

BASERATES, THE PRESUMPTION OF GUILT, ADMISSIBILITY RULINGS, AND ERRONEOUS CONVICTIONS

Michael J. Saks & D. Michael Risinger***

2003 MICH. ST. DCL L. REV. 1051

TABLE OF CONTENTS

INTRODUCTION	1051
I. BASERATE IMPACT IN LITIGATION: THE INVESTIGATIVE STAGE	1055
II. BASERATE IMPACT IN LITIGATION: THE NEGOTIATION STAGE	1058
III. BASERATE IMPACT IN LITIGATION: THE MOTION STAGE	1059
IV. BASERATE IMPACT IN LITIGATION: THE TRIAL STAGE	1061

INTRODUCTION

In our other co-authored contribution to this symposium, we questioned the commitment of the litigation system to rationality, and we also raised the issue of the fitness of humans to rationally process information under certain conditions, citing the debate between Amos Tversky, Daniel Kahneman and Gerd Gigerenzer.¹ This little piece takes up one ubiquitous and protean problem where Tversky and Kahneman may have the advantage over Gigerenzer in the courtroom.

An important but often largely invisible part of decision-making resides in the baserates on which much decision-making depends. The assumed (and often implicit) baserate affects the outcome of the decision being made. If the baserate is known or assumed to be low, the decision made is more likely to be negative; if the baserate is known or assumed to be high, the decision is more likely to be positive.

* Michael J. Saks is Professor of Law and Professor of Psychology at the Arizona State University, where he also is a Faculty Fellow at the Center for the Study of Law, Science, and Technology. Arizona State University, B.A., B.S., Pennsylvania State University, 1969, M.A., Ohio State University, 1972, Ph.D., Ohio State University, 1975, M.S.L., Yale Law School, 1983.

** Professor of Law, Seton Hall University School of Law, B.A., Yale University, 1966, J.D., Harvard University, 1969.

1. See D. Michael Risinger & Michael J. Saks, *Rationality, Research and Leviathan: Law Enforcement-Sponsored Research and the Criminal Process*, 2003 MICH. ST. DCLL. REV. 1023, 1025 n.12 (2003) and accompanying text.

Consider some illustrations of the impact of baserate assumptions. If a laboratory technician scrutinizing suspected cancer cells knows that the patient is young, the technician also knows that the incidence of cancer is low and the technician is less likely to judge the cells to be cancerous, whereas if the technician knows that the patient is old, the technician also knows that the incidence of cancer is much higher and the technician is more likely to judge the very same cells cancerous.² When radar operators and officers are protecting a ship during times of hostility, their expectation (baserate assumption) implicitly increases that approaching planes are unfriendly, and they are therefore more likely to perceive a threat and fire upon the approaching plane. During periods of calm their implicit expectation of hostility is reduced, they perceive less threat, and are less likely to order an attack on an approaching plane.³ Criminalists who look at two fragments of glass (one taken from a shattered crime scene window and the other from a suspect's outer clothing) which share every measured physical characteristic⁴ will judge these similarities to be more inculpatory if they believe there are many different types of glass in the "population" than if they think a few types are quite common.⁵ Mistaken assumptions about baserates, be they low or high, will lead to more erroneous decisions in one direction or the other.

Strongly skewed baserates present another interesting problem. Tests for detecting events that occur with relative rarity (e.g., diseases, medical malpractice, criminal conduct, violence by mentally ill persons, child abuse) will produce a high rate of false positive errors, no matter how accurate the ability to discriminate true from false, or positive from negative, outcomes.⁶ For example, in medical diagnosis, a condition which is found in 1 of every 250 people and which is screened with a test that is 93% accurate will yield positive (disease present) results of which 5% reflect people who truly have the disease and 95% are false positives. That is a lot of error produced by a highly accurate test, and it results from the skewed baserate. But people intuitively (and mistakenly) tend to equate the likelihood of accurate outcomes with the accuracy of the test, without taking adequate account of the baserate

2. See Victoria L. Phillips, Michael J. Saks, & Joseph L. Peterson, *The Application of Signal Detection Theory to Decision-Making in Forensic Science*, 46 J. FORENSIC SCI. 294 (2001).

3. See John A. Swets, *The Science of Choosing the Right Decision Threshold in High-Stakes Diagnostics*, 47 AM. PSYCHOL. 522 (1992).

4. Specific gravity, refraction index, etc.

5. See E. F. Pearson et al., *Glass and Paint Fragments Found in Men's Outer Clothing—Report of a Survey*, 16 J. FORENSIC SCI. 283 (1971).

6. Unless the test is one hundred percent accurate.

and the impact it has on the outcomes of the decisions.⁷ For example, if a screening test were developed to identify children likely to become future felons, most people would be inclined to approve or disapprove of the test based on the test's "accuracy". Yet even if the test were highly accurate, it still would yield a high rate of false positive errors because of the skewed baserate: a very high ratio of future non-felons to future felons.

A number of formal models have been developed in various literatures addressing the impact that baserates and assumptions about baserates have on decisions, such as Signal Detection Theory and models based on Bayes Theorem.

Signal Detection Theory has demonstrated that decision-makers with equal ability to perceive stimulus information, but who form different implicit decision thresholds for deciding when evidence is sufficient to declare a signal to exist or not to exist (be it the detection of cancer, enemy aircraft, or a forensically meaningful match), will produce different decision profiles. Where the decision-maker knows or assumes that signals are present with a high probability, a lenient threshold will tend intuitively to be set (because the opportunity to score many hits is felt to be high and the risk of false positives is felt to be low). Where the decision-maker knows or assumes that signals are relatively rare, the threshold will tend intuitively to be set higher because the opportunity to score hits is felt to be low and the risk of false positives is felt to be high.⁸

In Bayes' Theorem, the baserate is contained in the initial prior probability. The impact of that prior probability is well known among evidence scholars.⁹ One well developed example has involved paternity testing, where Bayesian analysis was performed to determine the "probability of paternity."¹⁰ How high or low analysts set the prior probability (the probability without taking into account any case-specific evidence) that the man suspected of paternity is indeed the father has a large impact on the

7. See Michael J. Saks & Jonathan J. Koehler, *What DNA "Fingerprinting" Can Teach the Law About the Rest of Forensic Science*, 13 *CARDOZO L. REV.* 361 (1991).

8. See Phillips, Saks, & Peterson, *supra* note 2. This is captured in the old aphorism of medical diagnosis that when you hear hoofbeats, you think horses, not zebras.

9. See, e.g., Symposium, *Decision and Inference in Litigation*, 13 *CARDOZO L. REV.* 253 (1991); see generally Symposium, *Probability and Inference in the Law of Evidence*, 66 *B.U. L. REV.* 377 (1986); Richard O. Lempert, *Modeling Relevance*, 75 *MICH. L. REV.* 1021 (1977).

10. DAVID L. FAIGMAN, DAVID H. KAYE, MICHAEL J. SAKS, & JOSEPH SANDERS, *The "Probability of Paternity,"* in *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* § 26-1.5 (2002) [hereinafter *MODERN SCIENTIFIC EVIDENCE*]; see also D.H. Kaye, *The Probability of an Ultimate Issue: The Strange Cases of Paternity Testing*, 75 *IOWA L. REV.* 75 (1989).

eventual probability that emerges from the testing and analysis. Traditionally that value has been set at fifty percent,¹¹ but that is not an empirically derived number in any sense, and other, lower values have also been used. As a result, in response to identical evidence, laboratory analysts who assume that the initial probability of paternity is equal to the probability of non-paternity are more likely to reach a conclusion of paternity than those who begin with a lower prior probability (and less likely than those who begin with a higher prior probability).¹²

Base rate assumptions can, of course, be found at work inside the trial system, and their effects can be noted (or overlooked) at every stage of the legal process. For example, consider judgments of the accuracy of the tort system in sorting cases of alleged medical malpractice. Defendants condemn the system, pointing to the finding that four of every five malpractice filings are judged to be without merit, but such a result is virtually unavoidable given the base rates which form the input to the legal system's sorting process. The base rate of negligently injured hospital patients is about one percent, which means of course that ninety-nine percent of patients have not been victims of medical malpractice.¹³ Of those negatives, there is only one false positive filing error in every thousand cases of non-malpractice. On the other hand, of the positives (the patients who are truly the victims of malpractice), only one in thirty-five files a claim. These are highly favorable odds for the tortfeasors.¹⁴ That ratio of false positives to true positives is 4:1 and is as much an unavoidable result of the underlying base rate as is the high rate of false positives even with a highly accurate medical screening test, such as the one discussed earlier.

Base rates come in two forms: objective and subjective. Objective base rates reflect the actual frequencies of positive and negative cases that

11. On what authority, one might ask? This is not an empirically determined probability, but a guess and a convention. Given that, should it not be under the control of the legal factfinder rather than being delegated by default to experts who have no expertise on the question?

12. See FAIGMAN, KAYE, SAKS, & SANDERS, *supra* note 10; see also Kaye, *supra* note 10.

13. See HARVARD MEDICAL PRACTICE STUDY, PATIENTS, DOCTORS, AND LAWYERS: MEDICAL INJURY, MALPRACTICE LITIGATION, AND PATIENT COMPENSATION IN NEW YORK (1990).

14. The risk of being a defendant in a suit filed by a patient who *has not* been a victim of malpractice is .001 and the risk of being a defendant in a suit filed by a patient who *has* been a victim of malpractice is .029. Both values are quite low; thus, the risk of a doctor being sued is very low whether the doctor committed malpractice or not. And the risk of being sued by a patient who *has been* the victim of malpractice is twenty-nine times as great as the risk of being sued by a patient who *has not been* the victim of malpractice.

enter the decision-making system, as shown by some form of quantified data. Subjective baserates are assumptions made by decisionmakers about the frequencies that underlie any given decision, not based on any quantified data, but merely reflecting impressions derived from experience. Subjective baserates also come in two forms, explicit and implicit. Explicit subjective baserates are consciously conceived and expressed, sometimes by magnitude words like "common" or "uncommon," sometimes by using numerical expressions of probability in a subjective and metaphorical way. Implicit baserates are not specifically expressed, but they are nevertheless present in many decisions by implication. Such implicit baserates are the hardest to see, and so their effects, though profound, may be elusive. Errors caused by erroneous implicit baserate assumptions are therefore the most difficult errors to identify and correct.

I. BASERATE IMPACT IN LITIGATION: THE INVESTIGATIVE STAGE

Various observers have noted that when a suspect becomes salient to the police, investigators focus too quickly and too exclusively on that individual.¹⁵ Investigators pursue a confirmatory strategy.¹⁶ They search for facts which are consistent with the hypothesis of guilt, a strategy which is more prone to mistaken inculpatory conclusions than a disconfirmatory investigative strategy. Adopting instead the disconfirmatory investigative strategy which is followed (ideally, at least) in conventional science,¹⁷ one would test the hypothesis of guilt by pursuing a falsification strategy. If the suspect is indeed the culprit, certain facts should not be so. Those would be inquired into and,

15. See also C. RONALD HUFF ET AL., *CONVICTED BUT INNOCENT: WRONGFUL CONVICTION AND PUBLIC POLICY* (1996); Arye Rattner, *Convicted But Innocent: Wrongful Conviction and the Criminal Justice System*, 12 L. & HUM. BEHAV. 283 (1988); MARTIN YANT, *PRESUMED GUILTY: WHEN INNOCENT PEOPLE ARE WRONGLY CONVICTED* (1991). The concept was most famously captured in Johnny Cochran's "rush to judgment" assertions in the trial of O.J. Simpson.

16. ARTHUR S. REBER, *THE PENGUIN DICTIONARY OF PSYCHOLOGY* 151 (2d ed. 1995) (defining confirmation bias as "[t]he tendency to seek and interpret information that confirms existing beliefs").

17. See *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 593 (1993) (citing discussions of falsifiability in treatises written by professors M. Green, C. Hempel, and K. Popper).

if they are found, the hypothesis is weakened. If disconfirming evidence cannot be found, the hypothesis is strengthened.¹⁸

Students of the problem of erroneous convictions have characterized many erroneous outcomes as being the result of a “presumption of guilt.” Reviews of such cases often have found that investigators too hastily and on too little evidence came to believe in their own hypothesis of guilt, searched for facts consistent with that hypothesis, and on finding some of those facts came to believe more firmly in the suspect’s guilt.¹⁹ Had they proceeded by subjecting their hypothesis to potential disconfirmation they would have been more likely to discover its weaknesses, and an erroneous prosecution and conviction would have been less likely.²⁰

These phenomena also can be understood as a failure to take proper account of the baserate. When a prime suspect comes into the investigative cross-hairs, and a fact consistent with guilt is found, the investigator could ask: How likely would the existence of this fact be if this suspect *were not* the perpetrator? Another way to state this is: How many other people would yield the same sort of evidence?²¹ The higher that probability (or frequency), the less inculpatory value the evidence has. Investigators no doubt have some intuitive feeling for these values—they do not regard any and all facts about the suspect as inculpatory. However, they usually do not address the issue directly and explicitly, but only implicitly and intuitively.

18. As to the generally understood primacy of skeptically proceeding by attempting to falsify, see DONALD B. CALNE, *WITHIN REASON: RATIONALITY AND HUMAN BEHAVIOR* 220 (1999) (“If the working hypothesis withstands all attempts to refute it, new knowledge can be claimed.”). Of course, the falsification strategy is unfamiliar to the common sense reasoning of most people, including most police investigators and judges. See *Daubert*, 509 U.S. 579, 598-601 (Rehnquist, C.J., concurring in part, dissenting in part).

19. See HUFF ET AL., *supra* note 15; see also Rattner, *supra* note 15; see also YANT, *supra* note 15.

20. The power of the “presumption of guilt” is seen most clearly in cases where strongly exculpatory evidence comes to light during the search for inculpatory evidence and is discounted or ignored rather than recognized as weakening the hypothesis that the suspect is the perpetrator. In *United States v. Brown*, No. CR 99-184 ABC (C.D. Cal. Dec. 1, 1999) (unpublished order), numerous fingerprints were found on the vital documents of the case, but none belonged to the defendant; and security camera pictures of the perpetrator were of someone other than the defendant. In *United States v. Rutherford*, 104 F. Supp. 2d 1190 (D. Neb. 2000), two different eyewitnesses to the criminal act at issue excluded the defendant, and there was evidence that the defendant was far from the crime scene when it occurred. In neither case did the exculpatory evidence shake the government’s belief in its hypotheses concerning these defendants, and in both cases the government went to trial with little more than a handwriting expert’s opinion that the defendants were the writers of the documents in question in the cases.

21. For example: similar clothing, similar geographic location, similar opportunity, similar motivation, similar attributes for purposes of (eyewitness or forensic) identification.

Bayesian theorists would ask an additional question: How likely would the existence of this fact be if this suspect *were* the perpetrator. Then they would create a likelihood ratio out of the answers, with the answer to the second question as the numerator. If the resulting likelihood ratio were small, the investigator would know that the item of evidence did not actually prove much. However, influenced by their “presumption of guilt” methodology, investigators tend to overvalue seemingly inculpatory evidence and deprive themselves of the opportunity to more accurately evaluate potentially exculpatory evidence.²²

Signal detection theorists would note that here, and throughout the legal process, assumptions of low baserates regarding the existence of apparently inculpatory circumstances that are in fact more common (as well as the investigator’s motivation to prove a favored hypothesis) lead to a downward shifting of the decision threshold, increasing the probability of reaching a positive conclusion (both true positives and false positives).

This problem is manifested in crime laboratories in several ways. First of all, the theory of individualization depends entirely on the frequency of occurrence of various attributes in the population (that is, the baserate of attributes of fingerprints, handwriting, DNA and so on). That an apparently matching “questioned exemplar” found at a crime scene (such as a hair) and a “known exemplar” (that is, known to be from the defendant) actually came from the same person or other common source, is more likely to the same degree that the baserate of attributes in the population is small. The more uncommon the attributes, the less likely a coincidental match.

One might assume or expect that “experts” were in possession of good information about the baserates upon which their conclusions depend. But only DNA examiners have explicit objective data on such population baserates, and so they can make an accurate report as to the probability of a coincidental match. Unlike DNA analysis, however, other individualization specialities rely on subjective impressions and assumptions about baserates, generally assuming them to be very low. They often go even further, and assume that, essentially, there is no baserate aside from the instance in front of them, that is, they assume that there is no other person or object in the population that could have the same features of hair, handwriting, striations, or whatever. These assumptions about tiny or nonexistent baserates result in an underestimation of the probability of coincidental matches²³ and increase the chances of an erroneous inculpation.

22. See *infra* notes 23-25 and accompanying text.

23. See FAIGMAN, KAYE, SAKS, & SANDERS, *supra* note 10, §§ 27-31 at 346. See also Michael J. Saks, *Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science*, 49 HASTINGS L.J. 1069 (1998).

An additional interesting problem exists for individualization analysts in crime labs. They have good reason to believe that evidence is not submitted to them at random, but instead comes from investigators who have been led by other incriminating facts to think that the evidence they are submitting will link suspects to crimes. Thus, under the procedures of the conventional crime lab,²⁴ examiners inevitably believe that the likelihood is high that any given item of submitted evidence will be inculpatory, which will lead to a lowering of the decision threshold for declaring positives, and, all else being equal, a higher risk of declaring evidence to be inculpatory when it is not. Put differently, the examiner implicitly assumes before analysis that there is a high likelihood that the evidence, after it is analyzed, will result in an inclusion. Some suggestion that this might be happening comes from a detailed study of four different crime laboratories that found that, on average, fewer than 10% of all reports disassociated a suspect from the crime, a rate that seems low.²⁵

II. BASERATE IMPACT IN LITIGATION: THE NEGOTIATION STAGE

Litigators settle many more cases than they try. Such negotiations are thought to be conducted “in the shadow of the law.” They also take place “in the shadow of the jury.” In other words, the way that cases are resolved in negotiations is conditioned on expectations about how the cases will be resolved if they go to trial. This too can be viewed as a baserate problem. The problem of interest here is, which baserate? For example, litigators can have in mind the risk of losing associated with the entire category of tort cases or more refined categories such as auto crash cases or product liability cases, or even more refined categories, taking into account more and more attributes of the particular case at hand. The more refined the category of cases used as the reference group for the case at issue, the more accurate the predictions of

24. For alternative procedures that would tend to eliminate this problem, see D. Michael Risinger, Michael J. Saks, William C. Thompson, & Robert Rosenthal, *The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion*, 90 CAL. L. REV. 1 (2002).

25. See JOSEPH L. PETERSON ET AL., *FORENSIC EVIDENCE AND THE POLICE: THE EFFECTS OF SCIENTIFIC EVIDENCE ON CRIMINAL INVESTIGATIONS* 117 (1984). One suggestion for overcoming this problem is that evidence be presented in the form of a lineup, so as to radically alter the baserates for examiners. In a lineup, the examiner knows that there is only one chance in five or seven or ten (depending on how many foils are added to the evidence lineup) that guessing or assuming will by itself produce a correct result, and that errors will be visible to all. By contrast, the usual analysis is done in a show-up format with a high assumed baserate (a high prior probability) for inculcation. In any event, in the normal procedure there is little or no possibility for observers to know when the examiner has made an error. See Risinger, Saks, Thompson, & Rosenthal, *supra* note 24.

outcome will be. But if they become too fine-grained, their predictive potential will degrade, as the sample size in the relevant sub-category becomes smaller and smaller. The challenge is to find the optimal sub-category to treat as the reference population for the case at issue, where predictive power will be maximized.

Most clearly in civil cases, an interesting additional concern is to discount the amount of damages by the probability of a plaintiff's verdict, in order to estimate the expected value of the case. Again, the baserate is the key to the best approximation of an answer, to the extent that an accurate baserate can be determined.²⁶

III. BASERATE IMPACT IN LITIGATION: THE MOTION STAGE

Commentators have puzzled over the apparent phenomenon that in ruling on motions concerning the admission or exclusion of evidence, all else being equal, courts tend to grant motions made by civil defendants and prosecutors, and to deny motions made by civil plaintiffs and criminal defendants.²⁷ While there are many explanations for the disparate rulings, they have been especially hard to explain, and particularly apparent, in *Daubert* hearings at which the various parties seek to have expert evidence excluded or admitted.²⁸ Though scientifically literate commentators seem to agree that, in general, forensic science evidence offered by government prosecutors is among the weakest evidence offered to courts—weaker, for example, than evidence of causation offered by tort plaintiffs²⁹—civil defendants have been far more successful in having evidence excluded than criminal defendants.

Baserate assumptions may be one of the factors at work in these situations. First, why would courts admit forensic science expert testimony on behalf of which only scanty research can be offered to try to meet the tests posed by *Daubert*? Part of the answer might be that judges are filling the data gaps with their own speculative—but strongly felt—senses of the rates of accuracy and error. Without any actual data with which to evaluate, for

26. Despite much folklore to the effect that lawyers and insurance settlement agents “know the value of a case” and can predict case outcomes with impressive accuracy, empirical studies exploring these predictive skills have failed to find much more than noise, until the case comes very close to the time of resolution. See discussion and studies cited in Michael J. Saks, *Do We Really Know Anything About the Behavior of the Tort Litigation System – And Why Not?*, 140 U. PA. L. REV. 1147, 1223 n.263 (1992).

27. See D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 ALB. L. REV. 99 (2000).

28. See *id.*

29. Compare relevant chapters (that is, epidemiology and toxicology versus forensic science) in FAIGMAN, KAYE, SAKS, & SANDERS, *supra* note 10.

example, bitemark expert testimony, what would a court imagine it to be? No doubt the assumed error rates are far smaller than the data forensic dentists have thus far reported would indicate.³⁰ Judges who have approached *Daubert* analysis by reversing the burden of proof – responding to the proponent’s failure to produce supportive data by requiring the opponent of admission to prove that error rates are not low, and otherwise to admit the evidence,³¹ seem likely to be making such generous baserate assumptions. Where these notions—the assumed risk of error in one direction or the other—come from we cannot say. Perhaps from Sherlock Holmes and Patricia Cornwell and CSI, or from campaigns by forensic scientists to implant exaggerated impressions of accuracy in the minds of the public and the legal profession.³² But from whatever source, they exist—and they affect decisions.

A second and more speculative, though perhaps more obvious, baserate that may be contributing to these decisions is that reflected in the ratio of outcomes in a category of cases. If a judge knows nothing else about the facts of a criminal case, the judge knows that at the end of the day eighty to ninety percent of defendants will be found guilty.³³ Awareness of this baserate might influence the judge’s interpretation of evidence, much as knowing the age of a patient can influence the cancer lab technician’s interpretation of the cells on the patient’s slide. Thus, awareness of the baserate of criminal trial outcomes might affect a judge’s assessment of the merits of a *Daubert* challenge, if only by changing the decision threshold: what it takes for a proponent to cross the threshold of admissibility.³⁴ Some judges may implicitly regard a ruling to exclude a seemingly highly inculpatory item of evidence as risking a distortion of the likely correct (judging from the verdict baserate) outcome of the case. Obviously, judges are supposed to let the trial evidence lead them to their judgments of a defendant's guilt or non-guilt, or

30. See C. Michael Bowers, *The Scientific Status of Bitemark Comparisons: Proficiency Testing of Board Certified Odontologists*, in MODERN SCIENTIFIC EVIDENCE, *supra* note 10, § 30-2.1.3[1] at 543 (reporting a false positive error rate of 64%).

31. See, e.g., *United States v. Starzeczyzel*, 880 F. Supp. 1027 (S.D.N.Y. 1995) (concerning proffered handwriting expert opinion); *United States v. Llera Plaza*, 188 F. Supp. 2d 549 (E.D. Pa. 2002) (concerning proffered fingerprint expert opinion).

32. See SIMON A. COLE, *SUSPECT IDENTITIES: A HISTORY OF FINGERPRINTING AND CRIMINAL IDENTIFICATION* (2001); see also John I. Thornton, *Criminalistics - Past, Present, and Future*, 11 LEX ET SCIENTIA 1 (1975).

33. Federal juries returned convictions in 84% of the criminal cases submitted to them for a verdict. BUREAU OF JUSTICE STATISTICS, U.S. DEP’T OF JUSTICE, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS 2001 (Ann L. Pastore & Kathleen Maguire eds., 2002) (Table 5.17, Disposition of cases terminated in U.S. District Courts, fiscal year 2000).

34. Or what it takes for an opponent to prove non-admissibility to those judges who invent such a requirement.

in a jury trial to not even concern themselves with ultimate outcomes, but expectations borne of baserates could make some evidentiary rulings in the case at bar more likely than others.

To suggest a possible explanation for decisions is not to suggest that they are proper or justifiable. A fair amount of what the law of evidence and procedure tries to do is to correct for frailties in human decision-making. Rulings about the admission of evidence in the case at bar should not be an artifact of beliefs about the outcomes of all other cases. We have elsewhere suggested how distortions in decision-making could be corrected in the work of crime laboratories.³⁵ In trials, the most effective tool for accomplishing those corrections has likely been the bifurcation of decision-making between judges and juries, to which we turn next.

IV. BASERATE IMPACT IN LITIGATION: THE TRIAL STAGE

Juries—or any factfinders separate from the investigators and the court—are the law’s last best chance to undo distortions produced by the processes described above. The instruction to juries in criminal cases to adopt a presumption of innocence and to require proof beyond a reasonable doubt is a dual attempt to undo the possible distorting effects of baserates. The “presumption of innocence” aims to replace any baserate assumptions that jurors might have which might draw them toward expectations of guilt. It seeks to impose a prior probability approaching zero.³⁶ The instructed standard of proof is a direct attempt to induce the factfinders to set a high decision threshold.³⁷

When asked to predict the likelihood that they would end up voting to convict in a criminal case on which they do not yet have any evidence, mock

35. See Risinger, Saks, Thompson, & Rosenthal, *supra* note 24.

36. Compare this to paternity testing analyses which begin with a 50:50 presumption (or prior probability or baserate). See *supra* notes 10-12 and accompanying text. The reason we say “approaching zero” is to account for the simple mathematical fact that if the prior probability were truly zero, it could not be affected or revised by new information.

37. On the other hand, factfinders are presented with evidence the evaluation of which requires them to fill in baserates of their own. For example, the perpetrator of a crime is said to have been wearing a pair of brand X sneakers and the defendant is found to own a pair of brand X sneakers. How inculpatory that pair of facts is depends on how common brand X is. If most of the folks in town own brand X, the evidence does not zero in sharply on the defendant, but if brand X is an uncommon brand owned by very few in the area, then the defendant can be more tightly connected to the crime scene. Many of the facts jurors are required to evaluate require them to exercise their intuitions about the commonness, and therefore the diagnosticity of such facts, rather than to be given empirical data on those necessary foundational facts.

jurors offer the prediction of a 50% chance of voting to convict.³⁸ That is not a particularly low prior probability, but it is considerably lower than the 85:15 ratio that judges expect due to their knowledge of criminal trial outcome baserates. However, it is misleading to think that jurors thereby do not honor the “presumption of innocence” based on these responses. Some evidence exists to suggest that jurors set their prior probabilities lower than they think they do. Research that has come at the question more indirectly, by extrapolating backwards to what the mock jurors’ starting points must have been before they had any evidence, has found that the jurors had starting assumptions very close to zero (innocence), but to which they attached very little weight, so that the presumption of innocence was abandoned as soon as the first piece of inculpatory evidence was presented.³⁹

However, once at least certain types of evidence are admitted—even weak evidence such as expert opinion predicting a defendant’s future dangerousness⁴⁰—there appears to be little that can be done to undermine its impact. Research which has examined the capacity of weak cross-examination, strong cross-examination, and strong cross plus rebuttal expert to undo the effects of expert evidence has found that none of it made much difference. Once the expert evidence was in, nothing undid its effects more than marginally.⁴¹ Similarly, studies of mock juror responses to evidence and arguments about damages find that once these are presented, their major impact is not undone by counter evidence and counter arguments by the other side.⁴²

These findings raise concerns about the adequacy of current practice in regard to giving jurors (or judges) information about the probative value of

38. See Thomas M. Ostrom, Carol Werner, & Michael J. Saks, *An Integration Theory Analysis of Jurors' Presumptions of Guilt or Innocence*, 36 J. PERSONALITY & SOC. PSYCHOL. 436 (1978).

39. See *id.*

40. See Shari Seidman Diamond et al., *Juror Reactions to Attorneys at Trial*, 87 J. CRIM. L. & CRIMINOLOGY 17 (1996). It is true that this study involved sentencing phase evidence of dangerousness, so that one should perhaps be cautious in generalizing to other trial contexts. However, results even in that context were startling. See also James M. Shellow, *Experts: The Limits of Cross-Examination*, 34 SETON HALL L. REV. 317 (2003).

41. See *id.*

42. See Allan Raitz et al., *Determining Damages: The Influence of Expert Testimony on Jurors' Decision Making*, 14 L. & HUM. BEHAV. 385 (1990); Mollie W. Marti & Roselle L. Wissler, *Be Careful What You Ask For: The Effect of Anchors on Personal Injury Damages Awards*, 6 J. EXPERIMENTAL PSYCHOL.: APPLIED 91 (2000). This is not to say that educating jurors to reduce the weight they have given to an initial witness, expert or non-expert, is not effective in any context. Studies have found expert testimony concerning eyewitness error to reduce jurors’ confidence in the accuracy of eyewitness identifications. See MODERN SCIENTIFIC EVIDENCE, *supra* note 10, § 15-2.3.2 at 256 (2002).

evidence, including information on the accuracy of expert witnesses. That is, what good does it do to tell factfinders that the evidence they just heard has a lower probability of being correct than they assumed it had? The research, though limited, seems to confirm the legal maxim about the impossibility of putting toothpaste back into its tube.

Thus, although the law of trials is designed to prevent jurors from being biased by knowledge of baserates, it often relies on cross-examination and rebuttal experts to undo the harm of misleading or exaggerated expert evidence, which may not work as well as the law imagines.⁴³ The most effective cure for that may be exclusion of the misleading expert testimony. That, in turn, takes us back to the problem of judicial decisions on *Daubert* motions and to the causes of erroneous or misleading expert testimony, which are themselves (partly) baserate problems in need of solutions.

43. But here is something to note: It appears that informing jurors about the baserate of expert accuracy after hearing the testimony is largely ineffective. So it is not baserate information per se, but what role it plays that seems to matter most. In the role of a prior probability, it appears to be potent. In the role of adjusting the weight of testimony already received, it is weak.