precise explanation. His lack of precision here is therefore understandable.

Second, why should we rely on confidence, which is admittedly a very shaky measure? Much empirical research shows that confidence has a weak linkage to accuracy, for reasons including that the evidence-providers can easily manipulate the level of confidence as by the order of presentation, that the fact-finder’s generation of confidence falls subject to all sorts of biases such as the coherence shift’s confirmation bias, and that confidence depends on many external factors including self-image. Here, I suppose, his argument is that humans can do no better with any other alternative measure.

Third, and most important, how does the felt level of confidence translate into whether or not the evidence meets the standard of proof? His suggestion is that the standard of proof demands a specific level of confidence from the fact-finder in its view on fact X. Thus, a high criminal standard demands great confidence, whereas a preponderance standard would settle for a shaky conclusion. The criminal fact-finder and the noncriminal fact-finder would both be looking at whether they somehow “believe” the fact is true, but might have to be much more confident in that belief if the case were a criminal one. This attempt at brief description shows that confidence is no self-explanatory tool for handling ambiguous

---

Note that I am not contending that the criminal standard could not contain a confidence measure in addition to, rather than in lieu of, a required probability. Maybe the presence of that extra something would explain why Professor Engel intuits that the criminal standard has a subjective element that is lacking in the preponderance standard. See Engel, supra note 10, at 436, 441–42, 461–63 (saying that the high standard looks for a personal conviction). Further, the culturally fraught criminal standard of proof and its wordy formulation could convey, on an emotional level, the message to the fact-finder’s unconscious that there is this need for personal conviction based on confidence in guilt, plus a very high probability of guilt. See id. at 460–62. But see James Q. Whitman, The Origins of Reasonable Doubt: Theological Roots of the Criminal Trial 4 (2008) (arguing that the reasonable-doubt standard originated “to make conviction easier, by assuring jurors that their souls were safe if they voted to condemn the accused”).

16. See generally David Dunning, Self-Insight: Roadblocks and Detours on the Path to Knowing Thyself 6–9, 37–61 (2005) (describing the tendency for people to be overconfident and positing a theory of how we develop self-judgment).
evidence and also that confidence would be an especially clumsy tool for administering a preponderance standard in a legal system choosing to deploy it in pursuit of error-minimization.17

A. Confidence’s Prescriptive Problems

Yet, this confidence test is not how the standard of proof should work. A focus on probability is the mathematical route to minimizing the expected cost of erroneous decisions. The standard of proof therefore asks how likely the fact is, but the confidence measure answers with how sure the fact-finder is in its belief. In brief, confidence is a kind of second-order probability that is irrelevant to the first-order probability with which standards of proof should be concerned.18

Confidence can at most serve as an imperfect surrogate for probability, or likelihood.19 It may be that as probability goes up, confidence tends to go up too.20 But correlation does not mean that the two measures are

17. See supra note 15 (explaining the difficulty of expressing preponderance in terms of confidence). This difficulty likely explains why Professor Engel switches, at one point, to a different explanation of how confidence would translate into various standards of proof:

The easiest way to translate these techniques of calibrating judgment into mental mechanisms is by manipulating the threshold. . . . [D]epending 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. . . . For the remaining elements of the story model, one needs a richer decision criterion.

Engel, supra note 10, at 461–63; cf. supra note 9 (considering Professor Allen’s similar explanation, and observing that comparison of competing stories does not work well for standards higher or lower than preponderance).


19. In statistical terminology, “likelihood” (the chance that the data would be observed, given a hypothesis as true) is not wholly equivalent to “probability” (the chance that a hypothesis is true, given the observed data). E.g., RICHARD M. ROYALL, STATISTICAL EVIDENCE: A LIKELIHOOD PARADIGM 5–6, 28 (1997). But for most people, likelihood means probability. I use it here in that way, with perhaps the connotation of an intuitive measure of probability and with the benefit of conforming to the common legal, and probabilistic, usage of “more likely than not.”

20. “Subjects are most confident in their decision if they believe the case to be clear, i.e., if
equivalent. Confidence does not necessarily imply a high probability, nor does the absence of confidence necessarily imply a low probability. To focus on confidence is either to focus on the wrong thing or to conflate confidence and probability.

A telling counterexample to the equivalence of confidence and probability comes from a hypothetical that I posed on a recent examination. The disputed fact was which of two dead people had been driving a very negligently crashed car and which was the passenger; the evidence was that the car’s owner was driving on departure for the long trip and normally did not let others drive, but the accident scene provided no proof on the driver’s identity, and no other proof was available. In that hypothetical, the fact-finder could be highly confident that on all available evidence the owner was slightly more likely than not the driver. If the standard of proof turns on estimated likelihood, the fact-finder could find liability under a preponderance standard but could not find it beyond a reasonable doubt. Given the fact-finder’s sky-high confidence, however, an approach that ties the standard of proof exclusively to confidence would, albeit totally unacceptably, conclude the car owner was the driver under the criminal standard!

My point is that a fact-finder could be very confident in its determination that the fact was just slightly more likely than not, so that confidence and probability are not equivalent. But a confidence theorist, converting the vague definition of confidence into a virtue, might respond that the fact-finder in my hypothetical should not be very confident because of the paucity of evidence and the availability of a plausible alternative story: the theorist would prefer to say that we want the fact-finder to be confident in the “truth” rather than confident that the evidence shows a fact to be more likely than not. I would rejoin with an altered hypothetical in which there is a mountain of evidence all pointing to the car owner being slightly more likely the driver, including scientific evidence establishing a point estimate of sixty percent. More generally, I would rejoin that the
common-law fact-finder is not supposed to hold an unavoidable paucity of evidence against the burdened party, but is instead in such a situation supposed to decide likelihood based on the evidence.24

Epistemology provides another route to expose the difference between confidence and probability. The epistemological aim of evidence law is that the fact-finder should construct a belief that corresponds to the outside world’s truth.25 Probability is a measure of the chance of that correspondence existing between finding and reality. An alternative, but less desirable, approach would be for the fact-finder to seek an evidential construct (or agree to a construct) that coheres with the rest of the world image inside the fact-finder’s head. Confidence is a measure of the degree of that coherence between the new belief and all the old beliefs. Psychologists might approve of this strategy of looking to confidence, so celebrating the Civil Law’s “subjective” approach over the common law’s “objectivism.”26 Such a strategy is understandable: people prefer looking to an inside feeling rather than to any outside measure—and so might operationalize an inquiry about probability of a fact by changing it into an inquiry about confidence that the fact obtains—because of their assumption that one has better cognitive access to the state of one’s own mind. But

---

24. In other situations, fact-finders can properly use common experience of the world or sometimes even specific knowledge of the case. See, e.g., J. Alexander Tanford, An Introduction to Trial Law, 51 Mo. L. Rev. 623, 700 (1986) (“Jurors are encouraged to use common experience, but are prohibited from becoming a kind of ‘expert witness’ . . . .”); see also THOMAS CATHCART & DANIEL KLEIN, PLATO AND A PLATYPUS WALK INTO A BAR . . . UNDERSTANDING PHILOSOPHY THROUGH JOKES 53–54 (2007). Take, for example, the following story:

A defendant was on trial for murder. There was strong evidence indicating his guilt, but there was no corpse. In his closing statement, the defense attorney resorted to a trick. “Ladies and gentlemen of the jury,” he said. “I have a surprise for you all—within one minute, the person presumed dead will walk into this courtroom.”

He looked toward the courtroom door. The jurors, stunned, all looked eagerly. A minute passed. Nothing happened. Finally the lawyer said, “Actually, I made up the business about the dead man walking in. But you all looked at the door with anticipation. I therefore put it to you that there is reasonable doubt in this case as to whether anyone was killed, and I must insist that you return a verdict of ‘not guilty.’”

The jury retired to deliberate. A few minutes later, they returned and pronounced a verdict of “guilty.”

“But how could you do that?” bellowed the lawyer. “You must have had some doubt. I saw all of you stare at the door.”

The jury foreman replied, “Oh, we looked, but your client didn’t.”

Id.

25. See Allen & Leiter, supra note 9, at 1497–503 (explaining naturalized epistemology in the context of the law of evidence).

26. See Engel, supra note 10, at 436–37, 441–42, 461–63; cf. supra note 15 (attempting to put the distinction between subjective and objective to a more productive use).
philosophers instruct us that the inside world does not provide steadier
ground than the outside world. And more to my point, the law has no
reason to interest itself in any such internal feeling, which would be
irrelevant to societal needs. Instead, society wants the law to pursue an
optimal outcome in terms of reality. Accordingly, the law by its standard of
proof seeks to force the fact-finder, in the final decisional step, to link its
inside mental state to the outside real world.

B. Confidence’s Descriptive Difficulties

Regardless of the desirability of looking to probability, it could be
that many fact-finders in actuality look to confidence when called upon to
apply a standard of proof. But given the overall fogginess of the supposed
interaction between felt confidence and standards of proof, it was
foreordained that Professor Engel would provide no empirical showing
that fact-finders misinterpret standards of proof so as to impose only a
confidence test. His Erfurt Experiment measures subjects’ confidence
levels and subjective probabilities, but fails to show which drives
decision. All he shows is “that different standards of proof are
psychologically feasible.”

Other researchers have not advanced the ball much farther. Despite
considerable evidence that judges meet difficulty in conveying any standard
of proof to the jury, and that juries have difficulty in quantifying the
standards of proof, there is some evidence that is more hopeful. Finders of
a fact can focus separately on the question of how probable it is that the fact
is true. No need exists for them to perform a quantification step in that

developing an anti-luminosity argument.

28. See Engel, supra note 10, at 460–61. The experiment receives fuller description in Andreas
Glöckner & Christoph Engel, Can We Trust Intuitive Jurors? An Experimental Analysis (Max Planck

29. Engel, supra note 10, at 463.

30. See Clermont, Procedure’s Magical Number, supra note 1, at 1141 & n.110, 1146–47 &
nn.126–27, 1149 & n.135; see also Engel, supra note 10, at 458–61 (highlighting that the beyond-a-
reasonable-doubt standard translates into widely varying probabilities for the lay juror); Dorothy K.
Kagehiro, Defining the Standard of Proof in Jury Instructions, 1 Psychol. Sci. 194 (1990) (reviewing
research on juror comprehension that found if the standard of proof was stated in terms of probability,
then it had its anticipated effect; but when the standard was not so stated, it did not have its anticipated
effect); Elisabeth Stoffelmayr & Shari Seidman Diamond, The Conflict Between Precision and
(2000) (discussing experiments in which mock jurors failed to distinguish between standards of proof).

31. See Clermont, Procedure’s Magical Number, supra note 1, at 1143–48 (arguing that
decision-makers should continue to use the “seven categories of uncertainty . . . : (1) slightest
standards of proof can conceivably exist, how might it actually operate? The processing of evidence itself could have yielded a rough measure of probability, if it were in part Bayesian in nature, or perhaps yielded merely a level of confidence, if not at all Bayesian. I readily admit that it is rather implausible that fact-finders apply any rigorous probabilistic technique in complex processing of the evidence, and that they may indeed be wholly nonprobabilistic. However, there is nothing inconsistent or infeasible about even the nonprobabilistic fact-finder subsequently fulfilling, at least in some intuitive way, the straightforward probabilistic task of (1) estimating subjective probability, by looking at the evidence to measure roughly the probability of fact \( X \)'s existence as against the probability of not-\( X \), and (2) comparing it to a set of possible levels: (1) impossibility, (2) reasonable possibility, (3) substantial possibility, (4) equipoise, (5) probability, (6) high probability, and (7) almost certainty); cf. Engel, supra note 10, at 437–38, 460–61 (describing an experiment showing that subjects could measure subjective probability).

32. See Clermont, Procedure’s Magical Number, supra note 1, at 1146 & n.126, 1149 & n.136 (arguing that “quantification ordinarily is not a necessary or desirable step in applying the standard of proof”).


34. See supra note 28.

35. See Friedman, supra note 2, at 2044–46 (discussing a theory “[c]losely associated with Bayesian probability theory,” called Bayesian decision theory, which posits that decision-making is “based on the expected value of any course of action”). But see Allen & Jehl, supra note 9, at 930–36,
scale of coarse gradations, such as asking whether the fact is more likely
than not or whether it is almost certain. That is to say, a nonprobabilistic
model of the evidence-processing step, such as the story model, could be
consistent with this more probabilistic approach to the standard-of-proof
step. And this more probabilistic approach is entirely feasible. After all,
we perform such a task constantly in real life. We all competently apply
standards of proof similar to the law’s. I might decide to go see a
particular movie, rather than stay home, only if going to the movie is more
likely than not to deliver more pleasure. Countless decisions rest on cost
minimization, where one decides on a particular route over all available
alternatives because it is likely the best way to go, or the way one will be
happiest. Other decisions may entail a higher standard, because one
weighs errors in one direction more heavily than in the other direction. I
might decide to marry a particular person only if I am really, really sure
that this is the route to go.

I further admit that it is easier to imagine a law-trained judge
undertaking a deliberate and probabilistic approach to the standard of proof.
But I can also picture jurors performing this task. First, the judge delivers
careful instructions that require the jurors to approach the standard of proof
by a separate step and in probabilistic terms. Second, the task is really not
so demanding, as it requires only an intuitive comparison of rough
probability to a gradated scale. Third, jury deliberations could help
immensely in performing the task. Group discussion could reinforce the
duty to relate one’s inner conviction to likelihood in the outside world and
help to explain how to perform the task. We indeed know that jurors often
do not reach a decision until the deliberations phase, and that jurors can
use the standard of proof as an argument during deliberations.

Consequently, rather than assume that fact-finders in applying the
standard of proof invoke some unspecified and unproven confidence-based

36. See generally Clermont, Procedure’s Magical Number, supra note 1 (exploring the
psychological tendency to handle probability by coarse categories).
37. See Shari Seidman Diamond, Beyond Fantasy and Nightmare: A Portrait of the Jury, 54
BUFF. L. REV. 717, 758–61 (2006) (discussing the effect of jury deliberations on juror decision-making);
Hannaford et al., supra note 11, at 640 fig.3 (a graph detailing the percentage of jurors that made up
their mind at different stages of a trial; compared to other stages, more jurors formed a decision in the
deliberations phase); David H. Kaye, Valerie P. Hans, B. Michael Dann, Erin Farley & Stephanie
Albertson, Statistics in the Jury Box: How Jurors Respond to Mitochondrial DNA Match Probabilities, 4
J. EMPIRICAL LEGAL STUD. 797, 815–18 (2007) (explaining the effect of jury deliberations in a mock
criminal trial for robbery in which expert testimony on mitochondrial DNA sequencing was presented).
38. See Robert J. MacCoun & Norbert L. Kerr, Asymmetric Influence in Mock Jury
Deliberation: Jurors’ Bias for Leniency, 54 J. PERSONALITY & SOC. PSYCHOL. 21, 30 (1988)
(suggesting that “deliberation tends to amplify the effects of these instructions”).
approach, which would be both inaccurate and undesirable to apply, I shall continue to assume that the fact-finder does, even if in some rough and intuitive manner, what the law with good reason actually tells it to do: settle on a sense of subjective probability and compare it to the standard of proof’s coarse gradation.

IV. LAW’S ROLE

Whether one accepts as the manner of implementing the standard of proof either (1) a confidence-based process in the unconscious or (2) the traditional view of a more deliberate (even if largely intuitive) comparison of rough probability to a gradated scale, there remains the question of the law’s contribution to the mental process. On this question, a psychologist sees the law’s formulation of a high standard of proof as a way to set a somatic marker that on an emotional level conveys to the fact-finder’s unconscious the message to protect the defendant; the high standard drives home the duty to take fact-finding seriously and to demand a considerable feeling of confidence before finding for the burdened party. I tend to see the expression of a higher standard as a way to inform the fact-finder that the burdened party must provide a stronger showing in terms of probability. Importantly, we agree that no matter how the law’s standards actually work, they seem to succeed in their mission of tilting the playing field in the normatively desired direction.

The common law is obvious in trying to accomplish more, however, by adjusting the standard of proof to the circumstances of the case. This observation brings me back to why both the common law and the Civil Law apply a high standard to criminal cases, but only common-law jurisdictions lower the standard for noncriminal cases. Oddly, although Professor Engel set out to explain this very difference between the common law’s and the Civil Law’s standards of proof, his psychological detour has no explanatory power whatsoever on the divide. His bottom line reduces to this: “The heart of the conflict between American and Continental law is therefore the different assessments of cost and benefit [of unconscious decision-making]. Whereas Continental law is attracted by the benefit, American law is scared of the cost.”39 Quite simply, Engel’s explanation has to be inadequate

39. Engel, supra note 10, at 438. His logic appears to be that humans do confidence levels, not probability; that the only thing any standard of proof can do is impose a high-confidence requirement; and that because low confidence means more errors, the law should always require high confidence and never tolerate the low confidence permitted by a preponderance standard: therefore, the Civil Law is generally right in its intime conviction approach, and the common law inexplicably ceases to care about errors in its noncriminal cases. See id. at 464–66. Even assuming his psychology to be sound, his logic leads astray because of its faulty premises regarding the incidence of error under legal standards: he does
because it draws no distinction between noncriminal and criminal cases. Whether both systems proceed by confidence or by probability, or whether the Civil Law goes on the basis of confidence and the common law on probability, the question remains why the common law alone sometimes lowers its standard of proof. That is to say, for an adequate explanation, he would have further to explain why the Civil Law’s standard treats all cases alike, but the common law has decided to treat some cases one way, other cases another way, and still other cases a middle way. He does not and could not explain this because his confidence-directed concerns are orthogonal to the question at hand. Most likely, no psychology-driven theorizing can explain why the common law chooses to treat different cases differently because the explanation lies in the goals of the procedural system.

The common law in its standards of proof consistently tries to minimize expected error costs. Because costs of error obviously differ in different settings, the common law is forced to apply multiple standards of proof: preponderance of the evidence, clear and convincing proof, and proof beyond a reasonable doubt. The elevated standards help to avoid the high social costs of false positives, as in convicting the innocent, and the preponderance standard minimizes expected social costs in any setting where false positives are no more costly than false negatives.

The Civil Law, by contrast, applies basically the same standard of proof in all settings. In all cases, in the face of differing costs, the Civil Law sends to its fact-finder the strong signal to protect the nonburdened party. It follows that the Civil Law is not trying solely to minimize expected error costs. The Civil Law must have another purpose. My best surmise remains that it is seeking the perceived legitimacy of the higher standard. But it pursues that goal at the expense of the burdened party, whoever that party is after the law’s burdens and presumptions play out in a particular case.

In sum, the law does more by its standard of proof than signal the fact-finders. It also expresses a choice among the many possible goals of procedure. Given their different purposes, the common law and the Civil Law have chosen different but appropriate standards of proof. And, as already suggested, the law of each of the two systems has at least partially succeeded in expressing that standard to the fact-finders. Thus, the law appears to be relatively effective in this domain.

not realize that a lower standard, whether based on confidence or on probability, is in terms of error-minimization a more efficient and fair standard for noncriminal cases. See supra notes 14 & 15. Also perplexing is that his logic would seem to paint the Civil Law, in its avoidance of its fact-finders’ relying on best hunches in noncriminal cases, as being more rather than less wary of unconscious decision-making than the common law.
CONCLUSION

The psychological mechanism for implementing standards of proof remains to be discovered. For the time being, science gives law no reason to abandon its traditional hope that its intuitive fact-finder roughly estimates the subjective probability of the burdened party’s version of fact, and then compares the probability to a set scale of coarse gradations of probability, such as asking whether the version appears more likely than not or whether the version appears almost certain.

Meanwhile, the Civil Law and the common law apply different standards of proof in noncriminal cases, namely, the high intime conviction and the low preponderance standard. Again, psychology provides no magical explanation of this divide. Regardless of how the mind works in applying the standard of proof, the explanation for the difference in the two systems’ laws likely lies in their pursuing different goals when they formulated their standards of proof.