Standards of Persuasion and the Distinction between Fact and Law

Richard D. Friedman

University of Michigan Law School, rdfrdman@umich.edu

Available at: https://repository.law.umich.edu/articles/174

Follow this and additional works at: https://repository.law.umich.edu/articles

Part of the Civil Procedure Commons, Courts Commons, Criminal Procedure Commons, and the Evidence Commons

Recommended Citation

STANDARDS OF PERSUASION AND THE DISTINCTION BETWEEN FACT AND LAW

Richard D. Friedman*

The invitation to respond in these pages to Gary Lawson's very interesting article, *Proving the Law*, was tempting enough. But what made it irresistible was Professor Lawson's comment that he is "addressing, with a brevity that borders on the irresponsible, subjects well beyond [his] depth."¹ Now, that's the kind of debate I really like. Let me jump right in.

A principal question raised by Lawson, which I find quite interesting, may be phrased in general, and purposefully ambiguous, terms as follows: Before an actor treats a proposition as a valid proposition of law, what standard of persuasion should that proposition meet—that is, to what degree must the actor be persuaded of the merits of the proposition? This is a question that plainly might, and probably should, admit of different answers in different circumstances. Thus, a broader meta-question is: On what basis should a legal actor or observer determine the standard of persuasion that a proposition should meet?

Part II of this Essay approaches this meta-question in general terms. It concludes that only in certain circumstances is the question even coherent. Absent certain simplifying assumptions, it will not help the legal actor in deciding on an optimal course of conduct to determine a single standard of persuasion on a proposition of law in the given circumstances. An accurate decision requires a more complex determination of the expected value of each possible course of conduct. The complexity of this task may explain why, as Lawson notes, this entire matter has nearly been overlooked; whether or not legal actors happen upon sensible ways of dealing with uncertainty concerning the validity of legal propositions, it is probably too difficult for them to articulate just what they are doing.

Part III offers some tentative thoughts on a narrower conundrum presented by Lawson in the context of the criminal law. The issue may be posed briefly as follows: Given that a criminal defendant may not be convicted unless facts sufficient to support each element of the crime are

---

* Professor of Law, University of Michigan School of Law. Many thanks to Ron Allen.


2 "Valid" is a purposefully ambiguous term, because it has a meaning in both the descriptive and prescriptive senses. A proposition may be descriptively valid if it accurately reflects the state of the law. Or it may be prescriptively valid if it accurately reflects the interests that the lawmaker believes should underlie the law.
proven beyond a reasonable doubt, is it acceptable to allow the conviction to stand if the court is not persuaded to that degree that the legal propositions underlying the conviction are valid?

Before addressing these matters, I will offer some comments in Part I on a matter that underlies the discussion, both in Lawson’s article and in this Essay—the relation between law and fact.

I. THE LAW-FACT DISTINCTION

A. The Conceptual Distinction

According to Lawson, “the law-fact distinction . . . is purely a creature of convention.” This, it seems to me, is an overstatement. There is an analytic difference, independent of convention, between law and fact—though it is not always a distinction that will do us much good.

Lawson speaks of “a reality that exists independently of its acknowledgment by the conscious mind of a perceiver.” That strikes me as a serviceable definition of fact.

But this definition may not seem to distinguish law from fact. Lawson argues that “[f]rom an epistemological perspective, every positive propositional claim about the law in the form ‘the law is X’ is a factual claim of one sort or another.” Lawson calls this point an “innocent” one, and I suppose that it is, so long as it is limited to descriptive statements of law. Without defining “descriptive statement of law” too precisely, I mean the term to cover a statement of what certain officials or other significant actors have declared, or will declare, the law to be, or of what their conduct has indicated, or will indicate, they believe the law to be, or even merely of what their decisions have been, or will be, in given situations. The matters described by such a statement may well be deemed to be aspects of reality that exist independently of their conscious acknowledgment by the perceiver, and so factual. But these are factual matters of a very particular type, and usually can be distinguished from ordinary factual questions. If, for example, a litigation concerns an accident, then the questions “What have the courts said in the past that they would do about situations such as this,” and “What are they likely do about this one?” might be called factual. But they concern the past and prospective conduct of legal officials in determining legal norms that may be applied in this litigation. That is a quality absent from an ordinary

4 Lawson, supra note 1, at 863.
5 Though the following discussion takes a somewhat different approach, I believe it is substantially in accord with the excellent treatment of the subject in Henry P. Monaghan, Constitutional Fact Review, 85 Colum. L. Rev. 229, 232-37 (1985).
6 Lawson, supra note 1, at 866.
7 Id. at 863.
8 Id.
The conception of the concept of fact and law, both of which must be determined in adjudication, are distinct. We can therefore begin by thinking of two functions that, in theory, are also distinct. The fact-finding function is to determine that part of reality that is relevant to the adjudication of the action. We might think of this function as the reconstruction in imagination of that portion of reality, as if making a mental film. The law-determining function, then, is to prescribe the consequences to be attached to that aspect of reality.

This description usually understates the complexity of the two functions, however. Facts, at least disputed facts, usually cannot be determined to, or nearly to, a certainty; thus, the fact-finding function is to reconstruct in imagination various possible accounts of the relevant portion of reality, assigning a probability to each. And the law-determining function must take this uncertainty into account, prescribing the consequences not simply for a given factual state but for a given distribution of possible factual states; put another way, the law-determining function requires not only the determination of substantive norms but also the determination of standards of persuasion with respect to the facts.

The fact-finding and law-determining functions could be performed without any articulation at all, though the results would not be satisfactory. If an adjudicator were allowed to act like Justice Frankfurter’s "kadi under a tree," he might determine the outcome of a case without articulating either the guiding principles or factual findings; he might simply say, "Plaintiff shall take two cows," or "Judgment for the defendant." But justice demands judgment according to principle, and fairness and efficiency make some measure of predictability at least desirable. The adjudicator should at least attempt to articulate the set of principles

9 See Terminiello v. Chicago, 337 U.S. 1, 11 (1949) (Frankfurter, J., dissenting) ("This is a court of review, not a tribunal unbounded by rules. We do not sit like a kadi under a tree dispensing justice according to considerations of individual expediency.").
according to which he is deciding the case. Stating such principles, together with the outcome of the case, implies findings of facts (or, more precisely, of probability distributions of possible facts) that, under those principles, lead to that outcome.

B. The Burden of Articulation Imposed on the Fact Finder

The demands for articulation increase greatly if the pure fact-finding function is given to one actor and the pure law-determining function is given to another, for then those two actors must communicate with each other. Imagine first that the entire burden of articulation is imposed on the fact finder. That is, the fact finder communicates to the law determiner what really happened, and the law determiner decides, given those facts, what consequences to impose. Clearly, this is not a burden that the fact finder can carry.

The fact finder may of course get the facts wrong, but that need not trouble us here; our system tolerates a substantial probability of incorrect fact-finding with no apparent distress. A more serious problem is the one highlighted in section A: The fact finder is unable to determine with certainty what occurred. Thus, its report to the law determiner would have to include a probability distribution of possible accounts. This greatly magnifies the scope of the fact finder’s task.

Another problem, even putting factual uncertainty aside, is that, without categorizing and evaluating, the fact finder cannot possibly articulate all the separate aspects that constitute a real factual setting. If the fact finder really could produce a movie (or a polysensual counterpart that included smell, touch, taste, and emotion as well as sight and sound) and present it to the law determiner, that might solve the problem. But no fact finder composed of humans can do this. Instead, the fact finder must rely on words.

One reason the fact finder must rely on words in its actual determination of the facts is that to a large extent it must rely on witnesses. A witness is highly unlikely to be able to provide a demonstration of just how slippery the road surface was, or to make a statement such as, “The kinetic coefficient of friction of the road was 0.85.” The witness’ mind

---

And, in appropriate cases, what will happen (e.g., how bad the plaintiff’s disability will be), what would have happened in the past given a particular set of conditions (e.g., how much the plaintiff would have been able to work had the accident not occurred), and what would happen in the future given a particular set of conditions (e.g., how much would the plaintiff be able to work had the accident not occurred).


12 I use “witnesses” here in a broad sense to include not only an in-court witness but also an out-of-court declarant whose statement is admitted for the truth of what it asserts.
begins to evaluate and categorize reality as soon as her senses perceive and her brain records that reality.\(^{13}\) And, in attempting to communicate that reality, she must again evaluate and categorize by making choices in the use of words. Thus, instead of speaking in terms of the coefficient of friction, she is likely to say something like, "The road was very slippery."

Even if the witness were able to speak precisely in terms of the coefficient of friction, the fact finder would probably be unable to do much with that information—at least without translating it into terms with which it is more familiar, such as "very slippery." And, especially if the evidence that the fact finder receives is of the "very slippery" type, the fact finder is unlikely to be able to be more precise in presenting its own findings. Moreover, this problem is multiplied many-fold as the fact finder attempts to describe not just one particular circumstance but all aspects of a complex, dynamic situation.

In short, the limits of language and of the human mind mean inevitably that articulated factual findings will be evaluative categorizations. One problem with such categorizations might be that they include normative content—"very slippery" might mean in effect "too slippery to be driving the way the defendant was driving." Putting that problem aside for now, the difficulty is that such categorizations may not give the law determiner all the information she needs. How slippery, she may ask, is very slippery? Suppose the defendant's conduct, as described by the fact finder, should be deemed tortious given some degrees of slipperiness that might be considered very slippery, but not given other degrees of slipperiness. The case then is indeterminate.\(^{14}\)

C. The Burden of Articulation Imposed on the Law Determiner

Now consider the polar alternative: Suppose that the entire burden of articulation is put on the law determiner. In other words, the law determiner articulates with great specificity what facts (or, more pre-

---


As one introductory psychology text comments:

[O]ur senses provide us with raw data about the external world. However, without interpretation, this raw information remains what William James (1890) called "a booming, buzzing confusion." . . . Ultimately, it is the brain that interprets the complex flow of information from the various senses. Using sensory information as raw material, the brain creates perceptual experiences that go beyond what is sensed . . . . In other words, we use sensory information to create perceptions that are more than just sums of various parts. We tend to fill in the missing information, to group various objects together, to see whole objects and hear meaningful sounds rather than just meaningless bits and pieces of raw sensory data.


\(^{14}\) One way the law determiner might respond to this problem is to assume that the fact finder has found, or would find, the most extreme possible set of facts that could reasonably be found on the state of the evidence. Thus, for example, a court considering a motion for judgment notwithstanding the verdict will "construe the evidence and inferences most favorably to the non-moving party." Bruno v. Western Elec. Co., 829 F.2d 957, 962 (10th Cir. 1987). This is not a happy solution; the facts hypothesized by the court may be far different from the facts found by the jury.
Standards of Persuasion

cisely, what distributions of factual possibilities) will support what result, and the fact finder, following those instructions, simply chooses the result. To model this view, one might think of every expectable combination of factual findings as a point in a field of infinite dimensions. The law determiner must create a template that will be applied to that field, determining the consequences for given areas—that is, for given sets of points. A simple template would have just two areas, perhaps one marked "Liability" and the other marked "No Liability." In a particularly simple form of such a template, just one factor would separate the two areas, at least in the range of points covering ordinarily expectable findings. More complex templates, or rules, would often be necessary, however, to yield acceptable results. In any event, the fact finder need only determine in which area it believes the point representing reality lies.

I believe this theoretical model has its appeal. In pure form, though, it cannot be translated into operational reality. It might have one advantage over the model presented in section B, in which the fact finder does the articulation: If the law determiner is a court and the fact finder is a jury, the former is in all probability more articulate than the latter. This point cannot be taken too far, however; the court is unlikely to be able to determine precisely what facts support what judgment. Thus, the court will not be able to give instructions in the form, "If you find that it is more likely than not that the defendant's speed, as expressed in miles per hour, divided by the coefficient of friction, was 65 or more, and if . . . ." Moreover, even if the court were perfectly articulate, the jury would still in all probability be unable to work sensibly in such terms.

There is a further problem. Carrying the full burden of articulation means that the court must articulate a rule indicating what the result should be for each possible combination of facts that the jury might plausibly find. If the rule depends on one or two factors, or if each factor may be considered separately, articulating the rule so fully may not be too difficult. In many settings, however, the task would be intractable, requiring the court to articulate the interplay of a long series of factors across a many-dimensional field of possible facts.

D. Resolution in Tension

In reality, then, the functions of law-determination and fact-finding cannot be allocated in pure form. Fact finders cannot perfectly articulate the facts that they find. Indeed, juries are not ordinarily expected to be articulate at all. A trial court sitting without a jury is expected to articu-

15 In fact, I have been puttering intermittently for some time with an unpublished paper called A Spatial View of Law, which I would be delighted to share with interested readers and which, among other things, elaborates on this model.
late its factual findings, but it may use conclusory statements to jump the gap between its articulations of factual reality and of legal standard. On the legal side, courts articulate generalized norms, and may attempt to articulate more particular applications. But with respect to most, and perhaps all, legal questions, there comes a point in the articulation of the standard where the courts are unwilling and perhaps unable to be more precise. When it reaches this point, a court might still impose a standard, albeit an unarticulated one. Sometimes, however, by using such open-textured words as “reasonable,” a court leaves to the jury the determination of the applicable standard in the particular case.

Courts might speak of a standard such as “reasonable” as a question of fact. We should not be fooled. The jury in such a case does more than determine an aspect of reality. It also determines the norms that will be applied in that case. In a real sense, it makes law—but inarticulate law that is good for the one case only. Thus, the courts, in stating general principles, are the wholesalers of law, but often the retail work is done by the jury.

Nor should we be satisfied by referring to “mixed questions of fact and law.” That term may often be accurate enough, but it is potentially a cop out, because it can obscure the complexity of what occurs. Both the court and the jury may have to consider legal standards and facts simultaneously. But that does not mean that the function performed by either lies on a continuum between fact-finding and law-determination. Rather, there may be aspects of both fact-finding and law-determination in the functions of both the court and the jury.

When the court decides whether the plaintiff’s case is sufficient to go to the jury, it need not determine what the actual facts are. It must, however, determine whether the jury might reasonably find facts that are sufficient to support the claim. This involves both a form of fact-finding,

See, e.g., FED. R. CIV. P. 52(a); see generally Edward H. Cooper, Civil Rule 52(a): Rationing and Rationalizing the Resources of Appellate Review, 63 NOTRE DAME L. REV. 645 (1988).

If the fact finder were articulate, the court could make a conclusory statement like, “On the state of facts found, I conclude that the defendant is not liable.” Given an inarticulate fact finder, however, a court wishing to impose a standard without articulating the standard to the fact finder must keep, or take, the case away from the fact finder. Then the conclusory statement must be something like, “On no set of facts that could reasonably be found would the defendant be liable.”

I am not referring here to jury nullification, which is defined in THOMAS GREEN, VERDICT ACCORDING TO CONSCIENCE (1985), as “the exercise of jury discretion in favor of a defendant whom the jury nonetheless believes to have committed the act with which he is charged.” Id. at xiii. Jury nullification is conduct in conflict with judicial instructions; I am discussing here conduct implementing and supplementing those instructions.

Cf. Smith v. Allwright, 321 U.S. 649, 669 (1944) (Roberts, J., dissenting) (“The reason for my concern is that the instant decision, overruling that announced about nine years ago, tends to bring adjudications of this tribunal into the same class as a restricted railroad ticket, good for this day and train only.”).

I speak here of a civil plaintiff’s case, but the same points would hold for other parties against whom a case or issue might be decided without reaching the fact finder.
at least to the extent of determining what facts might reasonably be found, and law-determination. In an easy case, the court might not have to perform either function very precisely. An easy case would be one in which, on any view of the governing legal standard that the court finds plausible, it is clear either that the jury might reasonably find facts sufficient to support the plaintiff's claim, or that the jury may not so find. In such a case the court need not, for the sake of deciding whether the case belongs before the jury, determine just what the applicable legal standard is, or even the bounds of what facts the jury might reasonably find. In a harder case, the court might have to decide whether a combination of factual findings exists that the jury might reasonably make and that satisfies the legal standard the court prefers—but even then, the court need not, simply to determine whether the case is for the jury, articulate the standard it chooses. And, easy case or hard, the court need not even determine the legal standard for purposes of instructing the jurors, if it is satisfied to leave to them the retail work of deciding the standard for the particular case.

When a jury is asked to resolve an issue such as the "reasonableness" of the defendant's conduct, it, too, must perform both a fact-finding function and a law-determining one. In an easy case, it, too, may not need to perform either function with great precision; especially given the jury's inarticulateness, the two functions may blur together, in that the jury may be uncertain to what extent it is finding facts and to what extent it is determining standards. If, for example, the jury decides that the facts of the defendant's conduct lie within a given range, and that this entire range lies outside any plausible view of the zone of reasonable conduct, then it may decide that the defendant's conduct was unreasonable—without determining, much less articulating, precisely what that conduct was or what the standard of reasonableness is. In a harder case, the jury might have to be considerably more precise, both as to the facts and the governing standard. But, unless the jury is required to return an extraordinarily detailed special verdict, it need not be articulate at all.

This very lack of articulateness may indeed be attractive to the court; to the extent the court feels comfortable leaving the determination of the standard in the particular case to the jury, the court need not worry whether the standard actually chosen is consistent with those applied in other cases. Of course, this inarticulateness has a very substantial downside, for it may represent a partial abandonment of the goals of predictability and of decision according to understood principles.

In some cases, there may be more positive reasons—in contrast to the jury's black box nature—for leaving some of the task of selecting legal standards to the jury. For one thing, where the general principle is that prevailing community standards are to supply the norm, the jury

21 See supra note 17.
may be more likely than a court to act in accordance with those standards; "reasonable speed" and "ordinary care" are good illustrations.

Moreover, an efficient allocation of authority does not break down neatly along, or even necessarily very close to, the law-fact division. Thus, it often makes sense for general facts, which are significant across a range of cases, to be determined at the top of the pyramid, ensuring uniformity and obviating the need for repeated litigation.\(^2\) And, by the same token, it often makes sense for the particularized standard in a given case to be determined at the bottom of the pyramid by those most familiar with the case. Some narrow decisions of policy in most systems—political, enterprise, military, and so forth—are often made well below the top of the pyramid, and there is no reason why the same should not be true in judicial systems as well. Indeed, appellate courts sometimes rather explicitly recognize the value of devolution of this sort, by purportedly leaving certain questions to the discretion of the trial court.\(^2\) Leaving the choice of particularized standards to the jury rather than to the trial court accomplishes much the same efficiency benefit. It does so, however, without pressuring the trial court, for better or for worse, to articulate those standards and without creating the possibility, for better or for worse, that the standard chosen will become a precedent for future cases.

Once again, there is a substantial downside if the fact finder is too readily granted the authority to select the governing standard. It makes little sense for a jury, or for that matter a randomly selected trial judge, to decide broad issues of political, economic, and social policy. Past a point, appellate judges must ensure that, if such a policy decision is to be made within the judicial system, they, rather than the fact finder, will

\(^{22}\) Thus, it is generally understood that legislative facts—facts that go to determine what the law is—are for the courts to determine. See, e.g., FED. R. EVID. 201(a), advisory committee's note, 56 F.R.D. 183, 201-03 (1973). But as the statement of legal principles becomes more particularized, it is not always easy to determine whether a given fact should be deemed adjudicative or legislative. The difficulty may be that a court often has a choice of how to treat a given fact. For example, an appellate court might say, "Because we believe factual propositions A through N to be true, we hold legal proposition P to be true." On the other hand, it might hold that the trial court should instruct the jury, "If you find that propositions A through N are true, then you should act in accordance with legal proposition P." Either of these approaches might be appropriate in a given case; the choice will often depend on how recurrent propositions A through N are. The work of Professors Walker and Monahan on this general topic is very important. See, e.g., Laurens Walker & John Monahan, Social Facts: Scientific Methodology as Legal Precedent, 76 CAL. L. REV. 877 (1988); Laurens Walker & John Monahan, Social Frameworks: A New Use of Social Science in Law, 73 VA. L. REV. 559 (1987).

\(^{23}\) See, e.g., United States v. Paradise, 480 U.S. 149, 184 (1987) ("We have recognized that the choice of remedies to redress racial discrimination is a balancing process left, within appropriate constitutional or statutory limits, to the sound discretion of the trial court.") (quotation marks and citations omitted); Solem v. Helm, 463 U.S. 277, 290 (1983) ("Reviewing courts, of course, should grant substantial deference to . . . the discretion that trial courts possess in sentencing convicted criminals."); Piper Aircraft Co. v. Reyno, 454 U.S. 235, 257 (1981) ("The forum non conveniens determination is committed to the sound discretion of the trial court.").
make the decision—either by articulating the governing standard with sufficient precision or, particularly where that seems inadvisable or too difficult, by removing the case from the fact finder.24

Thus, the allocation of authority over law and fact is the product of various tensions. On the one hand, there is a clear analytical distinction between law and fact, and the allocation of law determination to the courts and fact-finding to the jury is (assuming the value of the jury in the first place) presumptively desirable. But this allocation is impossible to effectuate in pure form, largely because of limitations on the power of articulation. Moreover, even if courts were articulate enough to state a legal rule for every possible factual finding, they often would not be willing to do so. And often a question, though analytically factual, is sufficiently general that it is most efficiently decided by the courts, just as some questions, though analytically legal, are sufficiently particular that it may be best to leave decision of those questions to the fact finder.

Thus, I have no debate with the commonplace that the allocation of authority over law and fact is in large part a matter of convention or function. But that commonplace should not make us despair of finding any conceptual distinction between law and fact.

II. ATTEMPTING TO DETERMINE A STANDARD OF PERSUASION FOR PROPOSITIONS OF LAW

The essence of Lawson's argument is his contention that "[f]or any given proposition in any given context, one needs a standard of proof that expresses the total weight or magnitude of the evidence required for a justified assertion of that proposition."25 This requirement, he asserts, applies to propositions of law as well as to propositions of fact; indeed, a theory of interpretation is "radically incomplete and indeterminate" unless it includes a governing standard of proof.26

24 The rule of reason in antitrust provides an interesting illustration. In Continental T.V., Inc. v. GTE Sylvania Inc., 433 U.S. 36, 49 (1977), the Supreme Court said, "Under [the rule of reason], the factfinder weighs all of the circumstances of a case in deciding whether a restrictive practice should be prohibited as imposing an unreasonable restraint on competition."

The Court cannot have meant what it said, for that would mean that in any case decided by the rule of reason the jury or trial court sitting without a jury would in effect adjudicate not only the legality of the practice before it but also the entire panoply of economic, political, and social issues that might go into determining whether the restraint on competition is "unreasonable." Not surprisingly, appellate courts in cases under the rule of reason have, notwithstanding the open-ended language of Sylvania, exercised substantial control over the fact finder, see, e.g., Continental T.V., Inc. v. GTE Sylvania Inc., 694 F.2d 1132 (9th Cir. 1982) (on remand; affirming summary judgment for defendant), and at times they have attempted to make more particular statements of the rule, see, e.g., Valley Liquors, Inc. v. Renfeld Importers Ltd., 678 F.2d 742, 745 (7th Cir. 1982) (Posner, J.) ("The plaintiff in a restricted distribution case must show that the restriction he is complaining of was unreasonable because, weighing effects on both intrabrand and interbrand competition, it made consumers worse off.").

25 Lawson, supra note 1, at 869.

26 Id. at 874-75.
In this Part, I will express partial agreement with Lawson. Section A examines what I will call the descriptive mode, in which an actor attempts to determine the validity of legal propositions according to norms that already have been, or will be, established by other actors. I will begin with a simple model. Given the restrictions of that model, I agree with Lawson that the actor must, either explicitly or implicitly, incorporate into her thinking what I will call a standard of persuasion—a test of how probable the validity of a proposition must appear for her to treat the proposition as valid. But then, still remaining within the descriptive mode, I will introduce some complexities. I will argue that, unless the actor can make strong simplifying assumptions, she cannot pose her decision problem in terms of the determination of a single standard of persuasion. Lawson is correct to treat the validity of a legal proposition as a probabilistic matter. But the notion of a standard of persuasion—what he calls a standard of proof—does not always suffice. Instead, it is sometimes necessary, if one wishes to analyze a problem fully, to adopt a more complex approach to selecting the decision with the greater expected value.

Section B adds a further variation by examining the prescriptive mode, in which a legal actor attempts to determine what the law should be. In this context, descriptive uncertainty as to the state of the law may still enter into the decision, but it is augmented by uncertainty as to what the consequences will be of choosing one rule or another. Thus, again, except in limited circumstances, the court cannot choose a rule of law by testing alternative candidates against a single standard of persuasion. Instead, it must make the more complex determination of what rule optimizes expected value.

A. The Descriptive Mode

1. A Simple Model: Known Facts, Two Possible Legal Rules.—Let us begin with a simple model. An actor must decide, in a known factual setting, between two alternative courses of action. The only uncertainty is what the law is. The actor has no control over what the law is, but must attempt to discover it. There are only two candidates for the governing legal rule. If one rule governs, then the actor’s interests (broadly

---

27 Lawson expresses standards of proof in conventional probabilistic terms. Id. at 869. He also suggests that such standards might be expressed on an ordinal scale, by determining whether a material proposition is “as or more certain” than some yardstick proposition. Id. at 870. Such an ordinal scale, referring to levels of certainty, should be considered probabilistic in nature. I believe, though, that thinking in terms of ordinal probabilities sacrifices valuable information. For example, in comparing the expected values of two possible decisions, D₁ and D₂, one must take into account both the consequences of each contingency under each decision and the probability of those contingencies. If the consequences of the first contingency are mildly superior under D₁ and the consequences of the second contingency are vastly superior under D₂, it becomes important to assess not only whether the first contingency is superior to the second, but also by how much. See infra note 29.
defined to include altruistic and social interests recognized by the actor) will best be served by acting in one way, and if the other rule governs, those interests will best be served by acting in the other way.

In this setting, the actor faces a simple ordinary case of deciding under uncertainty, with a two-by-two matrix. The problem has the same structure as the one she faces in deciding whether to carry an umbrella to work on a cloudy morning. In that context, there are, simplifying somewhat, two possible relevant states of the world (rain and no rain), and she can make one of two possible decisions. There are thus four possible outcomes. Does she carry the umbrella, meaning she will have wasted effort if the weather stays dry but will be protected if it rains? Or does she leave the umbrella home, so that her walk will be pleasantly uncluttered if it does not rain but quite unpleasant if it does? The decision depends on the probability that she assigns to rain and the relative value of each of the four possible outcomes.

Similarly, in the simplified legal setting considered here, there are two possible decisions the actor can make, and two possible given states of the world—that is, legal rules. Suppose, for example, the actor's choices of conduct are ACT-PERMISSIVE and ACT-RESTRICTIVE, and the rule is either RULE-PERMISSIVE or RULE-RESTRICTIVE. The four possible outcomes are ACT-PERMISSIVE & RULE-PERMISSIVE, ACT-PERMISSIVE & RULE-RESTRICTIVE, ACT-RESTRICTIVE & RULE-PERMISSIVE, and ACT-RESTRICTIVE & RULE-RESTRICTIVE. The optimal decision thus depends on an evaluation of the payoff for each of these four outcomes, and the probabilities of the alternative possible rules of law. Put another way, the actor should conduct herself in a given way only if she regards the rule favoring that conduct as having a probability at least as great as a certain threshold level, and that level will be determined by the relative payoffs of each of the four possible outcomes.

If, for example, the actor is a municipal bond lawyer asked to decide the taxable status of a new offering, she will probably regard as devastating the consequences of treating the bonds as tax exempt if they are not, and in substantial part because of this valuation she will probably not act on the basis that the bonds are exempt unless she is very sure.

If, on the other hand, the actor is a litigator who is contending after

---

28 The decision depends on an evaluation of the positive as well as the negative outcomes. The evaluations are relative, so that without loss of generality one outcome may be assigned a unitary value, thus setting the scale for the others; nevertheless, the remaining three must be evaluated in relation to that one. For this reason, I prefer not to use the term "regret matrix." A regret matrix is a simplified form of utility matrix, see Richard Lempert & Stephen A. Saltzburg, A Modern Approach to Evidence 162 & n.23 (2d ed. 1982), in which only incorrect decisions are evaluated. Sometimes this may be a useful simplification, but it is important to remember that all four outcomes must be evaluated.

29 Here is a simple mathematical model of the problem. The actor can make either of two decisions, $D_1$ or $D_2$. The law is in either of two states, $L_1$ or $L_2$. The probability of state $L_1$ is $p_1$, and the probability of state $L_2$ is $p_2$, which equals $(1 - p_1)$. If the actor selects $D_1$ and the law is in state
the fact that the bonds are exempt and she is deciding whether to argue a particular proposition of law in her favor, she will probably decide to make the argument even if it has a relatively low chance of success; the incremental costs of making the argument are relatively low and the pay-off, if the argument is correct, is high.

A third actor is the trial judge. Suppose that she is not trying to be creative but is simply trying to determine "what the law is." Suppose further that—both for her socially oriented interests and for her desire to compile a good record in the appellate court—she places the same value on a correct decision for one party as on a correct decision for the other, and that the same holds true of incorrect decisions. This judge probably will select the rule of decision that she believes more likely than not represents the law.  

2. Multiple Possible Legal Rules.—Now let us complicate the simple model by relaxing one of its restrictive assumptions, that there are only two possible legal rules. Lawson correctly recognizes that a decision in this context cannot depend on selecting the one rule among many that seem most probably correct. Indeed, if the number of possible legal rules is very large—and if there is a continuous range of possible rules, the number of individual possible rules is infinite—the probability of each individual rule will tend to be very small, near zero.

The same problem affects the fact-finding process. There is an infinite number of possible stories that might account for the evidence, each varying only infinitesimally from some other stories; thus, I have characterized each possible story as being "infinitesimally thin." For exam-

L₁, then consequence C₁₁ follows, if D₁ and L₁, the consequence is C₁₂, and so forth. The expected value of making decision D₁ is then

\[(p₁)(C₁₁) + (1 - p₁)(C₁₂),\]

and similarly the expected value of making decision D₂ is

\[(p₂)(C₂₁) + (1 - p₂)(C₂₂).\]

D₁ is the better decision if the expected value of D₁ is greater than that of D₂, and this will hold if and only if

\[
\frac{p₁}{1 - p₁} > \frac{C₂₂ - C₁₂}{C₁₁ - C₂₁}.
\]

The fraction on the left-hand side of this inequality represents the odds that state L₁ is the law. The numerator of the fraction on the right represents the extent to which, assuming L₂ is the law, the consequences of decision D₁ are superior to the consequences of decision D₂. Similarly, the denominator of that fraction represents the extent to which, assuming L₁ is the law, the consequences of D₁ are superior to those of D₂.

30 In terms of the mathematical model presented in note 29 supra, the judge regards C₁₁ as equal to C₂₂ and C₁₂ as equal to C₂₁. Thus, the fraction on the right side of the inequality in note 29 equals 1. That is, for D₁ to be the better decision, the odds that L₁ is the proper legal rule must be 1 or greater.

31 Lawson, supra note 1, at 891-92.

a story in which the accused stabbed the victim at one second after 2 p.m., one millimeter below the victim’s left ventricle, differs in two respects from a story in which the accused stabbed the victim at two seconds after 2 p.m., two millimeters below the victim’s left ventricle—but the differences are not great. The fact finder’s job cannot be to select the single most likely story, or even to compare the single most likely story favoring one party with the single most likely story favoring the other party. Instead, the fact finder must group bunches of infinitesimally thin strands, and then compare these thicker threads together.

A similar approach might be used by an actor in the descriptive mode trying to determine which of multiple possible legal rules governs a known factual situation. Assume as before that only two possible courses of action, such as ACT-PERMISSIVE and ACT-RESTRICTIVE, are possible. We may categorize as a permissive rule any possible rule the consequences of which are better, on the known facts, given ACT-PERMISSIVE than given ACT-RESTRICTIVE. A corresponding categorization would yield a group of restrictive rules. If, on the known facts, all permissive rules have the same consequences as each other given either course of conduct, then we may aggregate these rules, effectively treating them, as applied to the known facts, as one rule; the restrictive rules may, subject to the same conditions, be similarly grouped.

\[ P(C_{1,n} - C_{2,n}) > 0. \] (*)

Now assume that, for all i between 1 and m inclusive, D, leads to a better result than does D, that is, \( C_{1,i} > C_{2,i} \), Correspondingly, assume that, for all j between m + 1 and n inclusive, \( C_{j,1} > C_{j,2} \), and that for all rules \( L_{1} \) through \( L_{m} \), \( C_{i,1} = C_{i,2} = \ldots = C_{i,m} = C_{i,1} \), and \( C_{2,1} = C_{2,2} = \ldots = C_{2,m} = C_{2,1} \), and that for all rules \( L_{m+1} \) through \( L_{n} \), \( C_{i,m+1} = C_{i,m+2} = \ldots = C_{i,n} = C_{i,1} \), and \( C_{2,m+1} = C_{2,m+2} = \ldots = C_{2,n} = C_{2,1} \).
This aggregating assumption will sometimes hold true. Suppose there are two permissive rules—rules that favor a permissive result on the known facts—but that the rules would draw the dividing line in substantially different places; the conduct in question would be well within the realm of tolerable conduct given one rule, but very close to the line given the other rule. It may be that the actor need not worry about this difference. That is, it may be that the consequences of the permissive conduct are the same whichever of these rules is the actual one, and that the same holds true of the restrictive conduct. For example, the bond lawyer preparing an issue may conclude that if the bonds are indeed exempt the consequences of her decision—positive if she has decided to treat them as exempt, negative if she has not—will be the same however close the legal question may be. The lawyer may take comfort from Justice Holmes' statement that "the very meaning of a line in the law is that right and wrong touch each other and that anyone may get as close to the line as he can if he keeps on the right side."\(^{36}\)

If this appears true in the given case, the lawyer would be able to aggregate all permissive rules. And if she is also able to aggregate all restrictive rules (and if all rules are either permissive or restrictive so that they fit into one aggregation or the other\(^{37}\)), then her choice between the two groups of rules is essentially as easy as in the simple model presented in section A, in which she chooses between two rules: There are the

\[
= C_{2n} = C_{2r}. \text{ Then rules } L_1 \text{ through } L_m \text{ may be grouped together, and so may rules } L_{m+1} \text{ through } L_n. \text{ Let } P = (p_1 + p_2 + \ldots + p_n) \text{, which is the probability of the disjunctive proposition } (L_1 \text{ or } L_2 \text{ or } \ldots \text{ or } L_m). \text{ This means that } (1 - P) = (p_{m+1} + p_{m+2} + \ldots + p_n), \text{ and the expression in (***) may be simplified to}
\]

\[
P(C_{ij} - C_{ij}) > (1 - P)(C_{ij} - C_{ij})
\]

which in turn may be expressed as

\[
\frac{P}{1 - P} > \frac{C_{ij} - C_{ij}}{C_{ij} - C_{ij}}.
\]

This inequality is the exact counterpart of the one in note 29 supra. In other words, given the aggregating assumption, the multiplicity of possible rules does not complicate the problem, once the two groupings of rules are formed. The circumstances in which \(D_i\) is the better decision thus may be expressed in terms of the odds of a single (albeit aggregate, disjunctive) proposition.

If, however, the aggregating assumption cannot be made, we are left with the more complicated condition in (**). The circumstances in which this condition are satisfied cannot be expressed in terms of the odds of a single proposition, even an aggregate, disjunctive one.


\(^{37}\) I am putting aside the possibility of neutral rules, which have the same consequences given either course of action, for such rules add more complexity than interest to the analysis. A neutral rule is an example of a rule that does not fit into either of two alternative aggregations of rules, and so prevents expression of a simple standard of persuasion in terms of the odds of one aggregation, as in note 35 supra. But because neutral rules do not favor either course of conduct, if the only rules that do not fit into either of the aggregations are neutral ones, the actor might still express a slightly altered standard of persuasion, in terms of the probability of one aggregation divided by the probability of the other aggregation. The sum of those two probabilities would be less than 1 because of the possibility of neutral rules.
same four possible outcomes, and the value of those outcomes determines how probable the permissive group of rules must be before the actor takes the permissive path.\textsuperscript{38}

But the aggregating assumption will not always hold. An actor may determine that missing by an inch is not as bad as missing by a mile. Suppose the actor is the trial court, attempting to determine in the face of substantial legal uncertainty the legal status of given conduct in which a defendant, civil or criminal, engaged. The court may decide that the consequences of incorrectly characterizing the defendant's conduct as illegal may be much less if the conduct is close to the line than if it is well within the realm that the law tolerates. There are at least two reasons for this.

First, it may be that the distance of the defendant's conduct from the dividing line reflects the net social benefit of the conduct. Thus, if the actual legal rule (the one that the court of last resort would apply) would barely tolerate the defendant's conduct, this may well be because the conduct is barely net socially beneficial, or perhaps even net socially detrimental. And if that is so, it suggests that the social detriment of incorrectly penalizing the conduct—putting aside the defendant's reliance interest—is relatively small, if it exists at all.

Second, if the law really is unclear, the extent of the defendant's reliance interest may depend in part on how far the conduct lies from the line that the court of last resort would draw, if it were to draw a clear line. The reliance interest is strong, and the negative consequence of violating it is great, if that line is such that the defendant's conduct clearly lies on the good side of it. But if the location of the line is unclear, so that a fuzzy area exists in which conduct might be on the good side of the line and might be on the bad side, conduct in that fuzzy area creates a much weaker reliance interest, even if eventually it is determined that the conduct does lie on the good side.

Thus, the aggregating assumption does not necessarily hold. When it does not, the actor cannot treat all permissive rules as a group and all restrictive rules as another group. And if that treatment is impossible, then, in deciding whether to take the permissive course or the restrictive one, the actor cannot pose the issue in terms of how probable the permissive group of rules must be, as compared to the restrictive group, for the permissive course to be preferable.\textsuperscript{39} That is, in this context it does not work to speak in terms of a single standard of persuasion for a given legal proposition, even an aggregated proposition. Instead, the actor must make a far more complex determination: Taking into account the probability of each potential rule, and the consequences of that rule (al-

\textsuperscript{38} For a reasonably simple mathematical demonstration of this point, see \textit{supra} note 35.

\textsuperscript{39} The point is presented in mathematical terms in note 35 \textit{supra}. 

931
ternatively under one course of action and the other), which course of action has the greater expected value?

If there is a substantial number of potential rules, this question is impossible for a human actor to answer without the aid of simplifying heuristics. Perhaps actors often use the particularly powerful heuristic of treating all permissive rules as the same and all restrictive rules as the same.\textsuperscript{40} That would eliminate the computational complexity nearly altogether and would allow the actor to pose her question in terms of what standard of persuasion the permissive group of rules must satisfy to make the permissive course of conduct preferable. But this heuristic would also disregard much of the complexity of the situation, and so might lead to results that do not closely reflect the actor's evaluations of the potential consequences of her actions.

3. Factual Uncertainty, With a Unitary Decisionmaker.—So far in this Part, I have assumed that the facts have been determined and are known with certainty by the legal actor, leaving only the law uncertain. That is a great simplification, of course. Sometimes the legal actor must act before the facts are determined. In any event, as I have emphasized,\textsuperscript{41} the facts can rarely be determined with anything close to certainty. The legal actor must therefore take into account not a single factual setting but a range of possible settings.

Let us first examine the problem assuming that a unitary actor determines both law and fact. If the actor confronts a broad range of both legal and factual possibilities, she must deal with a very complex matrix. One rule may favor the permissive conduct only under certain very narrow factual settings, another may be slightly more generous, another yet again slightly more generous, and so forth.

Assume as before that only two possible courses of action, one permissive and the other restrictive, are open to the actor. Her task as before is to select the one that has the higher expected value. She might first sort into two groups all possible combinations of factual setting and legal rule—a group of permissive combinations, meaning all those in which the consequences will be better if the actor chooses the permissive conduct, and a correspondingly defined group of restrictive combinations. As in subsection 2, in which the actor assumed factual certainty, it may be that she can make an aggregating assumption:\textsuperscript{42} If all permissive combinations of fact and rule have the same consequences as each other given either the permissive or the restrictive course of conduct, then she may effectively treat the entire group of combinations as one; the restrictive combinations may, subject to the same conditions, be similarly aggregated.

\textsuperscript{40} I do not claim to have any psychological insight on this point.

\textsuperscript{41} See supra Part I and also subsection 2 of this section.

\textsuperscript{42} See supra text accompanying note 35.
If this aggregating assumption is correct, then the actor may determine a level of probability that states how likely it must be that the actual combination of factual setting and legal rule falls within the permissive group to make the permissive course of conduct better.\textsuperscript{43} Note that this is a standard of persuasion on the combination of factual setting and legal rule—\textit{not} on either component taken separately.

Moreover, it may very well be that the aggregating assumption is not valid. Subsection 2 pointed out problems with that assumption in the simpler situation in which the factual setting could be taken as a given. Those problems are likely to become substantially more acute given factual uncertainty. In particular, it may be that consequences of a course of conduct that would only be mild given one set of facts near the line drawn by a plausible legal rule become much stronger given another set of facts much farther away from that line.

\section*{4. Law-determining and Fact-finding Functions Divided.---This subsection adds one more complexity to the descriptive mode: The law-determining and fact-finding functions are divided, so that the trial court must perform the function of determining what the law is, and the jury must make the necessary factual findings.}

If either the jury or the trial court were perfectly articulate, this

\textsuperscript{43} The mathematical model presented earlier can be extended to this situation. Assume as before that there are $n$ possible legal rules $L_i$, with $i$ being an integer between 1 and $n$ inclusive and $p_i$ being the probability that $L_i$ is the actual rule. Assume also that there are $r$ possible factual settings $F_h$, with $h$ being an integer between 1 and $r$ inclusive and $q_h$ being the probability that $F_h$ is the actual rule. We may treat the $q$ and $p$ terms as independent; what the abstract rule of law is does not depend on what the actual facts are in the given case. $C_{i,h}$ represents the consequence of selecting decision $D_1$ given legal rule $L_i$ and factual setting $F_h$. The expected value of $D_1$ is therefore represented by this matrix:

\[
\begin{array}{c}
p_1q_1C_{1,1} + p_1q_2C_{1,2} + \ldots + p_1q_rC_{1,r} \\
+ + + + + \\
p_2q_1C_{2,1} + p_2q_2C_{2,2} + \ldots + p_2q_rC_{2,r} \\
+ + + + + \\
\vdots \\
+ + + + + \\
p_nq_1C_{n,1} + p_nq_2C_{n,2} + \ldots + p_nq_rC_{n,r}
\end{array}
\]

A corresponding matrix (but with 2 in place of 1 as the first subscript of each of the $C$ terms) would represent the expected value of $D_2$. If the aggregating assumption stated in the text is true, then (1) for all combinations of legal rule $L_i$ and factual setting $F_h$ for which $D_1$ is preferable, (a) $C_{i,h}$ equals some constant $C'_1$, and (b) $C_{i,h}$ equals some constant $C'_2$, and (2) for all combinations for which $D_2$ is preferable, (a) $C_{i,h}$ equals some constant $C'_1$, and (b) $C_{i,h}$ equals some constant $C'_2$. Now let $P$ equal the probability that the actual combination of legal rule and factual setting is one for which $D_1$ is preferable. Then it could be shown, by means similar to those used in note 35 \textit{supra}, but much more extensive in notation, that the expected value of $D_1$ is greater than the expected value of $D_2$ if and only if

\[
\frac{P}{1 - P} > \frac{C'_2 - C'_1}{C_1 - C_2}.
\]

This is the exact counterpart of the final inequalities in notes 29 & 35 \textit{supra}.
division would not create much additional complexity. Suppose first that the jury could articulate just what probability distribution of possible factual settings it has found. The court’s task would then be less complex than that of the unitary actor in subsection 3, because the fact-finding aspect of that actor’s task would already be performed. Correspondingly, suppose that the court could articulate each possible distribution of factual possibilities. For each such distribution, the court would then have to determine the preferable course of action, in the manner suggested in subsection 3. It would then have to articulate to the jury what result the jury should reach for each distribution of factual possibilities.

The problem, of course, is one highlighted in Part I: Neither judge nor jury can articulate information disaggregated to such a high degree. In Part I, I suggested one aspect of a court’s likely response to this problem. The court may, rather than fully articulating the legal standard, leave to the jury much of the work of transforming general norms into particularized rules.\(^4\)

But this leaves unsolved the problem suggested by Lawson: If the court is uncertain about what the law is, then, to the extent that the court does articulate the law, what standard should it use for resolving its uncertainty? I think this problem is probably unsolvable without some powerful simplifying heuristics.

The court might use an aggregating heuristic similar to the one suggested in subsection 3, assuming that all possible combinations of fact and rule fall within one of two groups, and that within each group every combination has the same consequences as every other combination given either of the two possible results that the adjudication might achieve. But recall that a heuristic of this type yields a standard of persuasion only for a combination of factual setting and legal rule, not for either component taken separately.

Thus, while this heuristic helps, it does not, except in certain situations, allow the court to determine how confident it must be in the validity of a given proposition to act on that proposition as the law. If all the fact-law combinations that favor one result involve one legal rule and all the combinations that favor the other result involve the other rule, then there is really no question of fact at all to be resolved; the court can assume a given state of facts and determine the governing legal rule in the manner suggested in subsection 3. Correspondingly, if all the combinations that favor one result involve one factual setting and all the others involve another, then it does not much matter just what the court believes the law is; the only question that matters is how the jury resolves that factual issue.

The case may not be so neat, however. The court might believe, for example, that, given one plausible combination of factual setting and

\(^{4}\) See supra pp. 923-24.
Standards of Persuasion

legal rule, the restrictive result is preferable, but that given all other combinations the permissive result yields better consequences.

If the factual setting that may support the restrictive result seems likely, the court might simply consider whether, assuming those facts are true, the restrictive rule is sufficiently probable to make the restrictive result preferable; if it is, then the court would ask the jury to determine whether those facts are indeed true. This approach might be practical, because it first resolves legal uncertainty and then factual uncertainty, but this also means that the approach avoids some of the complexity of reality.

Alternatively, one could adjust the standard of persuasion on the facts: The more confident the court is of the restrictive rule, the less confident the jury must be that the facts fall within that rule. This would be a rather unusual and unstable use of a standard of persuasion; it would not be surprising, though, if courts implement something very much like this sliding scale in determining whether a case should be presented to the jury at all.

If this all seems quite complex, I do not deny it. But the complexity, I believe, is not a gratuitous result of the analysis presented here. Rather, it is the result of a complex situation in which a decision must be based on both factual and legal perceptions, each of which might in itself be very complex and uncertain, and the decision making authority is not unified.

B. The Prescriptive Mode

In section A, I focused on the descriptive mode, in which the legal standards are determined by someone other than the actor in question. In this section, I turn to the prescriptive mode, in which the actor—I will assume that it is a court of last resort—attempts to determine the optimal state of law.

First, assume that the facts are known. In the descriptive mode, an actor choosing between alternative courses of conduct has to assign a probability to each possible legal rule, and determine the consequences

\[\begin{align*}
\text{Suppose the following: Given legal rule L}_1 \text{ and factual setting F}_1, \text{ the consequences of decision } D_1 \text{ are C}_1 \text{ and the consequences of decision } D_2 \text{ are C}_2, \text{ with C}_1 \text{ greater than C}_2. \text{ Given all other combinations of legal rule and factual setting, the consequences of D}_1 \text{ are C}'_1 \text{ and the consequences of D}_2 \text{ are C}'_2, \text{ with C}'_2 \text{ greater than C}'_1. \text{ The probability of L}_1 \text{ is p}_1 \text{ and the probability of F}_1 \text{ is q}_1. \text{ Then, by using the condition in note 43 supra, } D_1 \text{ is preferable if and only if}
\end{align*}\]

\[
\frac{p_1q_1}{1-p_1q_1} > \frac{C'_2 - C'_1}{C_1 - C_2}.
\]

By relatively simple manipulations, this condition may be expressed as

\[q_1 > \frac{C'_2 - C'_1}{p_1(C_1 - C_2 + C'_2 - C'_1)}.
\]

Thus, the greater \(p_1\) is, the smaller \(q_1\) need be to make \(D_1\) the preferable decision.
for each possible rule of each possible course of conduct. In the prescriptive mode, the court does not have to assign a probability to a given rule; the court’s course of conduct consists of choosing the rule.

This does not mean that uncertainty does not enter into the prescriptive choice of a rule. In fact, as I will now argue, the consequences of any given possible rule may be very uncertain, and the court must take this uncertainty into account in selecting a rule. But the eventual question that the court must answer is not how probable it is that a given rule is the law; the court is making the law. And, except under very restrictive conditions, the court cannot even pose the question in terms of how probable it is that a given rule is the best rule. Instead, the court must ask which rule has the greatest expected value.46

The consequences of a given rule may be uncertain, even assuming certainty as to the facts of the particular case, in at least two respects. Most obviously, the court may be uncertain how the rule would actually affect interests or conduct outside the courtroom.47 Another aspect of uncertainty may stem from an element of description that might enter into the prescriptive decision. The court may believe that norms previously established—perhaps by constitutional, statutory, or regulatory text, or by judicial precedent, or even by natural law—prescribe a rule of decision, or a range of rules of decision, for the case. If so, the court probably will assign some positive value, not necessarily dispositive, to deciding in accordance with those norms. But the court may, like the actors discussed in section A, be uncertain just what those rules are.

However the uncertainty of the consequences of a rule may arise, the court must take that uncertainty into account in evaluating the rule. Suppose, to take a simple case, that the court’s choice is simply between two rules,48 and that there is just one uncertain contingency, which has two possible states, termed affirmative and negative. Given the affirmative state, one rule is better, and given the negative state the second rule is preferable. In such a case, the court might easily determine how probable the affirmative state must be to make the first rule preferable.49

Usually, however, the world is not that simple. One problem, of course, is that there may be many candidates, not just two, for the legal

---

46 I put aside in this analysis, as I have throughout this Essay, the difficulty of measuring different consequences that may appear to be incommensurable. That, of course, is a very serious difficulty, but it is beyond the scope of this Essay.

47 The uncertainty may arise from uncertainty as to some aspect of the state of the world that is not involved in the particular case, or that has ramifications for application of the rule beyond that case; in traditional terms, the court may be uncertain about material legislative facts. The distinction between adjudicative and legislative facts is addressed briefly in note 22 supra.

48 A somewhat looser assumption would be that the choice is between two groups of rules, all the members of each group having identical consequences, at least so far as is apparent from the case at hand.

49 Suppose the probability of the affirmative state of the contingency is $p$, that the consequence of selecting legal rule $L_1$ is $C_{11}$ given the affirmative state and $C_{12}$ given the negative state, and that
rule. Even if just one contingency affects the expected value of each rule, it seems a singularly inefficient approach to determine for each rule-to-rule comparison the probability of the contingency that would make the first rule preferable, and then determine whether the probability is above that threshold. It is better to assess the probability of the contingency and determine which rule, given that probability, has the greatest expected value.

Moreover, there will usually be more than one contingency that affects the expected value of some, or all, of the possible rules. Those contingencies will probably affect the value of the rules to different extents; furthermore, the consequences of various combinations of contingencies may not be a simple addition of the consequences of the individual contingencies. Thus, multiple contingencies cannot usually be aggregated and treated as one contingency. There may be numerous circumstances—numerous combinations of probabilities of the various contingencies—that might make one rule preferable to another. Therefore, one cannot express the choice between rules in terms of the threshold probability of a single contingency or aggregated group of contingencies. Instead, the court must assess the likelihood of each contingency and the value of each possible rule under each contingency.

In other words, if the situation entails even a fair amount of complexity, there is no substitute for the bottom-line determination of which rule has the greatest expected value. I certainly agree with Lawson that selection of that optimal rule requires taking uncertainty into account. But, unlike him, I do not believe that the court's conclusion should usually be phrased in terms such as, "We believe that it is at least x% probable that this rule is the best." Instead, the court should say, in effect,

the consequence of selecting legal rule L₂ is C₂, given the affirmative state and C₂, given the negative state. The expected value of selecting L₁ is then

\[ p_1(C_{1,1}) + (1 - p_1)(C_{1,2}) \]

and similarly the expected value of selecting L₂ is

\[ p_1(C_{2,1}) + (1 - p_1)(C_{2,2}) \]

These are exactly the same expressions of expected value as in note 29 supra, and so the conditions in which L₁ is preferable may be expressed in the same terms:

\[ \frac{p_1}{1 - p_1} > \frac{C_{2,2} - C_{1,2}}{C_{1,1} - C_{2,1}} \]

This congruence is not surprising. In the simple model, presented in note 29 supra, of decision making by an actor in the descriptive mode, the actor had to choose between two available courses of action given a contingency—the legal rule—that had two possible legal states. In the model presented here, the actor must choose between two available legal rules given some other contingency that has two possible states. The problems facing the actor in the two situations are therefore analytically similar, though the symbols have somewhat different meanings in the two models.

50 This is a statement of cardinal probability. Lawson would also allow standards of persuasion to be expressed on an ordinal scale, but such expressions would still be probabilistic in nature. See supra note 27.
“Taking into account all possible contingencies, and the probabilities of those contingencies, we believe that this rule is the best.”

III. THE STANDARD OF PERSUASION IN DETERMINING THE ELEMENTS OF A CRIME

In this Part, I will narrow my focus from the generalities that have dominated Parts I and II and offer some tentative thoughts on an interesting conundrum that Lawson poses. A criminal defendant cannot be convicted constitutionally unless the fact finder concludes that, for each of the elements of the crime, facts constituting that element are proven beyond a reasonable doubt. This requirement reflects the fundamental and longstanding view that the negative consequences of an incorrect conviction are so great that they dwarf in magnitude the consequences of the other three possibilities—a correct conviction, a correct acquittal, and an incorrect acquittal. But, then, what about legal uncertainty? Suppose, to simplify matters somewhat, that the relevant facts are stipulated, or otherwise admit of no substantial uncertainty, but that the applicable legal standard is uncertain. Should conviction be deemed proper even if the court is not persuaded beyond a reasonable doubt that the legal standard is broad enough to include the given facts? To the extent, if any, that the answer is affirmative, why should courts treat legal uncertainty differently from factual uncertainty?

First, let us focus on the descriptive mode. Again for simplicity, assume that only one element of the crime is in question. Also for simplicity, assume that there are only two candidates for the definition of the element, one permissive and the other restrictive. (Alternatively, if there are more than two candidates, assume that they may be aggregated into two groups in the manner discussed in Part II, with each group being treated essentially as one definition.) For the trial court to act in accordance with the restrictive definition, the one that would yield a conviction, how sure must the court be that this is descriptively correct?

This strikes me as quite a difficult question. On the one hand, mistakenly branding an accused as a criminal is not clearly worse when the mistake lies in determining that the accused did something he did not in fact do than when the mistake lies in determining incorrectly that what the accused concededly did was criminal conduct. Either case yields the detested result of convicting the accused even though his conduct lay outside the realm proscribed by the law. For the accused, the two types of errors have the same effect; for future cases, the error of law is more dangerous.

51 See 4 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 358 (18th ed. 1829) (“[T]he law holds, that it is better that ten guilty persons escape, than that one innocent suffer.”).
52 See supra note 43 and text accompanying notes 42-43.
At the same time, however, it seems impractical to insist that the legal proposition supporting conviction be descriptively accurate beyond a reasonable doubt. Often the law is very open-textured, or fuzzy; there may be no precedents on point or even close. Thus, even though the trial court has some confidence that a given proposition of law is descriptively accurate, it may be unable to say, no matter how exhaustive an inquiry it makes, that the proposition is accurate beyond a reasonable doubt. One problem confronting the trial court may be that the law governing the case, if litigated to the end, would be whatever the court of last resort determined prescriptively to be the better rule, and the trial judge may not have perfect insight into the psychology of the judges on the high court. It seems intolerable to say that in all such cases the trial court ought to enter a judgment of acquittal, preserving the possibility of appeal by the prosecution and eventual conviction only, if at all, by withholding the judgment until the facts are determined.

Indeed, it appears to me that if the facts are determined against the accused and the trial court is reasonably confident—it is hopeless to attempt to quantify this standard—that the legal standard necessary to uphold the conviction is descriptively accurate, the court ought usually to enter the conviction, leaving it to the accused to appeal. Even if it is true that final conviction because of legal error is as bad as final conviction because of factual error, the same is not true of the initial entry of judgment. An error of law in the trial court is less dangerous because it can be corrected more easily.

The trial court should not rely, however, on the appeals process to correct errors that easily might have been prevented below. Perhaps, therefore, the trial court ought to supplement its "reasonable confidence" standard with an inquiry meant to preserve any reliance interest of the accused, just in case this is not an instance of a fuzzy law: As a prerequisite to entering a judgement of conviction, the court must be persuaded beyond a reasonable doubt that, assuming it made an exhaustive inquiry and analyzed the relevant materials correctly, it would not find that the permissive rule contended for by the accused is clearly a descriptively accurate statement of law.

In the hard cases, though, in which the law has not crystallized, the important work must be done by the courts that make the law—that is, by courts that act prescriptively. Suppose, then, that the trial court has entered a conviction on stipulated facts and that the accused has appealed to the court of last resort.

Prescriptive as that court may be, part of its function will still be descriptive. If the law has crystallized in the accused's favor, the defend-
ant will have a strong reliance interest against a retroactive expansion of
the zone of criminal conduct. Even if the law has not crystallized, the
court might be attempting in part to transform previously determined
general norms into operational particular rules of law. To the extent this
is the court's function, the court should weigh very heavily the danger of
branding the accused's conduct as criminal even though it lies outside the
area covered by those norms. Thus, the principle of lenity in construc-
tion of criminal statutes should have significant force.\textsuperscript{54}

Part of the court's task, though, will be purely prescriptive, to deter-
mine what the guiding norms should be. To the extent the court (or, for
that matter, the legislature) is engaged purely in prescription, the court's
chief concern in the descriptive mode—that it will wrongly brand the
accused as a criminal even though under a descriptively accurate state-
ment of the law his conduct was lawful—simply disappears. In the pre-
scriptive mode, we may still speak of correct and incorrect decisions, if
we wish, but the terms have different meaning; an incorrect decision
against the accused, for example, is one that categorizes the conduct as
criminal even though under a preferable norm the conduct would be
deemed innocent. In this sense, the negative value of an incorrect deci-
sion against the accused is not clearly greater in magnitude than that of
an incorrect decision in his favor. Nor is the positive value of a correct
decision against the accused (one that accurately reflects preferred
norms) clearly less than that of a correct decision in his favor.

Indeed, in some cases the court might well determine that a broad,
prophylactic criminal prohibition is preferable. It may be far better to
treat some conduct as criminal, even though that conduct seems on bal-
ance to be harmless or even net socially beneficial. It may be, for exa-
ample, that, although the court believes the conduct is net beneficial, the
court is risk averse and recognizes a plausible possibility that the conduct
is actually quite devastating; assuming that most people will comply with
the law, the principal downside of an overly broad definition of criminal
conduct might be to inhibit some activity that would be tolerable or even
beneficial, but the downside of an overly narrow definition might be to
tolerate activity that causes grievous harm. Or we may prefer, for pur-
poses of simplifying the rules of criminal law, to use a bright line rule
drawn well away from the fuzzier line that divides beneficial and detri-
mental conduct. Or it may be that law enforcement officials would have
so much more trouble detecting violations of a narrow proscription that
the broader proscription is favorable.

If we relax one of the simplifying assumptions I have imposed,
under which there are only two candidates for the applicable legal rule,
the point is even stronger. A full and delicate palette of shadings in word
choice is available for the expression of rules. Suppose, therefore, that

\textsuperscript{54} See Lawson, supra note 1, at 888.
the court, perhaps in attempting to set out a standard that will govern in varying factual settings, is choosing from among a wide range of possible rules, each only infinitesimally different from some of the others. Given this range of choice, it does not appear possible for the court to insist that it will choose one rule over its more permissive neighbor only if it is persuaded beyond a reasonable doubt that the first rule is preferable. The court is engaged in an optimizing task, and as it nears its decision it will likely have to select among standards that are very close together. Ideally, the court should feel that the standard it selects is preferable both to one a little more restrictive and to one a little more permissive, but the differences may not be dramatic or clear.

I have also conveniently pushed aside until now one other complexity raised by the discussion in Part I. Confronting it requires a return to the trial court. If, as I have contended, part of the prescription of standards controlling a particular case is actually performed by the jury, does that mean that the jury should not convict unless it is persuaded beyond a reasonable doubt of the validity of the standards it would prescribe against the accused? Perhaps, to the extent that the task of prescription is left to the jurors, they should approach it in the same way that the court of last resort does. And perhaps that in fact is what they do. Indeed, perhaps that is inevitably what they must do; to the extent that the jurors are engaged purely in prescription, they, no less than the court of last resort, may have to select the optimum standard from a range of possibilities, and that is a task that ordinarily does not allow decisions leaving no reasonable doubt.

Even if these musings are correct, I am not sure that they have much operational significance. The courts leave a law-defining function to the jury in large part, I have argued in Part I, because of limits on articulation. Courts, not to mention juries, are often uncertain to what extent the jury is finding facts and to what extent the jury is determining legal standards. It will hardly do, therefore, to enunciate an instruction like the following:

To the extent you are determining facts, you may not find the accused guilty unless you find that facts necessary to support each element of the crime are proven beyond a reasonable doubt. To the extent you are determining the governing standard of behavior, you should not reach a verdict of guilty under any standard unless you are confident beyond a reasonable doubt that the standard is within the general limits of the law that I have outlined for you, but within those limits you should select the standard that in your view optimizes what you deem to be the various interests at stake.

Some things are better left unsaid.

55 See supra pp. 923-24.
IV. Conclusion

Professor Lawson has made a substantial contribution by presenting the question of standards of persuasion on propositions of law. My overall conclusion is that the matter is very complex, perhaps far more than Lawson recognized, and that in some circumstances one cannot even speak meaningfully of a standard of persuasion; instead, one must ask the substantially more inclusive, and more complex, bottom-line question of what course of conduct has the greatest expected value. Implementation of a carefully articulated approach to that question, it seems to me, is not usually possible in the real world. We may be more comfortable, therefore, returning the entire matter to the closet that Lawson has dared to open. But for the sake of full understanding it is better that the problem be unveiled, and we should shed as much light on it as we can.