

Fall 1987

Decision Theory and Due Process: A Critique of the Supreme Court's Lawmaking for Burdens of Proof

Richard S. Bell

Follow this and additional works at: <https://scholarlycommons.law.northwestern.edu/jclc>



Part of the [Criminal Law Commons](#), [Criminology Commons](#), and the [Criminology and Criminal Justice Commons](#)

Recommended Citation

Richard S. Bell, Decision Theory and Due Process: A Critique of the Supreme Court's Lawmaking for Burdens of Proof, 78 J. Crim. L. & Criminology 557 (1987-1988)

This Criminal Law is brought to you for free and open access by Northwestern University School of Law Scholarly Commons. It has been accepted for inclusion in Journal of Criminal Law and Criminology by an authorized editor of Northwestern University School of Law Scholarly Commons.

DECISION THEORY AND DUE PROCESS: A CRITIQUE OF THE SUPREME COURT'S LAWMAKING FOR BURDENS OF PROOF

RICHARD S. BELL*

I. DECISION-THEORETICAL FOUNDATIONS

In its 1987 decision in *Martin v. Ohio*,¹ the United States Supreme Court again engaged in lawmaking for burdens of proof. Although the Court ratified the state's rule in this case,² the Court has, in the past, prescribed different rules as requirements of due process.³ The Court has justified its lawmaking by asserting that the burden of proof should apportion the risks of error in a way that favors the more important interests at stake in the trial.⁴ If an erroneous finding of fact F would harm one set of interests more than an erroneous finding of not-F would harm the other, the burden of proof should lie with the party who alleges F, and the standard of proof for F should be high enough to reduce the risk of an erroneous finding proportionately. This idea is founded in Bayesian decision theory,⁵ in which it is formalized as a rule for choice in conditions of uncertainty. It is a cornerstone of the Court's

* J.D., Yale University, 1973; M.A., Yale University, 1973; B.A., Northwestern University, 1964. Preparation of this Article was facilitated by the generosity of Jeanne Bauer.

¹ 107 S. Ct. 1098 (1987).

² Under Ohio law, the accused bears the burden of proving self-defense by the preponderance of the evidence.

³ *Santosky v. Kramer*, 455 U.S. 745 (1982) (requiring proof by clear and convincing evidence for termination of parental rights); *Addington v. Texas*, 441 U.S. 418 (1979) (requiring proof by clear and convincing evidence for involuntary civil commitment); *Mullaney v. Wilbur*, 421 U.S. 628 (1975) (requiring that the prosecution bear the burden of proof on the issue of provocation); *In re Winship*, 397 U.S. 358 (1970) (requiring proof beyond a reasonable doubt of the elements of the crime).

⁴ See *Santosky*, 455 U.S. at 755-56; *Addington*, 441 U.S. at 423-24, 427; *Patterson v. New York*, 432 U.S. 197, 208 (1977); *Mullaney*, 421 U.S. at 698, 700-01; *Winship*, 397 U.S. at 363-64.

⁵ For a thorough treatment of the decision-theoretical underpinnings, see R. JEFFREY, *THE LOGIC OF DECISION* 1-76 (1965).

due process doctrine for burdens of proof, and merits careful examination.

This Article begins with a description of the Bayesian foundations, which explain the Court's logic. It continues with a critical examination of the elements that give substance to the logic—a policy for errors and assumptions about factfinders' evaluations of evidence in trials of fact. The Court's reasoning about both of these elements has been deficient. Next, this Article offers a more acceptable set of assumptions and introduces a more satisfactory model of lawmaking for burdens of proof. This model, as applied to several due process issues, makes it plain that the Court's lawmaking has been unsound, if not erroneous. Finally, this model is recommended as appropriate for testing burdens of proof against due process requirements.

A. THE KAPLAN FORMULA

The intellectual origin of the Court's due process doctrine for burdens of proof appears to be Kaplan's article, *Decision Theory and The Factfinding Process*.⁶ It starts with the assumption that evidence about an issue of fact will never leave a rational trier of fact in a state of perfect certainty. The trier will be able to conclude only that the truth of a party's assertion is probable at some value of p between 0 and 1.⁷

⁶ Kaplan, *Decision Theory and the Factfinding Process*, 20 STAN. L. REV. 1065 (1968). See *Winship*, 397 U.S. at 370-72 (Harlan, J., concurring). See also *Patterson*, 432 U.S. at 208 (citing *Winship*, 397 U.S. at 372 (Harlan, J., concurring)); *Mullaney*, 421 U.S. at 701; *id.* at 706 (Rehnquist, J., concurring).

⁷ This assumption may seem too strong in one respect. Perhaps a factfinder at trial cannot assign a numerical value or even a qualitative value to the probability of some "facts," because of the logic of the concept rather than a lack of information. For example, the question of common law negligence is said to be a question of fact, but only because it is a question for the trier. A finding of negligence results from the application of a legal rule or standard to facts, such as facts about the defendant's conduct. See generally O. HOLMES, *THE COMMON LAW* 97-103 (M. Howe ed. 1963).

Although there is no difficulty in applying the concept of probability to the facts about defendant's conduct, the concept has no clear application to a finding that this conduct was or was not negligent. The trier's doubt about whether the defendant's conduct was negligent cannot sensibly be treated as doubt about likelihood. If the trier is in doubt, it is because the facts are on the margin of his conception of reasonable care. The trier must *decide* whether the defendant who acted in this way used reasonable care by applying his understanding of what is reasonable. This is not a decision that further information could help, as further evidence would help in finding a fact, unless such further information is about the defendant's conduct. Nor could anything show that the trier's decision is wrong. Another decision of some higher authority could override it, but only in adopting a standard different from the one applied by the trier. Further-

On this assumption, Kaplan and others⁸ have suggested that decision theory can provide a rule for deciding whether the trier should find the assertion true. Suppose that the outcome of a trial is determined by the trier's finding on *P*'s (plaintiff's or prosecution's) factual allegation *A* or, correspondingly, *D*'s (defendant's) factual allegation not-*A*. The trier's finding may be either right or wrong,⁹ and, if wrong, it may be an erroneous finding for *D* or an erroneous finding for *P*. Either error would be unwanted, and so it has for the trier an expected disutility, which equals the error's disutility multiplied by the probability of its occurrence. Let the disutility of an error for *P* be *D_p*, and the disutility of an error for *D* be *D_d*. Then, assuming that the trier wants to minimize the expected disutility of his finding, he should find for *P* just in case his estimate of the probability of *A* (*p_A*) exceeds 1 divided by $(1 + D_d/D_p)$.

The Kaplan formula¹⁰ is an arithmetical expression of the Court's due process doctrine for burdens of proof. It seems to provide a foundation for standards of proof that is as rational as decision theory itself. Moreover, as Kaplan points out, his analysis seems to be supported by the standards of proof actually used in trials of fact.¹¹ For example, if, in the ordinary civil trial, an erroneous finding for one party is no worse than an erroneous finding for the other party, then $D_p = D_d$. Assign each variable a disutility value of 1, and the Kaplan formula indicates that the trier should

more, it would be nonsensical to ask the higher authority about the probability that this standard is true and the trier's standard is false.

The question of negligence, therefore, raises difficulties about more than the relation of judge and jury, which Holmes discussed. It raises difficulties for anyone who would treat all questions for triers as questions of fact to which answers of probability can be given. Indeed, it raises difficulties for anyone who would suppose that the common law standards of proof are applicable to the issue. See Bohlen, *The Effect of Rebuttable Presumptions of Law Upon the Burden of Proof*, 68 U. PA. L. REV. 307, 315-16 (1920).

The assumption that factfinders can assign some numerical value to the probability of a fact must be taken literally. Trials of the issue of negligence fall within the assumption only if the actual conduct of someone is in dispute and the standard of reasonable care is given by the doctrine of negligence per se, a rule of law, or a clear application of the Hand formula. See *United States v. Carroll Towing Co.*, 159 F.2d 169 (2d Cir. 1947).

⁸ See Cullison, *Probability Analysis of Judicial Fact-Finding: A Preliminary Outline of the Subjective Approach*, 1969 U. TOL. L. REV. 538.

⁹ Of course, the trier's finding, if reasonable, is by convention always correct. There is good reason for this convention. See Ball, *The Moment of Truth: Probability Theory and Standards of Proof*, 14 VAND. L. REV. 807, 808 (1961). But institutional fact may not correspond to actual fact. Some theoretical discussions ignore this obvious truth. G. TULLOCK, *THE LOGIC OF THE LAW* 57 (1971). Furthermore, people may even deny it for various causes. See *id.* at 37. Such an approach is fatal to an understanding of burdens of proof.

¹⁰ See Kaplan, *supra* note 6, at 1071-72.

¹¹ *Id.* at 1072-73.

find for plaintiff only when pA is greater than .50. Notice that " $p = .50 +$ " is just the numerical expression of "more likely than not," which is, in turn, a widely-accepted interpretation of the common law standard of proof for ordinary civil trials—the preponderance of the evidence.¹² On the other hand, if, in the ordinary criminal trial, an erroneous conviction is deemed much worse than an erroneous acquittal, Dp is much greater than Dd . With appropriate numerical values for these variables, the Kaplan formula indicates a standard of proof for criminal trials that is very much higher than .50. The common law standard of proof for criminal trials—beyond a reasonable doubt—actually is thought to be very much higher than the civil standard of proof.¹³ Thus far, then, application of the Kaplan formula has some appealing results.

B. WEIGHING RELATIVE INTERESTS

These results, it seems, have caused the Court to accept the Kaplan formula as a principal ground of rulemaking for standards of proof.¹⁴ However, rulemaking by the formula is problematic. The formula does not rationalize any particular standard of proof unless

¹² *Id.* at 1072. See also *Burnett v. Reyes*, 118 Cal. App. 2d 878, 880, 256 P.2d 91, 93 (1953); MODEL CODE OF EVIDENCE Rule 1(5) comment para. (5) (1942); C. McCORMICK, EVIDENCE § 339 (3d ed. 1984); E. MORGAN, SOME PROBLEMS OF PROOF 84-85 (1956); Gausewitz, *Presumptions*, 40 MINN. L. REV. 391, 400 (1956); McNaughton, *Burden of Production of Evidence: A Function of a Burden of Persuasion*, 68 HARV. L. REV. 1382, 1385-87 (1955).

¹³ Kaplan, *supra* note 6, at 1073.

¹⁴ One need not believe that the Court actually calculates with the Kaplan formula. As an arithmetical expression of the Court's due process doctrine, the Kaplan formula is the logic of its lawmaking. To discuss use of the Kaplan formula, therefore, is to discuss the Court's reasoning about standards of proof.

It is not clear that Kaplan himself would approve this use of his formula. He seems to be of two minds about it. On the one hand, he repeatedly suggests that his model has application to the particular decision of a factfinder, rather than the decisions of a lawmaker. See, for example, his remarks on "the typical decision-theory problem," *id.* at 1065, "expected utility of a decision," *id.* at 1068, the "criminal trial," *id.* at 1070, what "the jury must feel" in order to convict, *id.* at 1071, and "designing . . . a proper matrix," *id.* at 1080. Kaplan confirms these suggestions when he states that

[i]n a criminal trial, as in any decision process, we must consider the utilities associated with differing decisions of the particular case at issue—not just the average utilities over many disparate types of criminal cases. Thus the rational factfinder should consider the disadvantages of convicting *this* defendant of *this* crime if he is innocent as compared with those of acquitting him if he is guilty.

Id. at 1073-74 (emphasis in original).

On the other hand, Kaplan also suggests that his model has application to decisions that are ordinarily associated with lawmaking. Thus, immediately after saying that the rational factfinder should consider the relative disutilities of erroneous verdicts for "*this* defendant" charged with "*this* crime," he illustrates with a distinction between two *types* of crimes, embezzlement and child molestation, rather than two particular cases. *Id.* at 1074. He may intend to caution against the blanket use of one standard of proof in

values are assigned to the variables D_p and D_d . These values alone determine the standard of proof to be adopted.¹⁵ The Court may well see ordinary civil trials, in which money or property is at stake, as cases in which one party's gain and other's loss are commensurate. In rulemaking for such cases, any values may be assigned to D_p and D_d , provided that they are equal, and the Kaplan formula indicates the traditional standard of proof.

On the other hand, the Court has adopted standards of proof that are supposed to be higher than .50+ or the preponderance of the evidence. It has required proof beyond a reasonable doubt in criminal trials and proof by clear and convincing evidence in some special non-criminal trials.¹⁶ But neither of these standards has been justified by the Court's declaration that disutilities of errors in such trials have any particular values. To declare only that D_p exceeds D_d is only to justify some standard of proof higher than .50+. It is not to justify any one of the possibilities.

Of course, the Kaplan formula works both ways, so the values of D_p and D_d can be ascertained if the standard of proof is given. A survey of federal judges found that they tend to associate "clear and convincing evidence" with a probability of .75 and "beyond a reasonable doubt" with a probability of .90.¹⁷ By these equivalences,

criminal trials. However, distinction among types of offenses is appropriate for legal rules, and Kaplan's point seems a more fitting consideration for a lawmaker.

Furthermore, Kaplan argues that inequalities of the disutilities of errors explain some legal requirements of higher standards of proof—proof by clear and convincing evidence in trials of fraud and proof beyond a reasonable doubt in criminal cases. *Id.* at 1072-73. It is natural to understand Kaplan as implying that lawmakers do or ought to apply his formula directly to the problem of "Determining the Probability Necessary to Return a Verdict" (title of the section in which these arguments appear). *Id.* at 1071.

Finally, though Kaplan justifies the vagueness of "beyond a reasonable doubt" by its permitting decisions that serve concrete policies for individual cases, he recognizes that "beyond a reasonable doubt" is supposed to describe a higher standard of proof than either "the preponderance of the evidence" or "clear and convincing evidence" and, thus, that it imposes some legal limit on the minimum value of pA for a guilty verdict in any case. *Id.* at 1037-77. This is to recognize that the standard of proof serves as a rule.

It is not surprising, therefore, that the Court has understood Kaplan as offering a technique for fashioning a rule of procedure. Other readers have understood him in the same way. See, e.g., Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1378 (1971).

¹⁵ Decision theorists may take an alternative approach, assigning values to correct findings as well as erroneous ones. See, e.g., R. JEFFREY, *supra* note 5, at 75-76; Cullison, *supra* note 8, at 564-66. This alternative approach may serve a special purpose. See Tribe, *supra* note 14, at 1379 n.161. However, the Court has only hinted at this approach. See, e.g., *Santosky*, 455 U.S. at 769; *Winship*, 397 U.S. at 367-68.

¹⁶ See *supra* note 3.

¹⁷ McCauliff, *Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees?*, 35 VAND. L. REV. 1293, 1325, 1328, 1332 (1982). See also McNaughton, *supra*

the requirement of proof beyond a reasonable doubt implies that Dp is deemed about ten times worse than Dd, and the requirement of proof by clear and convincing evidence implies that Dp is about three times worse than Dd. One could assume that the Court, in adopting higher standards of proof, has assigned corresponding values to Dp and Dd, but that would be speculation. There is no evidence that the Supreme Court Justices agree with the average federal judge in their numerical translations of the common law standards.

Even if one speculates in this way, the Court's lawmaking is problematic. To use the Kaplan formula in appropriate circumstances is logical. But logic is vacuous; it is the mere form of reasoning. Although substance is supplied by evaluating Dp and Dd, reasons for the evaluations are still needed. There ought to be reasons for concluding that Dp is ten times greater than Dd in criminal trials, and that Dp is three times greater than Dd in some special non-criminal trials. Simply inferring the Court's values from the required standards of proof is not only speculative but unsatisfying.

Of course, one may speculate further upon the reasons. Perhaps the Court has evaluated the relative disutilities of errors in criminal trials in light of Blackstone's principle: better that ten guilty go free than that one innocent be convicted.¹⁸ This ground of evaluation is plausible. The principle is surely part of common political morality. Nevertheless, how should one speculate about the Court's requirement that factual grounds for involuntary civil commitment be proved by clear and convincing evidence? The implicit evaluation of errors about defendants' mental illness, which makes Dp three times greater than Dd, is unsupported by any obvious rationale. With the consequences in view, an erroneous finding of mental illness may well be worse than an erroneous finding of no such illness, but is it three times worse? Why not worse by a factor of five or even eight? Is an erroneous finding of neglect in a proceeding to terminate parental rights also just three times worse than an erroneous finding of no neglect? Finally, how do the traditional notions of fairness associated with due process indicate the correct choice of a factor?

note 12, at 1389; Simon, *Judges' Translations of Burdens of Proof into Statements of Probability*, 1969 TRIAL LAW GUIDE 103.

¹⁸ 4 W. BLACKSTONE, COMMENTARIES *358. See *Ballew v. Georgia*, 435 U.S. 223, 234 (1978).

C. THE APPORTIONMENT OF ERRORS

The Court's acceptance of Bayesian decision theory as a foundation of due process doctrine for burdens of proof is problematic in other respects. The Kaplan formula, with disutility values for its variables, is supposed to indicate a standard of proof that will apportion erroneous findings about fact A in some particular way. This supposition is reasonable only if it is reasonable to assume that triers' evaluations of evidence about A, to which the indicated standard is applied, have a certain kind of distribution along a continuum. There is no reason to assume that triers' evaluations of evidence will be distributed in such a way, and there is very good reason to assume that they will not.

To understand the problem, imagine a program of lawmaking for trials in which factfinders estimate the probabilities of material facts at issue and apply numerical standards of proof that the Court has determined by means of the Kaplan formula. On the one hand, the Court has a policy in favor of correct decisions by factfinders.¹⁹ On the other hand, because there seems to be no feasible trial procedure that would ensure correct decisions in every case, the Court also has policies for factfinders' erroneous decisions. These policies correspond to the Court's estimates of the relative disutilities of errors in different kinds of trials, as inferred in Section B, above. Errors in criminal trials are to be apportioned so that ten guilty defendants are acquitted for each innocent defendant convicted. Errors in extraordinary civil trials also are to be apportioned in favor of defendants, but at a lesser ratio of three to one. Errors in ordinary civil trials are to be apportioned equally between plaintiffs and defendants.²⁰ In furtherance of these policies, the Court has established the standards of proof indicated by the Kaplan formula: for criminal trials, .90+; for extraordinary civil trials, .75+; and for ordinary civil trials, .50+.²¹

Has the Court done what it believes it has done? Will the standards of proof apportion factfinders' errors according to the Court's

¹⁹ "[W]e must be mindful [in assessing the interests of the individual and the state] that the function of legal process is to minimize the risk of erroneous decisions." *Addington*, 441 U.S. at 425. See *Matthews v. Eldridge*, 424 U.S. 319, 335 (1976); *Speiser v. Randall*, 357 U.S. 513, 525-26 (1958).

²⁰ *Addington*, 441 U.S. at 423; *Rosenbloom v. Metromedia*, 403 U.S. 29, 50 (1971); *Winship*, 397 U.S. at 371 (Harlan, J., concurring).

²¹ In the terminology of economics, the Court's policy for errors is its rate of indifferent substitution for errors in favor of *P* and errors in favor of *D*. On the assumption that a standard of proof will determine a rate of substitution for errors of the two kinds, the standard of proof that will effectuate the Court's policy is the standard that determines a rate of substitution equal to its rate of indifferent substitution.

policies? Decision theory proves that factfinders who meet certain qualifications will, in certain circumstances, arrive at final values of pA that must be supposed to have a certain kind of distribution.²² Decision theory also proves that, if the Court's standards of proof are applied to the values of pA in this distribution, the erroneous findings about A are apportioned according to the Court's policies.

Allowing for some tolerable roughness, the requisite qualifications and circumstances of the factfinders have quite simple descriptions. The circumstances are the factfinders' possession of full and accurate information about the probability of A before any evidence is received and about the probabilities that attach to each item of evidence that is received at trial.²³ The qualifications may be called perfect, decision-theoretical rationality. For the decision theorist, a perfectly-rational factfinder acts with Bayesian rationality. Such a factfinder begins with the initial probability of A , which, upon receipt of each item of evidence, he modifies by application of Bayes' Theorem until all evidence is received, and the factfinder reasons in this way only.²⁴ It must be supposed that perfectly-rational and fully-informed factfinders reach final values of pA in the wanted distribution.

D. CONTRAINDICATIONS

However, decision theory provides no endorsement of the Court's lawmaking unless factfinders are indeed fully informed and perfectly rational. There is every reason to assume that they are neither. Indeed, the actual process of cumulating evidence at trials is far from the decision theorist's ideal. One may accept some very general propositions of the Bayesian model,²⁵ but find that it fails as "a quantitative description of the ordinary process of weighing evidence."²⁶

²² The "must" here expresses the compulsion of reason.

²³ Failure to produce a possible item of evidence is treated as evidence about A in this connection.

²⁴ A clear account of Bayesian reasoning is found in Kaplan, *supra* note 6, at 1083-91, and in Tribe, *supra* note 14, at 1350-58.

²⁵ See, e.g., Ball, *supra* note 9, at 829.

²⁶ I. GOOD, PROBABILITY AND THE WEIGHING OF EVIDENCE § 6.1 (1950). Thus, suppose the testimony of nineteen witnesses supports A , but the testimony of twenty witnesses supports not- A . Given special assumptions—that the prior probability of A is exactly .50, that the probability of truthfulness is greater than .50 and equal for all witnesses, and that witnesses' credibilities are independent—it may be unreasonable to think that A is at least as probable as not- A . See Ball, *supra* note 8, at 824-25, 852-58; Trickett, *Preponderance of Evidence, and Reasonable Doubt*, 10 FORUM 75, 77-78 (1906). However, there is no good reason to suppose that such special assumptions would be made in all or, for that matter, any actual trials. For a thorough critique of the mathema-

Whether, and to what extent, factfinders in trials approximate the Bayesian ideal are unanswered empirical questions. But, empirical studies have indicated that such approximation is, at best, slight.²⁷ Furthermore, the empirical studies agree with considered opinions on the subject. It is significant, and only fair to mention, that Kaplan is of the same mind.²⁸ These opinions are based on more than the familiar disagreement about the probability of some proposition among people who have received and carefully reckoned with the same evidence.²⁹ Factfinders are ignorant of crucial statistical information relevant to assigning initial probabilities to material facts at trial, and their guesses about these initial probabilities can be aberrant.³⁰

Moreover, factfinders who have such statistical information may be unable to make proper adjustments in the initial probabilities upon receipt of more impressionistic evidence during trial, even if they apply Bayes' Theorem correctly. Consider, for example, the failure of a party to produce a possible item of evidence. To make a proper adjustment of the initial probability of the fact in light of such failure, the trier would have to reckon with such probabilities as are associated with the accessibility of the evidence and its relative cost to the party. The probabilities associated with cost, in turn, depend on the resources of the party and the other demands on those resources as determined by the party's preferences, litigation

tician's calculus of chance as applied to judicial factfinding, see L. COHEN, *THE PROBABLE AND THE PROVABLE* §§ 14-39 (1977).

²⁷ See, e.g., Edwards, *Conservatism in Human Information Processing*, in *FORMAL REPRESENTATION OF HUMAN JUDGMENT* 17 (B. Kleinmuntz ed. 1968); Slovic & Lichtenstein, *Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment*, 6 *ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE* 649 (1971). See generally Slovic, Fischhoff & Lichtenstein, *Behavioral Decision Theory*, 28 *ANN. REV. PSYCH.* 1 (1977).

²⁸ Kaplan, *supra* note 6, at 1084-91.

²⁹ Tribe, *supra* note 14, at 1348. For a criticism of the Bayesian model that is more technical and, in some respects, more thorough than the one offered here, see *id.* at 1358-68.

³⁰ Weld & Roff, *A Study in the Formation of Opinion Based upon Legal Evidence*, 51 *AM. J. PSYCH.* 609, 617 (1938). This is true even of judges, who should be better informed than laymen about many such facts. For example, it has been found that only 3.5% of the cars in accidents have mechanical defects and that mechanical defects of the cars play a part in only .25% of all accidents. James & Dickinson, *Accident Proneness and Accident Law*, 63 *HARV. L. REV.* 769, 770-71 (1950). But the New York Court of Appeals has said that "the probability that [an unexplained accident involving an automobile] occurred from a break in its mechanism is at least equally great" as the probability that the accident resulted from a lack of care in the automobile's operation. *Galbraith v. Busch*, 267 *N.Y.* 230, 234, 196 *N.E.* 36, 38 (1935), discussed in Jaffe, *Res Ipsa Loquitur Vindicated*, 120 *BUFFALO L. REV.* 1, 3-4 (1951). Perhaps the concept of probability is inapplicable to negligence, see *supra* note 7, but this conceptual difficulty would hardly excuse such a preposterous assertion.

strategy, estimate of the relative effect of the evidence on the trier, and expected value of a favorable verdict. All of these factors color a failure to produce evidence even if it exists, but a factfinder at trial could only be expected to apply very general commonsensical information in giving it significance. Furthermore, no one could have much confidence that this information will have a proper corrective influence on a statistically-based initial probability. Finally, non-rational forces may affect triers, deliberations and impair their abilities to give proper weight to evidence, even if the triers are fully informed about the requisite probabilities.³¹

Doubts about factfinders' information and rationality are manifested by the various rules intended to improve their weighing of evidence at trials. The requirements of two witnesses to treason and corroboration of an accomplice, the binding effect of judicial notice, *falsus in uno*,³² and rules as to attesting witnesses are straightforward quantitative or qualitative evaluations.³³ Some rules, most notably the hearsay exclusion, keep evidence from factfinders because they are deemed likely to miscalculate it.³⁴ For much the same reason,

³¹ Some such possible influences are obvious; others are not. For example, research has found that people tend to prefer a .50-.50 wager to a .25-.75 wager, although the average return is the same in both. Edwards, *Measurement of Utility and Subjective Probability*, in *PSYCHOLOGICAL SCALING: THEORY AND APPLICATIONS* 109 (H. Bullicksen & S. Messick eds. 1960); Edwards, *Probability-Preference Among Bets with Different Expected Values*, 67 AM. J. PSYCH. 56, 57 (1954); Edwards, *Probability Preference in Gambling*, 66 AM. J. PSYCH. 349, 357-58 (1953); Edwards, *The Reliability of Probability Preferences*, 67 AM. J. PSYCH. 68, 74-75 (1954); Edwards, *Subjective Probabilities Inferred from Decisions*, 69 PSYCH. REV. 109 (1962); Edwards, *Variance Preferences in Gambling*, 67 AM. J. PSYCH. 441, 450-51 (1954). If the .50-.50 probability has some general psychological attraction, triers will tend to evaluate *p*A at .50 even though Bayesian rigor would call for a somewhat higher or lower evaluation. For another example of non-rational influences, see the discussion of cognitive dissonance in Tribe, *supra* note 14, at 1383-84.

³² *Falsus in uno, falsus in omnibus* is the maxim for a rule applicable where a witness is found to have testified falsely about one material fact. Upon request of the opponent, the jury must be charged that it may disregard all of the witness' testimony.

³³ Cleary, *Evidence As A Problem In Communicating*, 5 VAND. L. REV. 277, 280 (1952). Of course, some such rules may be ill-suited to their purposes. For a criticism of one of these rules, see Note, *Falsus in Uno, Falsus in Omnibus*, 29 NEB. L. REV. 122 (1949).

The evaluation of evidence is extremely complicated, and common judicial sense may go awry in the formulation of rules for this business. For example, it is often said that an inference may not be founded on an inference and that direct evidence is stronger than circumstantial evidence. But it is usually easier to lie directly than to fabricate a convincing chain of circumstances. If the lawmaker were troubled only by the possibility that someone had framed a criminal defendant, he should be more wary of perjury by an "eyewitness" than an elaborate web of inculpatory circumstances. Abhorrence of "mere conjecture" or "speculation" is exaggerated relative to the precautions taken with testimony; after all, cross-examination is not an acid test of truthfulness. See generally G. TULLOCK, *supra* note 9, at 97-104.

³⁴ C. McCormick, *supra* note 12, § 245, at 728-29; E. MORGAN, *BASIC PROBLEMS OF EVIDENCE* 248 (1963).

other rules provide for exclusive types of proof of some facts.³⁵ Perhaps the largest and most notorious family of these rules for weighing evidence are presumptions founded entirely on the probative value of the basic facts, which lawmakers have established "to make more likely a finding in accord with the balance of probability."³⁶ If lawmakers had confidence in Bayesian factfinding, they would not adopt such rules.

Finally, the belief that factfinders are far from the Bayesian ideal is supported not only by empirical studies and common sense, but also by the many rules that limit the factfinders' information at trials. Notwithstanding the probative value of hearsay evidence,³⁷ factfinders may not consider it. Exclusion of relevant evidence that is deemed too inflammatory or distracting³⁸ may make a net improvement in the triers' deliberations.³⁹ But it does keep factfinders in ignorance, and it is justified by doubts about their rationality. Exclusion of evidence for reasons of policy extraneous to the factfinding process, which is the work of rules of privilege,⁴⁰ has the same effect. Even where rules do not exclude relevant evidence, because it cannot be excluded, a rule may require factfinders to ignore it.⁴¹ To the extent that they abide by the rule, the factfinders are prevented from considering significant information.

The decision theorist's assurance, that fully-informed and perfectly-rational factfinders will arrive at such values of pA as guaran-

³⁵ Statutes of frauds, for example, which require writings to support agreements concerning certain kinds of subjects.

³⁶ Morgan, *Instructing the Jury Upon Presumptions and Burden of Proof*, 47 HARV. L. REV. 59, 77 (1933). The National Commission on Reform of Federal Criminal Laws, proposing change for the criminal code's presumptions, said that it is necessary to advise jurors that the basic facts are "strong evidence of the fact presumed" because "the value of the basic facts . . . is not readily apparent to them." 1 WORKING PAPERS OF THE NAT'L COMM'N ON REFORM OF FEDERAL CRIMINAL LAWS 21 (1970). It may be said that rules requiring only such a charge do not concern what are properly called presumptions. Cleary, *Presuming and Pleading: An Essay on Juristic Immaturity*, 12 STAN. L. REV. 5, 16 (1959). Indeed, some authorities insist that no true presumption is founded entirely on the probative value of the basic facts. The present point is unaffected by these objections.

³⁷ See *supra* note 34 and accompanying text.

³⁸ See generally C. McCORMICK, *supra* note 12, § 185 at 544-46.

³⁹ Kaplan, *supra* note 6, at 1077.

⁴⁰ C. McCormick, *supra* note 12, § 72. The rules of privilege are not the only members of this class. Another is the rule of incompetency to testify as to sexual non-access for the purpose of bastardizing an infant, which has, in some jurisdictions, been modified only by statutes in derogation of the common law and strictly construed. See, e.g., *Sayles v. Sayles*, 323 Mass. 66, 80 N.E.2d 21 (1948).

⁴¹ For example, the privilege against self-incrimination may require that a jury be admonished against considering the criminal defendant's silence in determining the probability of his guilt. Clearly, this requirement is not based on the proposition that his silence is without probative value.

tee success with the Court's standards of proof, has little relevance to sound lawmaking practice. It would be unreasonable to assume that factfinders are fully informed and perfectly rational.⁴² This may be a fairly obvious conclusion. However, the preceding argument is not superfluous.⁴³ The factors that make trials differ from the deci-

⁴² There are also difficulties with the alternative employment of Kaplan's formula as a model for the particular decisions of factfinders. See *supra* note 14 and accompanying text. If factfinders were to make decisions in particular cases on grounds identified by the Kaplan formula, one formulation of the standard of proof for all trials, criminal and civil, would have been adopted long ago. Indeed, the phrase "beyond a reasonable doubt" would serve very well. Kaplan himself observed that the phrase has a comfortable elasticity that permits factfinders to tailor the standard of proof to the varying disutilities of errors in particular criminal cases. Kaplan, *supra* note 6, at 1073. His point has broader application.

Decision-making bears an obvious relation to action, and rational decision-making bears a like relation to reasonable action. Action in conditions of uncertainty about relevant facts is action affected by doubts about the consequences, and the weights of such doubts depend on both the estimated likelihoods of the consequences and the values attached to these consequences. See Ball, *supra* note 9, at 815-16. A reasonable doubt is a doubt of such weight as to make the proposed action unreasonable. The Kaplan formula, therefore, is a description of the logic of reasonable doubt or, more specifically, reasonable doubt where the proposed action is a verdict with some legal consequences. A finding of fact that is beyond a reasonable doubt is a finding that is not affected by a doubt of sufficient weight to make the finding unreasonable.

The Kaplan formula is supposed to have general application to all trials, civil and criminal. Therefore, if factfinders were to attach values to consequences *D_p* and *D_d* in particular cases and follow the logic of the Kaplan formula, the beyond a reasonable doubt formulation would be used in all trials because it expresses the Kaplan formula. But, the "beyond a reasonable doubt" instruction will not be used in all trials, and so factfinders are not to make trial-by-trial applications of the Kaplan formula. Indeed, various means are used to block or restrict such applications. See Kaplan, *supra* note 6 at 1074-77. The reason for these efforts to prevent factfinders' application of the Kaplan formula is not that they are deemed incapable of doing the arithmetic. The Hand formula is used in arguments about reasonable care, and this formula, like Kaplan's, is a logic of action in view of probabilities and consequences. Rather, part of the reason for these efforts may be a suspicion that factfinders in particular cases will boggle at what seems to be a gamble with the property, and especially the liberty, of a party. See *id.* at 1068; Trickett, *supra* note 26, at 83-84. Supplying factfinders with a rule for the standard of proof relieves them of any aversion to gambling, and introduces the broader perspective of a lawmaker who is ready and able to take responsibility for the decisions that guide the conduct of public institutions in important affairs. However, the main reason for using rules for standards of proof is a mistrust of policy-making by different factfinders in particular cases. C. McCORMICK, *supra* note 12, § 336, at 948. With regard to trials, mistrust is acute because factfinders' decisions are so difficult to review. Kaplan, *supra* note 6, at 1077. Furthermore, the relevant policy considerations are very complicated. As Kaplan has observed, "the goals of our criminal system as a whole may not only be somewhat unclear, but may also on occasion be at odds with one another." *Id.* at 1076. The same may be said of our civil system. However, in any such circumstances, the proper policy-making institution is not the *ad hoc* factfinder, but rather the lawmaker. Compare *id.* at 1076-77 with Tribe, *supra* note 14, at 1381-85.

⁴³ For one thing, the argument has shown something about the fairly obvious conclusion. It has shown that the conclusion does not merely provoke again the question what we may assume about factfinders' behavior, but also that this renewed question is an

sion theorist's ideal are factors that a reasonable lawmaker would consider in establishing rules for standards of proof. They are part of the empirical side of this rulemaking. Therefore, if the decision theorist's ideal does not constitute an empirically-acceptable assumption about the behavior of factfinders at trials, a fresh approach is needed, though the question remains the same: will the Court's standards of proof serve its policies for errors in the various kinds of trials?

E. MODELING EVALUATIONS OF EVIDENCE

The question is more practical than theoretical, and better for a social scientist than a decision theorist. Nothing in the extant literature answers it, but an answer could be found. By adapting a technique described by Savage for determining the probability that someone ascribes to a certain proposition's being true,⁴⁴ the social scientist could devise a way to instruct factfinders in applying a numerical standard of proof for issue A and a way to measure the values of pA that they reach.⁴⁵ Basing both the instruction for the standard of proof and the survey questions on the same technique makes them compatible measures of the same thing. Therefore, the standard of proof could be compared directly to the survey data. If the social scientist surveyed factfinders in a representative sample of trials of issue A, the data on their evaluations of pA would be projectable and the relative number of trials in which factfinders will arrive at the several possible values of pA could be predicted with confidence. Then, it could be predicted how any standard of proof will apportion findings of A and not-A.

empirical question. Probability theory has two camps—the *a priori* theorists and the empiricists—who, respectively, ignore and celebrate the shaved die. Those who would offer a theory of standards of proof constituted by the Kaplan formula alone—who would suppose that a .90+ standard of proof results in just one conviction of an innocent for every ten guilty verdicts—are intellectual kin to the *a priori* theorists. The model offered in the following section of this Article, which takes seriously the proposition that actual factfinding is not Bayesian, is in the empiricists' spirit. Non-Bayesian factfinders are the shaved dice of judicial trials.

⁴⁴ L. SAVAGE, FOUNDATIONS OF STATISTICS (1954). For two relatively informal descriptions of Savage's "personalistic theory of probability," see Kaplan, *supra* note 6, at 1066-67, and Tribe, *supra* note 14, at 1346-47.

⁴⁵ Both the instruction and the measurement require the factfinder to suppose an offer of two wagers: (1) that A is true; and (2) that, in a random drawing, a white ball will be drawn from a container of X white and Y black balls, giving suitable values to the variables X and Y. For example, the instruction for a .50+ standard of proof would give 50-50 values and would direct a finding for plaintiff just in case wager (1) is preferred to wager (2). The measurement of pA would give a series of values to the variables and would ask, for each set of values, which wager the factfinder preferred.

1. Ordinary Civil Trials

Let A be the decisive issue of fact in some ordinary civil trials. A finding that defendant did A gives judgment for plaintiff, and a finding that defendant did not- A gives judgment for defendant. Regarding such trials, it is natural to assume that the litigants' resources and abilities are approximately equal; that issue A is genuinely controversial; and that plaintiffs are right about A in half the cases, whereas defendants are right in the other half. At the very least, these assumptions are supported by our ideology, whereby Justice is blind to differences of persons and all disputes are fit to be tried without prejudice. It is also natural to assume that the factfinders in such trials are neither irrational in their evaluations of evidence nor equipped with extraordinary reasoning abilities or knowledge. On the basis of these assumptions, it is natural to assume further that, if the previously-described poll of factfinders were made, their evaluations of pA would show the normal distribution of a continuous variable. Plotting this data on a graph might yield the curve of Figure 1.

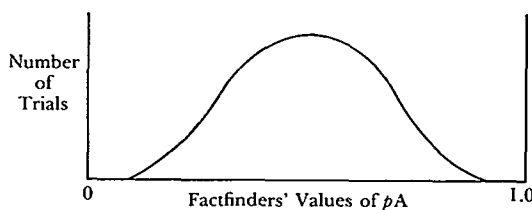


FIGURE 1

Call a curve like this a p -curve. As is apparent, a standard of proof for A will divide the p -curve and apportion findings of A and not- A in a particular way. Triers who apply a .50+ standard of proof, for example, would find for plaintiff in about half of the cases and for defendant in the rest. Different standards of proof apportion the findings differently. But the p -curve of Figure 1 does not show how many findings will be correct or erroneous, nor, in particular, how the errors will be apportioned.

The missing information would be available if, for each trial, it was known whether defendant actually did A or did not- A . But that would make nonsense of the whole enterprise, which was launched by the contrary assumption that evidence will never cause certainty about an issue of fact. Nevertheless, the social scientist could run a series of mock trials, realistic in the minutest details, except that he knows whether or not the defendant actually did A , and the defend-

ant actually did A in just half the cases. The social scientist could poll the factfinders and sort the p -curve data into two groups according to the actual merits of the trials. Such data would be eminently projectable and would improve the graph of Figure 1.

There is one assumption about the sorted p -curve data that should be added to the others. It is reasonable to assume that the mean of the values of pA for trials in which defendant actually did A will be higher than the mean for trials in which defendant actually did not-A. Indeed, this assumption simply expresses our belief that factfinders generally appreciate, to some extent, the probative value of evidence and that trying cases can resolve factual disputes more accurately, in the long run, than flipping a coin.

Given this assumption and certain others,⁴⁶ it might be supposed that a graphic display of the sorted p -curve data would look like Figure 2. Figure 2 shows two p -curves; the curve marked I dis-

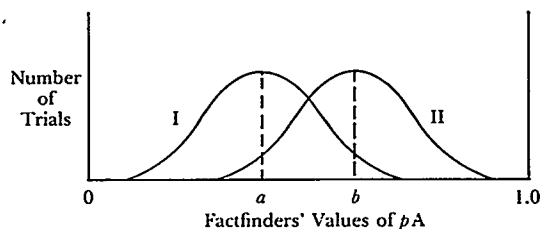


FIGURE 2

plays the data for trials in which defendant actually did not-A, and the curve marked II displays the data for trials in which defendant actually did A. Points a and b mark the mean values of pA for the two groups of data. In accordance with our assumption about factfinders' use of evidence, b is greater than a .

⁴⁶ For example, one important assumption is that the ranges of values of pA in the two groups of data will overlap. This is not a dubious assumption. It is implied by Trickett's observation that "[t]here can be evidence that fact X occurred, when it did not occur, and evidence that fact X did not occur, when it did occur, and, for the same reason, there can be more evidence that it occurred than that it did not occur, although it in fact did not occur. . . ." Trickett, *supra* note 26, at 78. The shape of the curves' slopes near the overlap is also plausible. The author of a formal model of deterrence, after defining y as the likelihood of punishment of a person who has obeyed the law and z as the increase in this likelihood resulting from that person's breaking the law, asserts: "Over the relevant range not only y but also the rate of change of y should increase as z increases, since high levels of z normally imply punishment of the less clearly culpable." Birmingham, *A Model of Criminal Process: Game Theory and Law*, 56 CORNELL L. REV. 57, 64 (1970). The truth of this assertion is far from obvious; but, if p -curves overlap in the way assumed here, Birmingham's assertion is true where y and z pertain to triers' findings that a person has broken the law.

2. The Standard of Proof and Distribution of Findings

If the factfinders in these trials apply the .50+ standard of proof (SOP) that the Court has adopted,⁴⁷ verdicts for plaintiffs (*Ps*) and defendants (*Ds*) will be distributed in the way illustrated by Figure 3. Trials represented by the area with left-facing diagonal lines will result in verdicts for plaintiffs, and trials represented by the area of horizontal lines will result in verdicts for defendants.⁴⁸

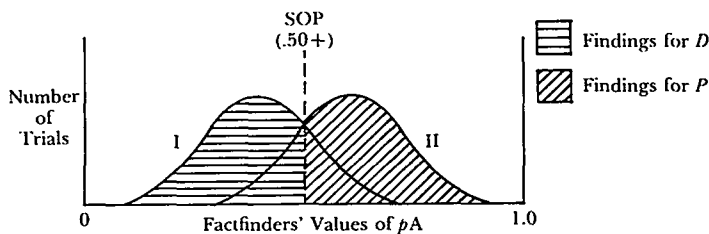


FIGURE 3

Now consider whether the Court's policy for factfinding in ordinary civil trials, which calls for a maximum of accuracy and apportionment of errors equally between plaintiffs and defendants, is served by the standard of proof indicated by the Kaplan formula. In Figure 4, right-facing diagonal lines and vertical lines have been ad-

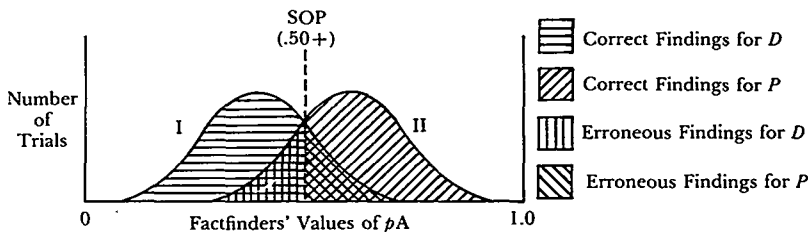


FIGURE 4

ded to represent the numbers of erroneous findings for plaintiffs and defendants, respectively. Inspection of Figure 4 reveals that errors are apportioned equally and that any other standard of proof would cause more errors. Therefore, it appears that the Court's policy for factfinding in ordinary civil trials is served by the standard of proof given by the Kaplan formula, provided that the *p*-curves for such trials have certain favorable properties. Moreover, reasonable

⁴⁷ See *supra* note 21 and accompanying text.

⁴⁸ Factfinding by the corporate decision of a jury calls for a minor reconceptualization of *p*-curve data. It might be imagined, for example, that the measured value of *pA* for each trial is the lowest reached by a member of whatever majority is required.

general assumptions about such trials favor the supposed configuration of p -curves, so the proviso may be deemed satisfied.

3. Criminal Trials

The Court has a different policy for errors in criminal trials. These errors are to be apportioned in a 10:1 ratio, favorable to defendants.⁴⁹ To enforce this policy, the Court has adopted a .90+ standard of proof, as indicated by the Kaplan formula.⁵⁰ Again, the question is whether the Court has done what it has aimed to do. The crucial factor is the configuration of p -curves for such trials. If these curves have the same properties that were supposed for ordinary civil trials, it is clear that the Court has blundered in setting the criminal standard of proof at .90+. As is apparent in Figure 5, a

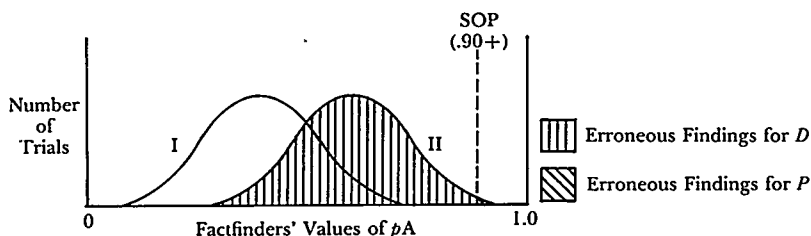


FIGURE 5

.90+ standard of proof will not yield the wanted distribution of errors. The .90+ standard will cause acquittals in all the trials in which defendants are actually innocent.⁵¹ But this standard will also cause acquittals in nearly all the trials in which defendants are actu-

⁴⁹ See *supra* note 19 and accompanying text.

⁵⁰ See *supra* note 21 and accompanying text.

⁵¹ Kaplan argues that:

[i]t would be comforting to be able to say that in all cases of doubt we would acquit rather than chance convicting an innocent man. But because no case is doubt free, unless we decide to avoid trying anyone, we will, if we try enough people, inevitably convict an innocent man. Moreover, we cannot avoid this by trying only a small number of defendants. Since the chance of error, however small, is random, we may convict an innocent man, not only in the long run but in our very first trial.

Kaplan, *supra* note 6, at 1071. If "we" refers to people other than triers of fact, the confusion of subjective "doubt" with objective "chance of error," which underlies this argument and similar arguments by many other writers, should be obvious. It may be that "we will, if we try enough people, inevitably convict an innocent man;" but, if this is so, it is because of the ways triers will evaluate evidence and apply standards of proof. Furthermore, it may be that we cannot be sure that only the guilty are convicted; but, if this is so, it is because we cannot be sure how triers will evaluate evidence and apply standards of proof. The fact that "no case is doubt free" is hardly relevant.

Compare also the assertion of Birmingham, described *supra* note 46, that with constant resources available for law enforcement, an increase in z will normally cause increase in y . Figure 5 represents one kind of exceptional situation; it shows that z may be increased, by reducing the standard of proof, without a resulting increase in y .

ally guilty. Furthermore, many of these erroneous findings for defendants are utterly uncompensated for by a reduction in the number of erroneous findings against defendants. Thus, the .90+ standard of proof will not only fail to effectuate the Court's policy for criminal trials; this standard will subvert the Court's general policy in favor of correct findings of fact and most other aims of the criminal trial process.

The Court has not gone wrong in its general aims for criminal trials, its policy favoring correct verdicts in such trials, or its policy favoring ten erroneous acquittals for each erroneous conviction. Rather, the Court has gone wrong in relying on the Kaplan formula alone to determine generally-applicable standards of proof. That formula may be well-suited to some needs of a one-time decisionmaker who is uncertain about crucial facts and is without authoritative guidance. However, it is not a dependable tool for the lawmaker. A lawmaker who undertakes to fashion a rule that will serve his policy for factfinding errors in a large class of trials must reckon with the kinds of empirical considerations that underlie p -curves. The bare Kaplan formula leaves these considerations entirely out of account.

II. LAWMAKING FOR COMMON LAW STANDARDS

A. MODELING EVALUATIONS OF PROOF

Sound practice needs better theory. To that end, the preceding model should have application in lawmaking for traditional common law standards of proof. These standards have accepted translations into probability values: "proved by the preponderance of the evidence" means ".50+ probable"; "proved by clear and convincing evidence" means ".75+ probable"; and "proved beyond a reasonable doubt" means ".90+ probable".⁵² This suggests that the probability continuum may be divided into six ranges of probability values, and that a different common law standard of proof may be associated with each range. For trials of any issue F , the empirical analysis of factfinders' evaluations of pF would allocate these evaluations among the six ranges of values: (C') 0 to $<.10$; (B') $.10$ to $<.25$; (A') $.25$ to $<.50$; (A) $>.50$ to $.75$; (B) $>.75$ to $.90$; and (C) $>.90$ to 1.0 . A factfinder who reached a value of pF within range A would find F proved by the preponderance of the evidence, but not by clear and convincing evidence nor beyond a reasonable doubt; within range B , he would find F proved by the preponderance of the

⁵² See *supra* note 17 and accompanying text.

evidence or by clear and convincing evidence, but not beyond a reasonable doubt; within range C , he would find F proved by any common law standard. Similarly, within range A' , he would find not- F proved by the preponderance of evidence, but not otherwise; and so on, for ranges B' and C' . Of course, the factfinder whose value of pF lay within ranges A - B would not necessarily find F as a fact; nor would a value within ranges A' - B' necessarily call for finding not- F . A finding of F or a finding of not- F depends on the applicable standard of proof, as well as the trier's evaluation of pF .⁵³

The common law does not identify standards of proof with ranges of probability values, so this suggested empirical analysis is unacceptable. Nonetheless, this analysis serves as a conceptual link between the two models. Lawmaking for common law standards of proof is concerned with the distribution of factfinders' evaluations of evidence about F in all trials of the issue. But here there may be six evaluations, rather than the indefinite number of possible values of pF . Evaluations of the evidence about F (Proof/ F) fall on a graduated scale defined by the common law standards: preponderance of the evidence (POE); clear and convincing evidence (CCE); and beyond a reasonable doubt (BRD). Factfinders evaluate Proof/ F as: (C') not- F proved BRD; or (B') not- F proved CCE; or (A') not- F proved POE; or (A) F proved POE; or (B) F proved CCE; or (C) F proved BRD.⁵⁴

This is the basic analysis for factfinders' evaluations of Proof/ F . A social scientist could obtain projectable data about such evaluations for trials in which F actually exists and for trials in which not- F actually exists. Furthermore, it is possible to make reasonable assumptions about the data. Consider ordinary civil trials in which F is at issue. On the assumptions about such trials that were made in Section E.1., *supra*, a graphic display of the data would take the form of Figure 6 where six sets of p -bars show the numbers of trials in which factfinders arrive at the several evaluations of Proof/ F .

Data for the trials in which not- F actually exists are identified with the numeral I, and data for the trials in which F actually exists are identified with the numeral II. As is apparent, the p -bars for issue F are similar to the p -curves for A in all relevant respects. (Compare Figure 6 with Figure 2). Furthermore, the two forms of empirical analysis have similar roles in lawmaking for standards of

⁵³ The analysis does not provide for evaluation of pF at exactly .50. This omission is of no concern here, but cases in equipoise have significance for the model of lawmaking and will get attention presently.

⁵⁴ The special transitivity relations described in the preceding paragraph are respected also in these evaluations.

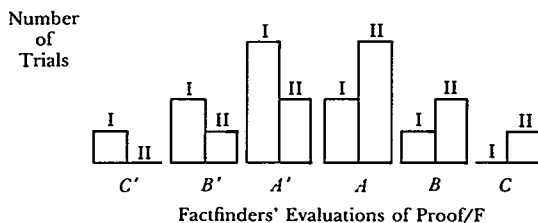


FIGURE 6

proof. Given the Court's policy for errors in ordinary civil trials and the supposed data on evaluations of Proof/F, Figure 7 shows that this policy will be enforced by requiring plaintiffs to prove F by the preponderance of the evidence.

B. THE BURDEN OF PERSUASION

Figure 7 also shows that the Court's policy for errors in ordinary civil trials will be enforced as well by requiring defendants to prove not-F by the preponderance of the evidence. The burden of persuasion for an issue is ordinarily treated as a simple function of the applicable standard of proof. This burden lies with the party who must satisfy the trier by at least the preponderance of the evidence in favor of the fact.⁵⁵ On this conception and the basic analy-

⁵⁵ See, e.g., MODEL CODE OF EVIDENCE Rule 1; C. MCCORMICK, *supra* note 12, § 339, at 956-59; E. MORGAN, *supra* note 34, at 20; Cleary, *supra* note 36, at 15-16; McBaine, *Burden of Proof: Degrees of Belief*, 32 CALIF. L. REV. 242, 254 (1944); Morgan, *supra* note 36, at 66-67. This usage is consistent with the frequent suggestion that the burden varies in degree, according to the applicable standard of proof or measure of persuasion. See, e.g., E. MORGAN, *supra* note 34, at 21-26; McBaine, *supra*, at 245-46, 254; Morgan, *supra* note 36, at 60. The suggestion, however, serves no great purpose and may cause confusion.

The burden of persuasion is sometimes called the risk of non-persuasion. E. MORGAN, *supra* note 34, at 19; F. JAMES & G. HAZARD, CIVIL PROCEDURE § 7.6 (3d ed. 1985); 920J. WIGMORE, EVIDENCE § 2485 (Chadbourn rev. ed. 1981 & Supp. 1986). By the same token, the risk of non-persuasion is said to lie with the party who must convince the trier that the fact at issue more likely than not exists. Ball, *supra* note 9, at 817; McNaughton, *supra* note 12, at 1383. Furthermore, the various standards of proof suggest that the risk of non-persuasion is a matter of degree. *Id.* Talk of a greater or lesser burden of persuasion can be harmless, but talk of a greater or lesser risk of non-persuasion is not. It may tempt one to suppose that the party having the greater risk of non-persuasion, in the sense of the higher standard of proof, *pro tanto* runs the greater risk of losing on the issue. This supposition is false, as the model of rulemaking for standards of proof makes plain. The risk-of-non-persuasion terminology has unambiguous sense where the ordinary civil standard of proof is applicable in the ordinary civil trial. In such a situation, the party who loses if the trier finds the two contradictory assertions equally likely to be true is the party who runs a risk that the other does not. Use of "risk of non-persuasion" should be limited accordingly.

There have been odd departures from the standard conception of the burden of persuasion. It has been suggested that the burden is borne by the party who loses if the evidence leaves the mind of the trier in equilibrium. See, e.g., McBaine, *supra*, at 243-44.

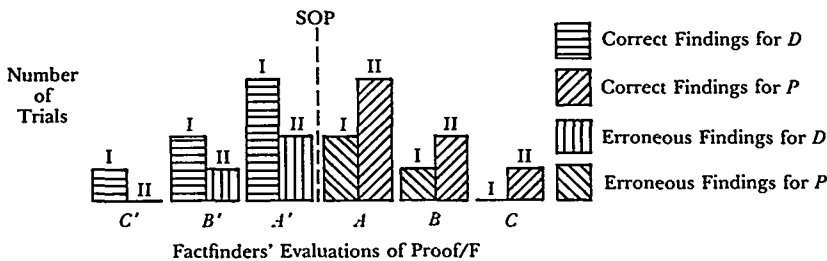


FIGURE 7

sis, the burden of persuasion might be allocated to either party, with no consequences whatsoever for the Court's policy.

By the basic analysis, only allocation of the burden of persuasion at a standard of proof higher than the preponderance of the evidence will affect the distribution of findings on an issue. The bare burden of persuasion is inconsequential. Moreover, the analysis has general application. Thus, the basic analysis indicates that the Court's lawmaking for bare burdens of persuasion in both civil trials⁵⁶ and criminal trials⁵⁷ is idle.

Of course, the basic analysis itself is founded on an assumption about the factfinders' evaluations of evidence about an issue of fact. It is assumed that factfinders in virtually all trials will conclude that either F or not-F has been proved by at least the preponderance of the evidence. The truth or falsity of this assumption is an empirical question. If this assumption is false, then the relative number of trials in which factfinders will conclude that neither F nor not-F has been proved by the preponderance of the evidence is also an empirical question. There is no good research on the subject. Nevertheless, the Court has acted as if it believes that the bare burden of persuasion has important consequences for factfinding,⁵⁸ and some

Where the ordinary civil standard of proof is applicable, this may identify the party with the burden. However, another standard of proof may require a rule that furnishes a solution if the trier is uncertain whether the standard is satisfied or not. This rule, in turn, might resolve the issue against either party, not just the party whose standard of proof is higher. Perhaps the most unusual description of the burden of persuasion yet proposed is a combination of the idea that the burden pertains to quandaries of the trier with the idea that the burden admits of degrees. See Underwood, *The Thumb on the Scales of Justice: Burdens of Persuasion in Criminal Cases*, 86 YALE L.J. 1299, 1300-01 (1977). This strange conception of the burden of persuasion serves only the special purposes of its author.

⁵⁶ Compare *Mobile J. & K.C.R.R. v. Turnipseed*, 219 U.S. 35 (1910) with *Western & Atl. R.R. v. Henderson*, 279 U.S. 639 (1929).

⁵⁷ Compare *Mullaney*, 421 U.S. 628 with *Patterson*, 432 U.S. 197.

⁵⁸ See *supra* note 56 and 57. See also *Speiser v. Randall*, 357 U.S. 513, 525 (1958). In *Turnipseed*, 219 U.S. 35, the Court upheld a legal presumption of defendant railroad's

lawyers agree.⁵⁹ An alternative modified p-bar analysis accommodates this belief.

The belief is founded on a special assumption about factfinders' evaluations of evidence: in a significant number of trials, the factfinder will stand paralyzed, unable to discern a relevant difference between one party's proof of F and the other's proof of not-F. More particularly, in these trials the factfinder will be unable to say either that plaintiff has proved F by the preponderance of evidence or that defendant has proved not-F by the preponderance of evidence. Call this the indiscernibility assumption. It can be accommodated easily by adding another set of p-bars to the scale that is used in the basic analysis of data about factfinders' evaluations. This set of p-bars marks the number of trials in which factfinders are unwilling to assign any of the six canonical evaluations to Proof/F. Instead, they give it a seventh evaluation: (X) neither F proved POE nor not-F proved POE. It is understood what having projectable data about these quandaries would be like, and there is no special difficulty in the idea that these data are sorted between trials in which F actually exists and trials in which not-F actually exists. Therefore, addition of this seventh set of p-bars still serves the program of the basic analysis.

The actual magnitude of the X bars for any issue is an empirical matter. However, some suppositions, again, are reasonable. It may still be assumed that triers' deliberations are efficacious and, in particular, that their evaluations are fairly acute. Therefore, the relative number of trials that leave factfinders in quandaries, though significant, is small for most issues.⁶⁰ Furthermore, if the set of trials sup-

negligence based on a showing that the accident involved railroad equipment and resulted in death. In *Henderson*, 279 U.S. 639, the Court invalidated a rule that shifted the burden of persuasion for the issue of negligence from plaintiff to defendant railroad on a showing of the same facts. In both cases, the preponderance-of-the-evidence standard of proof was applicable to the negligence issue, and the party who had to satisfy the standard for his side of the issue bore the burden of persuasion. Thus, there is only one practical difference between the two rules. Only if the jury found that neither defendant's proof nor plaintiff's proof satisfied the standard would the two rules make a difference in the result: under the *Turnipseed* presumption, defendant would prevail; with the *Henderson* shift of burden, plaintiff would prevail. The Court's treatment of this difference as crucial implies the belief that the allocation of a bare burden of persuasion has a significant influence on the outcome of trials. See also *Lavine v. Milne*, 424 U.S. 577, 585 (1976).

⁵⁹ J. MAGUIRE, J. WEINSTEIN, J. CHADBOURN & J. MANSFIELD, *CASES AND MATERIALS ON EVIDENCE* 1008 (6th ed. 1973). This belief is controversial. See generally C. McCORMICK, *supra* note 12, § 336.

⁶⁰ "No extended experience at the bar or upon the trial bench is required to produce a vivid realization that only in the . . . rarer instances is [the jury's] determination based upon the allocation of the burden of persuasion." Morgan, *supra* note 36, at 59. See also

plying data for the basic analysis is assumed to be similar in other respects to the set of trials supplying data for this alternative analysis, the numbers represented by the bars of set *X* will decrease the numbers represented by sets *A* and *A'*. The supposed results of modified analysis for ordinary civil trials of issue *F* have the form of Figure 8. It illustrates that allocation of the bare burden of persuasion to plaintiffs distributes erroneous findings markedly in favor of defendants. Allocation of the burden to defendants would have equal and opposite consequences.

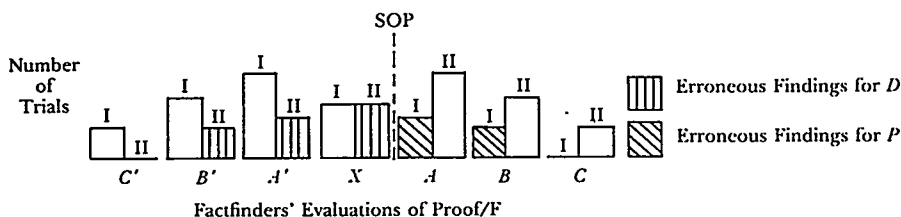


FIGURE 8

The existence and extent of indiscernibility in trials by common law standards of proof is a crucial issue of legislative fact. If indiscernibility is not assumed, lawmaking for the bare burden of persuasion is idle. Only the indiscernibility assumption and related other assumptions can support lawmaking to enforce a policy for errors by allocation of the bare burden of persuasion. The assumption is dubious and controversial, yet the Court, in such lawmaking, has neither argued for it nor even advanced it clearly. Therefore, even though due process requires that rules for burdens of proof distribute errors or risks of error according to the relative interests at stake in trials of fact, the Court's rulemaking has ignored a crucial part of the rationale.

C. THE BEYOND-A-REASONABLE-DOUBT STANDARD

In truth, assumptions of the kind that underlie p-bar analysis

Jaffe, *supra* note 30, at 10; Morgan, *Choice of Law Governing Proof*, 58 HARV. L. REV. 153, 191 (1944). Compare Morgan, *Choice of Law Governing Proof*, *supra*, at 191 n.89 with C. MCCORMICK, *HANDBOOK OF THE LAW OF EVIDENCE* § 322 at 686 (1954). Turnipseed, 219 U.S. 35, and Henderson, 279 U.S. 639, which seem to show a contrary view from the Supreme Court, see *supra* note 58, are both negligence cases. See Morgan, *Choice of Law Governing Proof*, *supra*, at 180, in which the author states: "The practical importance of fixing the burden of persuasion seems to have impressed courts and commentators more forcibly in negligence cases than in other litigation." Because the "fact" of negligence may not admit of probability or likelihood, see *supra* note 7, the factfinders' sense of relative likelihood is blunted, and they fall into significantly more quandaries about negligence than about other issues.

are all crucial issues of legislative fact for standards of proof. If findings about issue F are decisive in some kinds of criminal trials and assumptions about triers' evaluations of Proof/F indicate p-bars similar to those in either Figure 6 or Figure 8, then the Court's requirement of proof beyond a reasonable doubt will ensure that no innocent defendant is convicted. But, nine out of ten guilty defendants will be acquitted. Maybe the principles of fairness associated with due process permit, or even require, this result.

On the other hand, the Court has recognized that the innocent defendant may fairly be subjected to some risk of erroneous conviction, provided that the risk is justified.⁶¹ The state's interest in convicting the guilty provides such justification.⁶² Indeed, the state's interest may well justify a one-in-ten chance of erroneous conviction. If so, the requirement that F be proved beyond a reasonable doubt is excessive. As Figure 9 shows, proof by clear and convincing evidence exposes the innocent to a justifiable risk. Furthermore, the apportionment of erroneous verdicts at seven to one in favor of defendants is comfortably close to the traditional ideal.

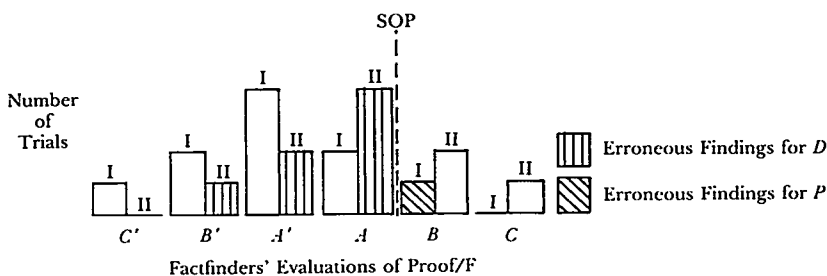


FIGURE 9

A convincing rationale for the requirement of proof beyond a reasonable doubt in criminal trials must deal with issues of ends and means. It needs not only sound brief for Blackstonian political morality, but also an argument that the beyond-a-reasonable-doubt standard of proof is necessary to realize the moral ideal. Such a rationale has not been offered, but it could be. As for ends, Blackstone's principle may be so well-established a conception of fair play and substantial justice that the Court's acceptance of it is unimpeachable. Due process may demand that errors be distributed in a 10:1 ratio favorable to the accused. As for means to this end, an

⁶¹ *Patterson*, 432 U.S. at 208.

⁶² *Id.*

argument for the beyond-a-reasonable-doubt standard must address the issues of legislative fact that underlie p-bar analysis.

There are very good reasons to assume that factfinders' evaluations of Proof/F in criminal trials differ from their evaluations of Proof/F in ordinary civil trials. For example, a criminal defendant is ordinarily brought to trial only after a preliminary finding of probable cause for the charge. Moreover, the resources and abilities of the parties are ordinarily unequal, with the advantage belonging to the state. Although it is assumed that F is a genuinely controversial issue, that just half of the defendants are actually guilty, and that triers do better than a coin-toss, it should not be supposed that p-bars for F resemble those of Figure 6 or Figure 8. The special characteristics of ordinary criminal trials imply that factfinders will give higher evaluations to Proof/F, even in cases in which the defendants are actually innocent. It may well be supposed, therefore, that p-bars for F resemble the p-bars in Figure 10. If so, proof of F beyond reasonable doubt is necessary to ensure the wanted apportionment of errors.

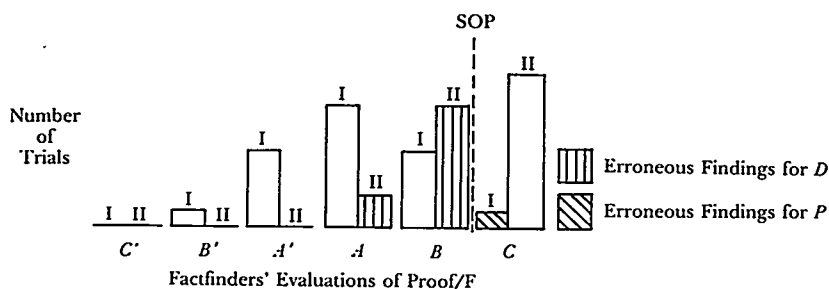


FIGURE 10

D. THE CLEAR-AND-CONVINCING-EVIDENCE STANDARD

The Court's requirement that some facts be proved by clear and convincing evidence also needs reason in policy and empirical considerations. It was speculated that the Court's policy for errors in proceedings for involuntary civil commitment and proceedings for termination of parental rights calls for a three-to-one apportionment in favor of defendants.⁶³ This policy will not be served by the prescribed standard of proof if factfinders' evaluations of proof conform to our assumptions about ordinary civil trials. Figure 9 shows that the clear-and-convincing-evidence standard will distribute errors in a ratio of seven to one. For all the Court has said, this policy

⁶³ See *supra* note 17 and accompanying text.

may be the actual policy for errors in such cases. However, nothing in the Court's opinions shows that the relative interests of defendants weigh heavy enough to justify such a policy. If the Court's lawmaking has a tenable rationale, it is found as much in special assumptions about factfinders' evaluations of proof as in special policies for errors.

The Court has had an inkling of this. In *Santosky v. Kramer*,⁶⁴ the Court described characteristics of New York's state-initiated permanent neglect proceeding to show how "numerous factors combine to magnify the risk of erroneous factfinding."⁶⁵ These included: the vague substantive standards and a tendency to bias; the inequality of the litigants' resources, skills, and access to evidence; and a one-sided ability to shape the historical record.⁶⁶ Not all of these characteristics are relevant to factfinders' evaluations of proof; the third, for example, is related to entrapment. But, unchecked bias and litigational power, combined in opposition to the defendant parents, are rightly assumed to cause relatively higher evaluations of Proof/Neglect, even in cases in which no neglect actually exists. Furthermore, assuming that these factors are not offset by others, a standard of proof higher than the preponderance of the evidence may be needed to apportion erroneous findings as policy requires. Therefore, it may well be supposed that the data on triers' evaluations of Proof/Neglect are displayed by Figure 11, where the requirement of clear and convincing evidence is shown to apportion errors in a 3:1 ratio favorable to defendants.

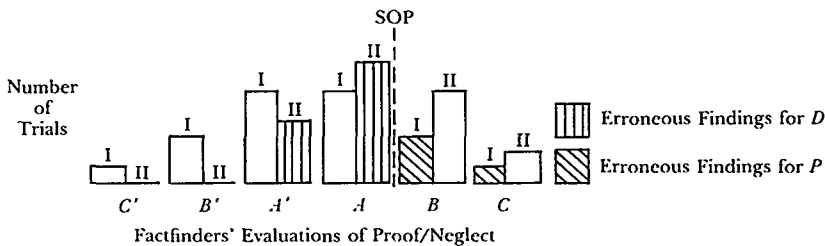


FIGURE 11

The form of the argument in *Santosky* suggests that the Court has recognized, if only vaguely, that its lawmaking for burdens of proof needs better rationalization. Although due process requires that the standard of proof enforce a policy for factfinding errors that corresponds to the relative interests at stake in certain kinds of tri-

⁶⁴ 455 U.S. 745 (1982).

⁶⁵ *Id.* at 762.

⁶⁶ *Id.* at 762-63.

als,⁶⁷ the policy alone does not justify the adoption of any particular standard of proof. Justification also demands tenable assumptions about the trials' general characteristics, and a sound argument that these characteristics will affect the distribution of factfinders' evaluations of proof in a particular way. Good reasoning in favor of some standard of proof needs not only well-founded conclusions about policy, but also well-founded conclusions about legislative fact. The *Santosky* opinion is closer to satisfying these demands.

III. IMPLICATIONS FOR DUE PROCESS

Implicit in the preceding argument is an explication of lawmaking for burdens of proof. As such, it is a model of rationality in lawmaking, and it is also a reconstruction of actual practice. Commentators note that burdens of proof are founded on considerations of policy.⁶⁸ These considerations are explicated by the policy for factfinders' erroneous decisions as a wanted ratio of errors. Commentators note also that burdens of proof are founded on such considerations as probability and relative access to evidence.⁶⁹ These considerations are explicated by the general assumptions about trials of an issue that may warrant special assumptions about factfinders' evaluations of proof.

The model is adequate, notwithstanding any controversy about the particular assumptions of legislative fact or the particular policies for errors adopted in the preceding illustrations. Its adequacy does not depend on the assumptions' truth or the policies' acceptance; nor does it depend on judgments about the reasonability of these assumptions or policies. Rather, the model is adequate as an acceptable logic of lawmaking for standards of proof. As such, the model serves due process review of burdens of proof. It makes plain the relation of lawmaking for burdens of proof to the traditional notions of fair play and substantial justice associated with due process.

The general form of constitutional challenges to burdens of proof is apparent. A burden of proof is inconsistent with due process only if it apportioned factfinding errors in an impermissible way. Perhaps some allocations of the burden of proof belong to categories that justify presumptions of unconstitutionality. A successful attack on any other allocation of the burden of proof, however,

⁶⁷ See *supra* note 3.

⁶⁸ F. James & G. Hazard, *supra* note 55, § 7.8, at 324-25; C. McCormick, *supra* note 12, § 337, at 952; Cleary, *supra* note 36, at 11-13; Morgan, *Some Observations Concerning Presumptions*, 44 HARV. L. REV. 906, 910-11 (1931).

⁶⁹ See *supra* note 68.

requires two things. First, it must establish a set of reasonable assumptions about trials in which the standard of proof is applied to an issue, upon which suppositions about factfinders' evaluations of proof may be based. Second, it must show that, upon any such tenable suppositions, the challenged standard of proof would apportion errors about the issue in a ratio that violates due process principles. If reasonable assumptions of legislative fact make it appear that the challenged standard apportions errors in a ratio consistent with due process principles, then the attack fails.

This Article offers no theory of reasonability for assumptions of legislative fact. The preceding discussion has exhibited the form and content of such assumptions. It need only be said again that the truth of such assumptions is an empirical matter, though their appropriateness and acceptability may be also a matter of ideology. Nor does this Article offer a theory of due process principles governing the apportionment of errors, though more should be said about how such principles may be critiqued.

If someone may be deprived of life, liberty, or property upon a finding of fact at issue in a trial, the burden of proof for that fact must be consistent with due process. Regarding policy for errors, due process implicates general principles of efficiency and fairness. The burden of proof, therefore, should serve to minimize errors, subject to constraints of distributive justice on the errors' incidence. These constraints are a matter of individual right to fair treatment. Lawmaking for burdens of proof is thereby bound to adopt a policy that shows respect for differences of interests affected by factfinding errors. Although no person has a right to a burden of proof that apportions errors in any particular ratio, he has a right to a burden of proof allocated upon consideration of his stake by a disinterested lawmaker whose deliberations are guided by a defensible theory of value. If the lawmaker has thereupon adopted a policy for errors about the issue of fact, this policy satisfies the requirements of due process.

The Court's several policies for errors may well be satisfactory. However, acceptance of any apparent policy of the Court is not acceptance of it as required by due process. The traditional principles of justice and fairness associated with due process may be generous enough to accommodate a range of policies for errors, any of which would pass constitutional muster. Furthermore, it would be wrong to assume that all issues in trials of some kind should be treated alike, so that, for example, erroneous findings about all facts necessary for criminal convictions are subject to the same policy.

In this light, the Court's review of burdens of proof in criminal

trials appears even more inappropriate. As a matter of course, the Court has ignored special considerations of legislative fact that may support an allocation of the burden of proof to enforce an acceptable policy. The Court has also ignored the special considerations of justice or fairness that may support a policy for errors about some element, even though that policy is less favorable to the criminal defendant. Because of its assumption that triers evaluate evidence with Bayesian rationality and that there is general proportional uniformity between the different interests at stake, the Court has ordained the beyond-a-reasonable-doubt standard of proof for all elements of a crime.⁷⁰ The consequences for due process jurisprudence have been appalling. A legislature's deviation from the Court's ordinance, which may well be justified by reasonable assumptions about factfinders' evaluations of proof or special considerations of distributive fairness in the apportionment of errors, engages only formal criteria to determine whether or not the affected material fact is an "element."

The theory of due process needs an acceptable model of lawmaking for burdens of proof—a model that accommodates fair, specific policies for errors and reasonable, specific assumptions of legislative fact. When due process review of burdens of proof is concerned with each of these two elements and is constrained by an appropriate deference to legislative judgment within well-founded theories of fairness and reasonability, the Court's lawmaking activity may end.

⁷⁰ *Winship*, 397 U.S. at 361.