Laboratory Error Seen Through the Lens of Science and Policy

*Margaret A. Berger**

INTRODUCTION

Professor Redmayne, in his interesting and helpful Article, *Expert Evidence and Scientific Disagreement*,[1] discusses the blurry demarcation line between scientific and "trans-scientific issues," the term he uses for scientific controversies that also "incorporate policy questions."[2] In looking at the debate on the admissibility of DNA evidence, he gives generally high marks to the 1996 report of the second National Research Council Committee (NRC2)[3] for managing a "delicate balancing act between science and policy."[4] He suggests, however, that the Committee does not deserve accolades for its treatment of laboratory error rates.

After commenting on NRC2's refusal to report combined match and error probabilities, and its failure to endorse the first National Research Council Committee's (NRC1) recommendation[5] that jurors must be told of error rates on proficiency tests,[6] Professor Redmayne concludes that this omission is per-

---

[1] Mike Redmayne, *Expert Evidence and Scientific Disagreement*, 30 U.C. DAVIS L. REV. 1027, 1031-35 (1997).

[2] *See id.* at 1031.

[3] *See* COMMITTEE ON DNA FORENSIC SCIENCE: AN UPDATE, NATIONAL RESEARCH COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE (1996) [hereinafter NRC2].

[4] Redmayne, *supra* note 1, at 1028-61 (discussing various viewpoints concerning scientific controversies). Professor Redmayne further explains: "Throughout most of the report, care is taken to mark a distinction between science and policy. Rather than compromising its legitimacy by proposing answers to policy questions, the report frequently leaves them to be decided by the courts." *Id.* at 1060.

[5] *See* COMMITTEE ON DNA TECHNOLOGY IN FORENSIC SCIENCE, NATIONAL RESEARCH COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE 89 (1992) [hereinafter NRC1] (suggesting that laboratory error rates be continually estimated in blind proficiency testing and disclosed to juries).

[6] NRC2, *supra* note 3, at 185, explains that it does not repeat the recommendation in

1081

haps indicative of the science/policy distinction being "less clear than the Committee would have us believe."[7] I take this comment to mean that Professor Redmayne disagrees with NRC2's characterization of the error issue as a "policy" matter that should be left to the courts,[8] and that he finds fault with the Committee's unwillingness to take a stand, as it did in other areas of uncertainty, on how to resolve the issue.

I do not agree with Professor Redmayne's conclusion. To my mind, NRC2 correctly treated the legal implications of laboratory error as a matter beyond its competence and mandate.[9] Laboratory error, as this Comment seeks to show, is an issue that lies indubitably on the "policy" side of the blurry line. Indeed, the appropriate characterization and treatment of this issue is a perfect illustration of Professor Redmayne's thesis about the difficulty of achieving closure in scientific controversies, and the important role burdens of proof can play in handling scientific uncertainty. However, I conclude that, in this instance, shifting the burden of proof is not appropriate. The treatment of laboratory error in the manner NRC2 refused to endorse, which may at first glance look like a scientific solution, is not the offshoot of scientific principles. It is a policy-driven conclusion inextricably intertwined with complex evidentiary, procedural, and cost-benefit concerns.

As Professor Redmayne states, NRC2 rejected two different solutions to the laboratory error rate issue: first, combining the probability of a match due to laboratory error with random-match probabilities (RMP);[10] and second, requiring prosecutors to inform jurors of laboratory proficiency-testing rates as a condition of admitting DNA evidence.[11] Part I of this discussion

---

NRC1: "Inasmuch as the purpose of our report is to determine what aspects of the procedures used in connection with forensic DNA testing are scientifically valid, we attempt no policy statement."

[7] Redmayne, *supra* note 1, at 1062.

[8] *See id.* at 1061-62.

[9] I should divulge that I was a member of NRC2. *See* NRC2, *supra* note 3, at iii.

[10] *See id.* at 217. The "random-match probability" is defined in the NRC2 Report as "the probability that the DNA in a random sample from the population has the same profile as the DNA in the evidence sample." *See id.* I use the term "RMP" to refer to any quantitatively based assessment of the probability of a random match. I do not thereby take a position on whether or not the RMP should be expressed as a likelihood ratio. *See id.* at 199-202.

[11] *See id.* "Proficiency-testing entails the testing of specimens submitted to the laborato-

considers whether these suggestions are scientifically mandated. If they are, then the Supreme Court's opinion in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*[12] might dictate their adoption in federal court.[13] After concluding that "science" does not demand either outcome, Part II moves beyond the scope of the NRC2 report to consider whether courts ought nevertheless to adopt these proposals as appropriate prudential conventions for dealing with the reality of laboratory error. I conclude that both suggestions are repugnant to evidentiary principles and unhelpful. The repudiation of these particular solutions, both on scientific and policy grounds, in no way implies, however, that laboratory error is not a serious problem, a conclusion with which NRC2 strongly agreed.[14] The final section of this Comment discusses possible solutions that could be accommodated within the governing legal framework and that would result in the admission of more reliable scientific evidence.

I. IS LABORATORY ERROR A "SCIENTIFIC" FACT?

A. Must the Error Rate Be Combined with the RMP?

Requiring the rate of profile matching due to laboratory error to be combined with the RMP would mean that an expert witness would not be allowed to testify to statistics that refer solely to the probability that a person randomly selected from the population has the same DNA profile as the evidence sample.

---

ry in the same form as evidence samples." *Id.* at 78. The NRC Report discusses open and blind proficiency testing. *See id.* at 78-79. This Comment refers to both kinds of tests when it discusses proficiency testing.

[12] 509 U.S. 579 (1993). In *Daubert,* the Supreme Court stated that a trial court may not admit scientific evidence without inquiring "whether the reasoning or methodology underlying the testimony is scientifically valid." *See id.* at 592-93; *see also* Margaret A. Berger, *Procedural Paradigms for Applying the* Daubert *Test,* 78 MINN. L. REV. 1345, 1359-63 (1994).

[13] *Daubert* challenges to DNA evidence might also be made on the ground that the egregiousness of the laboratory error in the case at issue means that no valid conclusions can be drawn. *See Daubert,* 509 U.S. at 579-93. This kind of objection to admissibility on the basis of error is unrelated to the proposals that this Comment evaluates, and will not be discussed further.

[14] One entire chapter of the Report is devoted to recommendations on how to reduce the risk of laboratory errors. *See* NRC2, *supra* note 3, at 75-88. Three of the eleven recommendations made in the Report are directed at minimizing the risk of laboratory error. *See id.* at 88.

This result is the goal of a number of commentators who have strenuously urged that a report on the frequency of matching DNA profiles is meaningless unless laboratory error is taken into account.[15] Furthermore, they claim that scientific principles compel the conclusion that only a combined statistic is significant in a court of law.[16] Their position is perhaps best approached with the help of a beguiling, but I will argue, ultimately misleading, baseball[17] analogy offered by Professor Jonathan Koehler, a leading proponent of this view.[18]

Professor Koehler's example is as follows: Suppose, he says that the fielder "makes throwing errors fewer than one time in a million, but makes fielding errors two percent of the time."[19] If we are trying to estimate the error rate on the fielder's next attempt (due to either throwing or fielding) the chance of an

---

[15] *See, e.g.,* Barry C. Scheck, *DNA and* Daubert, 15 CARDOZO L. REV. 1959, 1997 (1994) (concluding that DNA "matches" require reliable estimate of error rate). Professor Scheck states: "DNA evidence of 'matches' should not be admitted unless a laboratory can offer a reliable estimate of its false positive error rate, and in most cases, that error rate should be the only probability offered about the likelihood that the defendant was not the source of DNA trace evidence." *See id.; see also id.* at 1981 & n.77 (listing other authorities subscribing to this view).

[16] *See id.* at 1981 ("Without having reliable statistical estimates of both of these events — DNA profile frequencies and laboratory error rates — the jury does not have a scientifically valid basis under *Daubert* for evaluating the probability of a match.").

[17] Since this Comment is part of a Symposium on International Perspectives on Scientific Evidence, a very abbreviated and simplified account of baseball follows so as to make this baseball analogy comprehensible to foreign readers. Basically in baseball, the objective of the batter (that's the one holding the baseball bat) is to bat the ball and run to a base. He is unsuccessful when he bats the ball and is called "out" if (1) a fielder catches the ball before it touches the ground, or if (2) the ball touches the ground but the fielder throws it to the base to which the batter is running, and the fielder at that base catches the ball before the batter arrives. A fielder, therefore, has to be good at two different skills: he has to be good at fielding the ball, by catching it or scooping it off the ground, and then he has to have a good throwing arm.

[18] *See* Jonathan J. Koehler, *Why DNA Likelihood Ratios Should Account for Error (Even When a National Research Council Report Says They Should Not)*, 37 JURIMETRICS J. 425 (1997) [hereinafter *DNA Likelihood Ratios*]; Jonathan J. Koehler, *Proving the Case: The Science of DNA: On Conveying the Probative Value of DNA Evidence: Frequencies, Likelihood Ratios, and Error Rates*, 67 U. COLO. L. REV. 859 (1996) [hereinafter *Proving the Case*]; *see also* Jonathan J. Koehler, *DNA Matches and Statistics: Important Questions, Surprising Answers*, 76 JUDICATURE 222 (1993) [hereinafter *DNA Matches*]; Jonathan J. Koehler, *Error and Exaggeration in the Presentation of DNA Evidence At Trial*, 34 JURIMETRICS J. 21 (1993); Jonathan J. Koehler et al., *The Random Match Probability in DNA Evidence: Irrelevant and Prejudicial?*, 35 JURIMETRICS J. 201 (1995); 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).

[19] Koehler et al., *supra* note 18, at 211.

error is at least two percent because he has to field the ball before he can throw it. Even if the fielder's throwing improves so that he now throws imperfectly only once in a hundred million times, his overall error rate will not improve. Professor Koehler concludes that "just as the infielder's two percent fielding error rate sets a lower bound threshold for error estimates, the forensic scientist's laboratory error rate sets a lower bound for false positive match reports."[20] In other words, according to Professor Koehler, even if the probability of a random match of the DNA profile in question is one in a hundred million, if the laboratory makes errors two percent of the time, the RMP is irrelevant.

The notion of using a lower bound to deal with uncertainty echoes an earlier DNA controversy. As Professor Redmayne recounts,[21] uncertainty over the issue of population structure led NRC1 to recommend the ceiling principle,[22] a method of calculating allele frequencies that placed a lower limit on the size of the profile frequency. Professor Redmayne also reports strong criticism "by some scientists, who argued that the ceiling principle 'lack[s] any logical basis' [and] accused the Committee of 'neglecting established genetic principle, misleading the courts and disgracing the National Academy.'"[23] NRC2 agreed with those "who see no scientific justification for its use."[24]

Of course, the ceiling principle sought to deal with reservations about the impact of population structure, a very different kind of uncertainty than that addressed by the combined error and frequency statistic. But the essence of why both solutions are unscientific is the same. Application of the ceiling principle

---

[20] *Id.*

[21] *See* Redmayne, *supra* note 1, at 1058-59.

[22] *See* NRC2, *supra* note 3, at 35-36. Actually, NRC1 discussed both a ceiling principle and an interim ceiling principle. *See id.* at 35. Both were criticized by many in the scientific community, and NRC2 concluded "that neither ceiling principle is needed." *See id.* The NRC2 Report also recognized that some courts might continue to use the interim ceiling principle, and offers suggestions on how to improve application of the principle. *See id.* at 156-59.

[23] Redmayne, *supra* note 1, at 1058 & n.122.

[24] *See* NRC2, *supra* note 3, at 35; *see also* Government of the Virgin Islands v. Byers, 941 F. Supp. 513, 522 (D.V.I. 1996) ("Perhaps the most damning response was the suggestion that the ceiling principle is not science at all but merely an arbitrary policy decision designed to placate the defense bar.").

resulted in an arbitrary calculation that could not be the correct probability.[25] A combined error and frequency statistic cannot provide the correct probability either. Even if we assume for the moment that it is possible to establish a laboratory error rate, the insistence on invariably applying a lower bound in every case will result in inaccuracies in the great majority of cases, as well as the "excessively conservative" results produced by the ceiling principle.[26]

An argument that the RMP is an arbitrary and untested model that somehow becomes more scientific with the incorporation of an error model is flawed in two respects. In the first place, the RMP model is not arbitrary and untested. Professor Redmayne reminds us that new scientific theories become acceptable when they are seen to provide "a coherent theoretical framework."[27] A consensus has emerged, in at least a significant portion of the scientific community, on ways to calculate an RMP. More and more data have been amassed since the 1992 NRC report that support the robustness of the model.[28]

Second, the suggested model for calculating laboratory error in DNA cases produces an inexact, arbitrary error calculation for a particular laboratory. Even were it desirable on policy grounds to temper the impact of the random match calculation, there is no scientific basis for adjusting the RMP in the manner proposed.

B. Are Laboratory Error Rates "Scientific"?

The laboratory error rate is an amalgam of two different kinds of errors — interpretative mistakes and flaws in laboratory performance.[29] Mistakes in scientific interpretation occur when

---

[25] *See id.* at 157 (explaining criticism of arbitrariness).

[26] *See id.* at 158.

[27] *See* Redmayne, *supra* note 1, at 1032.

[28] *See id.* at 1061-62; *see also* NRC2, *supra* note 3, at 2 ("Much has been learned since the last report. The technology for DNA profiling and the methods for estimating frequencies and related statistics have progressed to the point where the reliability and validity of properly collected and analyzed DNA data should not be in doubt."). References to new data are incorporated throughout the report, especially in Chapter 4, Population Genetics. *See id.* at 89-124.

[29] *See* Koehler et al., *supra* note 18, at 203 n.7 ("As used here, 'laboratory error' includes all relevant human and technical errors. These include mislabelings, misrecordings, misrepresentations, case mix-ups, contaminations, and various interpretative

a laboratory erroneously applies a scientific principle. Examples of interpretation errors are those that may occur in VNTR typing,[30] for instance, if the laboratory declares matches on autorads beyond what is scientifically acceptable, or incorrectly draws a conclusion when only a single band appears.[31] Interpretative mistakes occur in connection with other forensic techniques as well, but courts have never excluded scientific evidence for a failure to include an estimate of the frequency with which such errors occur.[32]

Performance errors occur when a laboratory mislabels, misrecords, or misrepresents evidence, or contaminates samples or mixes them up. These problems are not unique to laboratories — the police can make these types of mistakes before the suspect, victim, or crime scene samples arrive at the lab.[33] Nor are these problems limited to DNA evidence, or scientific evidence in general. These are the ordinary mistakes, due to sloppiness or fraud, that plague all forms of evidence. When humans are involved, as they always are in the production of evidence, the error rate cannot be reduced to zero. Indeed, the principal goal of many standard evidence rules — such as best evidence, authentication, and hearsay — is to reduce

---

errors.").

[30] For a description of VNTR typing, see NRC2, *supra* note 3, at 14-23; *see also* United States v. Jakobetz, 955 F.2d 786, 791-93 (2d Cir. 1992).

[31] Scientific interpretation errors are considered throughout the NRC2 report, and rules and procedures are recommended for handling troublesome situations that pose the risk of a flawed scientific judgment. *See* NRC2, *supra* note 3, at 139-56.

[32] For surveys of errors made in connection with a variety of different forensic techniques, see Paul C. Giannelli, *The Admissibility of Laboratory Reports in Criminal Trials: The Reliability of Scientific Proof,* 49 OHIO ST. L.J. 671, 688-92 (1988) (discussing existence of contrived and incomplete laboratory tests); Edward J. Imwinkelried, *The Debate in the DNA Cases over the Foundation for the Admission of Scientific Evidence: The Importance of Human Error as a Cause of Forensic Misanalysis,* 69 WASH. U. L.Q. 19, 25-27 (1988) (discussing high error rates for blood and urine analyses, electrophoresis, and handwriting analysis).

[33] *See* William C. Thompson, *DNA Evidence in the O.J. Simpson Trial,* 67 U. COLO. L. REV. 827, 827 (1996). Professor Thompson states:

> [T]he extensive use of duplicate testing in the Simpson case greatly reduced concerns . . . about the potential for false positives due to poor scientific practices of DNA laboratories. On the other hand, the Simpson case revealed serious problems regarding the collection and handling of biological samples and the potential for cross-contamination of evidence, that are of general concern.

Id.

inaccuracies that are the consequence of ineptness or dishonesty.[34] It is questionable whether performance errors — just because they occur in a laboratory — should be counted in calculating the error rate on which commentators propose conditioning the admissibility of DNA evidence. We regularly admit other kinds of evidence, such as business records, without any thought of quantifying the error that undoubtedly attends their creation. We trust evidentiary rules to minimize error and expose it when it occurs.

However calculated, the laboratory error rate differs radically from the error rates in baseball. The baseball analogy is misleading. In baseball, each time the infielder handles the ball we know from the official scorer whether or not an error occurred. The result is an error rate that is based on each chance the infielder has for a play.

There is no error rate based on the real tests a laboratory performs.[35] Instead, a surrogate error rate is calculated on the basis of open or blind proficiency testing done on specially constructed samples.[36] Statisticians know that a meaningful estimate of a particular laboratory's error rate on proficiency testing cannot be ascertained. Such a calculation would require more proficiency tests than any laboratory will ever perform.[37] And

---

[34] *See* Edward J. Imwinkelried, *The Case Against Evidentiary Admissibility Standards that Attempt to "Freeze" the State of a Scientific Technique,* 67 U. COLO. L. REV. 887, 896-97 (1996) (stating that blood analyses are often excluded because of faulty chains of custody).

[35] *See* Koehler, *Proving the Case, supra* note 18, at 871 ("The fact is that there is usually no independent way to determine the veracity of reported DNA matches in case work."); *see also* John I. Thornton, *Courts of Law v. Courts of Science: A Forensic Scientist's Reaction to* Daubert, 1 SHEPARD'S EXPERT & SCI. EVIDENCE 475, 481 (1994) (stating that no statistics exist on the rate of error of most tests conducted in forensic laboratories).

[36] *See* Richard Lempert, *After the DNA Wars: Skirmishing With NRC II,* 37 JURIMETRICS J. 439, 447-48 (1997) (explaining why it is virtually impossible to conduct truly blind DNA proficiency testing).

[37] *See* NRC2, *supra* note 3, at 86 (arguing that laboratories would have to conduct unrealistically large number of proficiency trials).

> Suppose that two laboratories each have under specific conditions a false-positive error rate of 0.10% — one match per 1,000 nonmatching proficiency trials. To establish that rate accurately, it would be necessary for each laboratory to undergo many thousands of trials. If one laboratory were to pass 1,000 proficiency tests without error, the 95% upper confidence limit for the error rate would be 0.30%. If the other laboratory had made one error, the limit would be 0.47%. Those results are not significantly different statistically. Both laboratories could have a true rate of 0.10%.

even were it possible for a particular laboratory to conduct enough blind sufficiency testing to be statistically meaningful, arriving at a representative error rate would still entail considerable guesswork due to the nature of DNA testing and proficiency testing.[38] Because an accurate estimate of a particular laboratory's rate is unattainable either by counting errors it made in real cases or by looking at its record on proficiency testing, it is, therefore, pointless to argue that *Daubert* demands that DNA evidence "should not be admitted unless a laboratory can offer a reliable estimate of *its* false positive error rate."[39]

Of course, Professor Koehler understands that the probability of error in a particular laboratory cannot be computed.[40] Instead, he suggests that proficiency test results should be pooled in order to estimate an "industry-wide" error rate, and that this pooled rate will be reported as the applicable laboratory error in a particular case.[41] The pooled proficiency test approach pretends that all laboratories are the same, despite the reality that some laboratories are better than others, and despite the fact that case-specific information will be available about how the laboratory performed in the case at hand.[42] It also assumes, without proof, that sufficient proficiency testing is done to arrive at a meaningful estimate of industry-wide laboratory error. The

---

Id. The Report adds that "only the largest forensic laboratories could have performed DNA testing in as many as 1,000 cases; no laboratory performs more proficiency tests than case tests." *Id.*

[38] *See* Lempert, *supra* note 36, at 448-49. In discussing a laboratory in which one false positive error occurred in identifying 50 samples, Lempert explains that this should not be scored as a 2% error rate because 1225 comparisons were done, yielding a rate of .0816%. He adds, however, that in the real world this adjusted rate might be unrealistic because the error was due to the cross-contamination of evidence and crime samples, an error that could have resulted in numerous false positives in other cases. Professor Lempert concludes this account of the difficulty of arriving at an error rate by noting that the laboratory changed its protocol for handling samples after the error was detected. *See also* William C. Thompson, *Accepting Lower Standards: The National Research Council's Second Report on Forensic DNA Evidence*, 37 JURIMETRICS J. 405, 419 (1997) (questioning how error rate should be calculated if test is designed to find only false matches and not false exclusions). But are false exclusions relevant when a defendant is attacking a match? Perhaps laboratory error rates that include false exclusions require adjustment.

[39] Scheck, *supra* note 15, at 1997 (emphasis added).

[40] *See* Koehler, *Proving the Case*, *supra* note 18, at 874-75.

[41] *See id.* at 875 ("industry-wide error rates should be used when estimating a DNA laboratory's match report error rate").

[42] See *infra* notes 77-86 and accompanying text.

objective of proficiency testing is not to measure error, but to allow laboratories to evaluate their performance, a somewhat different question. Furthermore, the amount of testing that is required (by voluntary standards or legislative directive) is not determined solely by how much performance review is appropriate. Especially in an era of budgetary constraints, cost as well as benefit affects the volume of testing. Proficiency testing is expensive and time-consuming; personnel who could be working on pending cases are instead spending their time designing tests, taking tests, and grading tests. Therefore, the data available for calculating industry-wide error rates are the product of pragmatic and evaluative concerns rather than a scientific approach to quantifying error.

This leap from a real case error rate to a surrogate proficiency test-based error rate to a pooled error rate to the rate for a particular laboratory may be the best that a statistician can do. But does this mean that the result is "scientific"? Does this calculation provide a "satisfactory approximation" of the error that occurred in the laboratory in question?[43]

I would argue that it does not; not only for the reasons stated above, but also because Professor Koehler's baseball analogy is misleading in another respect as well. It suggests that the fielder's lifetime error rate is a constant that will apply to each subsequent play. But this is a meaningless fiction to any fan knowledgeable about the infielder's current playing abilities. Similarly, the historical industry-wide error rate will not be sufficiently helpful in evaluating a laboratory's performance in a particular case at a particular moment in time. The historical pooled rate will include performance errors that no longer occur because of quality control and quality assurance programs, as well as interpretive errors experienced with now obsolete methodologies.[44] The rate may not take into account the fact

---

[43] *See* David H. Kaye, *DNA Evidence: Probability, Population Genetics, and the Courts,* 7 HARV. J.L. & TECH. 101, 128 (1993). In discussing a very different issue involving DNA evidence — the degree of population substructure — Professor Kaye observed that few assumptions in applied mathematics or statistics "hold rigorously. . . . Since almost all scientific work proceeds on the basis of simplifying assumptions, the question is whether the simplification produces satisfactory approximations." *Id.*

[44] For instance, as laboratories abandon VNTR-based testing systems in favor of other newer techniques such as PCR-based systems and STRs, certain errors that only pertained

that error rates in homicides and assaults differ markedly from those experienced with rapes, where the presence of the rape victim's DNA in the crime sample ensures that certain kinds of error cannot occur.[45] Furthermore, because proficiency test rates are derived from simulations, they are impervious to the possibility that in real cases other evidence, such as the existence of an iron-clad alibi, retesting, or laboratory review procedures may unearth a mistake, which the laboratory will correct.[46]

Although Professor Koehler concedes that the error rate ought to be adjusted when "[c]ase-specific circumstances . . . suggest that error is more or less likely to arise," he relies on another analogy to argue that it would be unscientific[47] to ignore pooled error rates just because specific facts may alter the probability of error in a given case.[48] He claims that the invariable existence of "unique circumstances . . . does not itself provide reason to disregard industry-wide error rate estimates any more than the fact of human uniqueness provides physicians with reason to disregard medical studies about likely responses to various drugs and procedures,"[49] or the Federal Aviation

---

to a VNTR-based system will no longer occur. For a discussion of new technologies, see NRC2, *supra* note 3, at 22-23.

[45] *See* Koehler et al., *supra* note 18, at 208 (questioning whether erroneous match call on proficiency test between sample A and one fraction of mixed sample B should be scored as error). Professors Koehler, Chia, and Lindsey state that "'fraction errors' are unlikely to falsely incriminate an innocent suspect because the suspect matches the other fraction of the mixed sample." *Id.*

[46] These criticisms should in no way be taken to suggest the futility of proficiency testing. As the NRC2 Report emphasizes, such testing is an important component of quality control. *See* NRC2, *supra* note 3, at 78-80.

[47] *See* Koehler et al., *supra* note 18, at 206 n.15 ("Some may reject industry-wide error rates as insufficiently relevant to the instant case. The scientist responds by noting that the industry-wide error rate provides a baseline from which an estimate of the probability of error in the instant case can be made.").

[48] *See id.* An unadjusted error rate would run afoul of the "fit" or relevancy prong of the *Daubert* opinion. *See infra* notes 96-102 and accompanying text.

[49] *See* Koehler et al., *supra* note 18, at 206 n.15; *see also* Koehler, *Proving the Case, supra* note 18, at 873.

> No one would argue that the unique testing features associated with a particular laboratory in a particular case should be ignored when these features are demonstrably related to a reduced error rate. But it is a[n extreme form of the base-rate] fallacy to believe that the fact that such individuating features may exist denies relevance to industry-wide error-rate estimates.

Administration to ignore errors that lead to airplane crashes.[50]
These analogies miss the point. The question under discussion is
not whether, and under what circumstances, laboratory error
may have probative value.[51] The question is whether the scien-
tific validity of pooled error rates is such that, under *Daubert,*
evidence of a DNA match is meaningless unless it is either com-
bined with, or accompanied by, this pooled proficiency test-
based error rate.

The RMP conveys accurate information to the fact-finder
about the distinctiveness of the DNA profile in question. No
scientific principle dictates that this significance is worthless
without a pooled surrogate inexact approximation of an error
rate.[52] Whether it would be desirable on policy grounds to pro-
vide the jury with some estimate of error is the subject of the
next section. Thus far I have concluded only that "science" does
not dictate the exclusion of DNA evidence unless a laboratory
error rate is made part of the prosecution's case. Consequently
NRC2 rightly declined to make recommendations about han-
dling error in judicial proceedings as beyond its mandate to
deal with scientific issues.

Id.

[50] *See* Koehler, *DNA Likelihood Ratios, supra* note 18, at 430.

[51] *See, e.g.,* Edward J. Imwinkelried, *Coming to Grips with Scientific Research in* Daubert's
*"Brave New World": The Courts' Need to Appreciate the Evidentiary Differences Between Validity and
Proficiency Studies,* 61 BROOK. L. REV. 1247, 1256 (1995) (concluding that proficiency studies
should be usable to attack credibility, and that character rules that might bar such use
should be revised); Thompson, *supra* note 38, at 419 ("the rate at which a laboratory cross-
contaminates samples in proficiency tests would be useful for the jury to know if the jury
must evaluate the likelihood such an error occurred in a particular instance").

[52] The court can clarify for the jury precisely what the RMP means by instructing the
jury that it "assumes that no sample mishandling or laboratory error occurred." *See* NRC2,
supra note 3, at 198 n.93 (suggesting jury instruction be designed to "minimize the
possibility of cognitive errors"); *see also* United States v. Shea, 957 F. Supp. 331, 345 (D.N.H.
1997). In addressing the admissibility of RMP, the court concluded:

> Although I acknowledge that a jury could become confused concerning the
> meaning and potential significance of a random match probability estimate, I
> am confident that the risk of confusion is acceptably small if the concept is
> properly explained. Moreover, because such an estimate can be extremely
> valuable in helping the jury appreciate the potential significance of a DNA
> profile match, it should not be excluded merely because the concept requires
> explanation.

Id. at 345.

II. SHOULD THE PROSECUTION BE BURDENED WITH A LABORATORY ERROR RATE?

In discussing whether a pooled error rate ought to be made part of the prosecution's case as an appropriate prudential mechanism, I first consider the principal arguments that have been advanced in favor of this solution. Then I turn to the evidentiary and adversarial problems that such a proposal would create. My conclusion is that the rationale for imposing a burden on the prosecution is weak, and that such a shift would be inconsistent with significant evidentiary policies. Even though there may be some uncertainty associated with DNA evidence because of the possibility of error — as is true with all evidence — a radical realignment of the prosecution's burden of proof is not warranted.

A. *The Policy Arguments in Favor of Burdening the Prosecution*

Two principal arguments are offered in favor of making a pooled error rate a condition of admissibility. First, the peculiar nature of DNA evidence — or scientific evidence in general — means that jurors must be informed about laboratory error in order to assure more accurate verdicts. Second, the mandatory divulgence of error is justifiable as a prophylactic mechanism that will force laboratories to adopt more rigorous procedures.

1. The Nature of the Evidence

Mandating an error rate as a condition of admissibility is sometimes justified on the ground that jurors are overwhelmed by the singular characteristics of DNA evidence. In particular, the argument is made that once jurors hear the low probabilities associated with a RMP, they will find it inconceivable that error could have occurred. The astronomical odds to which prosecution experts refer allegedly blind them to the possibility that something may have gone wrong in the laboratory. Indeed, some commentators suggest that informing jurors about the odds of a random match may make more of an impression than the absolute terms experts use when they testify about matching

fingerprints or bullets or explosives, even though such conclusions rest on statistical assumptions as well.

Although the belief that jurors will overrate probabilistic evidence has been expressed with some frequency,[53] this speculation has not been confirmed. Virtually no empirical studies have as yet been conducted on juror reactions to extremely low probabilities. The few studies to date with mock jurors cut both ways. While they indicate that the subjects did not give error rates their statistical due, they also suggest that jurors undervalue match probabilities.[54] In other words, jurors seem to be underwhelmed rather than overwhelmed by statistical information of any kind.[55] Their failure to respond to statistical evidence is also corroborated by recent research conducted by Professor Koehler that indicates that mock jurors' conviction rates are unaffected by different rates of laboratory error.[56] Unavailable as yet are data on how jurors react to judicial instructions on error or to explorations of error during the cross-examination of expert witnesses who testify to DNA matches. Re-

---

[53] *See, e.g., Ex parte* Perry, 586 So. 2d 242, 254 (Ala. 1991) (quoting State v. Schwartz, 447 N.W.2d 422, 428 (Minn. 1989), and finding that DNA population frequency evidence creates a "potentially exaggerated impact on the trier of fact"); People v. Harbold, 464 N.E.2d 734, 749 (Ill. App. Ct. 1984) ("We believe that testimony to statistical probabilities [with regard to ABO blood-typing] encouraged the jury to disregard evidential risks traditionally weighed in determining guilt or innocence, and focused unfairly upon a numerical conclusion."); Commonwealth v. Curnin, 565 N.E.2d 440, 441 (Mass. 1991) (noting "aura of infallibility" of RMP); State v. Carlson, 267 N.W.2d 170, 176 (Minn. 1978) ("Testimony expressing opinions or conclusions in terms of statistical probabilities can make the uncertain seem all but proven, and suggest, by quantification, satisfaction of the requirement that guilt be established 'beyond a reasonable doubt.'"); Brown v. State, 751 P.2d 1078, 1079 (Okla. Crim. App. 1988) (statistical probability evidence "may have an undue psychological impact on the trier of fact"); *see also* State v. Futch, 860 P.2d 264, 277-78 (Or. Ct. App. 1993) (Riggs, J., dissenting) (jury overreliance on statistics cannot be ignored in case admitting DNA evidence), *aff'd*, 924 P.2d 832 (Or. 1996). The seminal article on the pitfalls and dangers of statistical evidence is Lawrence Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971).

[54] *See* Koehler, *Proving the Case, supra* note 18, at 882 n.62 (stating that "data suggest that jurors may not be as impressed with DNA matching evidence as many believe. Fewer than half of the subjects [in the University of Texas experiment] voted to convict in any condition.").

[55] *See* NRC2, *supra* note 3, at 197 (analyzing several studies that suggest jurors generally undervalue weight of probability evidence).

[56] *See* Koehler et al., *supra* note 18, at 213 (reporting that, in University of Texas experiment, "introduction of laboratory error rates (.1%, 2%) did not significantly affect conviction rates").

search is also lacking on whether jurors find more convincing a probabilistic account — which acknowledges the possibility that someone else could be the source of the evidence — than opinions couched in absolute terms, such as testimony by an expert that a fingerprint at the crime scene was left by the defendant. This lack of information on how lay persons process statistical information led NRC2 to recommend the need for behavioral research.[57]

Couching DNA evidence in probabilistic terms, which was originally done with fingerprints as well,[58] may well turn out to be an interim measure. As more and more genetic loci become available for analysis, courts may eventually conclude that each person's DNA profile is unique.[59] Experts would then be permitted to opine that a match signifies that the samples being tested come from the same source. Although a shift to testimony in absolute rather than probabilistic terms would moot the argument that error rates are a needed correction to the overpowering odds to which jurors are subjected, other arguments for making error rates part of the prosecution's burden would survive.

A second possible justification for a special error rule for DNA evidence rests on the evidence's arcane nature. Unlike fingerprint evidence, which rests on a theory that is readily understandable and visible — it is easy to see that the whorls on one's fingers do not match those of others in the room — DNA alleles are invisible, and theories of population genetics are complex.[60] The argument here is that the mysterious nature of DNA profiling makes it appear infallible, and therefore not subject to error.

Some commentators would extend this rationale — that jurors need to know that error does occur even in connection with abstruse and seemingly irrefutable evidence — to all scientific evidence.[61] They would require the prosecution to make an

---

[57]	*See* NRC2, *supra* note 3, at 203-04.

[58]	*See id.* at 56-57.

[59]	*See id.* at 136-38 (discussing point at which increasing discovery of genetic markers might lead courts to conclude that genotype is unique).

[60]	*See* State v. Bogan, 905 P.2d 515, 525 (Ariz. Ct. App. 1995) (Weisberg, J., concurring) (drawing distinction between DNA evidence and physical comparisons, such as fingerprints, shoe tracks, and ballistic evidence, that jurors can see).

[61]	*See* Koehler et al., *supra* note 18, at 216 n.53 ("These arguments apply with equal

error rate part of its case-in-chief because they believe that scientific evidence is so compelling that jurors will ignore the possibility of laboratory error.

Singling out DNA or other scientific evidence for special treatment is inconsistent with the law's reaction to the possibility of error in other contexts. There are other kinds of evidence that jurors may consistently overvalue.[62] A considerable body of empirical research suggests that jurors may accord eyewitness testimony far more probative value than it deserves.[63] Until now, the data concerning error in eyewitness cases were derived from simulations that perhaps do not reflect what happens when there is a genuine crime.[64] If, however, we are wary of extrapolating error from staged events, we now have another source for error rates. Data are accumulating, and could be collected more systematically, on how often an eyewitness identification in an actual case is contradicted by a DNA test that exonerates the

---

force to other forensic science techniques including fingerprinting.").

[62] *See* Daniel W. Shuman et al., *Juror Assessments of the Believability of Expert Witnesses: A Literature Review*, 36 JURIMETRICS J. 371, 371 (1996). For instance, empirical research is just beginning to assess the extent to which jurors are affected by the occupation of the expert witness. *See id.* at 379.

A subsequent article by these same authors concluded that, although police officers were believed considerably more than medical specialists, the results were not statistically significant and suggest that jurors do not form impressions of experts based on their occupations. *See* Daniel W. Shuman et al., *Assessing the Believability of Expert Witnesses: Science in the Jurybox*, 37 JURIMETRICS J. 23, 26-31 (1996). *But see* Michael J. Saks & Roselle L. Wissler, *Legal and Psychological Bases of Expert Testimony: Surveys of the Law and of Jurors*, 2 BEHAV. SCI. & L. 435, 442 (1984) (finding that jurors' assessment of witness believability is affected by expert's occupation).

[63] *See* Steven Penrod & Brian Cutler, *Witness Confidence and Witness Accuracy: Assessing Their Forensic Relation*, 1 PSYCHOL. PUB. POL'Y & L. 817, 822 (1995). The authors conclude that:

> Taken together, the survey studies, the prediction studies, and the mock juror experiments converge on a worrisome set of conclusions about jurors' abilities. Jurors appear to overestimate the accuracy of identifications (there are more convictions than accurate identifications), do not distinguish accurate from inaccurate eyewitnesses, and are generally insensitive to factors that influence eyewitness identification accuracy.

Id.; see also Brian L. Cutler et al., *Juror Sensitivity to Eyewitness Identification Evidence*, 14 L. & HUM. BEHAV. 185, 190 (1990) (finding that actual jurors, as well as mock jurors, have difficulty assessing eyewitness testimony).

[64] *See* Gary L. Wells, *Applied Eyewitness-Testimony Research: System Variables and Estimator Variables*, 36 J. PERS. & SOC. PSYCHOL. 1546, 1551-55 (1978) (discussing factors that make accuracy rates of eyewitness research unrepresentative).

accused.[65] In addition, a growing number of rape convictions based on eyewitness testimony are being overturned after post-conviction DNA testing. Despite these real case data that confirm the research findings on mistaken identifications, the law rarely acknowledges that jurors ought to be informed that their confidence in the accuracy of eyewitness testimony may be misplaced. Even in a case that hinges exclusively on eyewitness identification, not only is the prosecution under no obligation to furnish error rates, but usually the defense is precluded from calling expert witnesses to educate jurors about the prevalence of eyewitness error.[66] In light of this unwillingness, even for impeachment, to make use of eyewitness error about which we know quite a bit, it seems rather inconsistent to take the radical step of conditioning the admissibility of DNA evidence on pooled proficiency test-based error rates, even though there is virtually no data on how jurors process RMPs or error rates.

2. Error as a Prophylactic Measure

Some commentators justify making the laboratory error rate a part of the prosecution's case as a way to force laboratories into adopting more rigorous procedures.[67] This rationale applies

---

[65] *See* EDWARD CONNORS ET AL., CONVICTED BY JURIES, EXONERATED BY SCIENCE: CASE STUDIES IN THE USE OF DNA EVIDENCE TO ESTABLISH INNOCENCE AFTER TRIAL 20 (1996) (finding that in 21,621 cases in which 19 public and private laboratories had performed DNA testing, the DNA test results excluded suspects in about 23% of the cases). The laboratories' data failed to provide details; it is, therefore, impossible to determine in how many cases there had been eyewitness identifications, but since many undoubtedly involved rape — though these data were not available either — in which eyewitness identifications are customary, the error rate appears to be considerably higher than the 1% to 4% error rate that Professor Koehler has endorsed for DNA testing. *See* Koehler, *DNA Matches, supra* note 18, at 229. Professor Koehler's estimate is on the high side. Other commentators' estimates, which he cites, range from 0.0008% to 2%. *See* Koehler et al., *supra* note 18, at 218 n.47.

[66] *See generally* Roger B. Handberg, *Expert Testimony on Eyewitness Identification: A New Pair of Glasses for the Jury*, 32 AM. CRIM. L. REV. 1013 (1995) (arguing that courts should admit eyewitness expert testimony more often); Steven Penrod et al., *Expert Psychological Testimony on Eyewitness Reliability Before and After* Daubert: *The State of the Law and the Science*, 13 BEHAV. SCI. & L. 229 (1995) (examining grounds upon which state and federal courts have excluded eyewitness expert testimony); Wayne T. Westling, *The Case for Expert Witness Assistance to the Jury in Eyewitness Identification Cases*, 71 OR. L. REV. 93 (1992) (arguing for more liberal admissibility of eyewitness expert testimony).

[67] *See* Thompson, *supra* note 33, at 844-46 (describing laboratory errors in context of O.J. Simpson case).

equally to laboratories engaged in other kinds of forensic analyses, although the outcry about error has occurred principally in the DNA context. Errors occur with other forensic techniques,[68] and the consequences may be as catastrophic in their adverse impact on a defendant's life as a false DNA match. In a narcotics prosecution, for instance, the nature of the drug, the purity of the drug, and the weight of the drug all have severe legal ramifications.[69]

But even if we accept that prodigious error occurs and agree that something must be done to ensure higher laboratory standards, it does not follow inevitably that the admissibility of forensic evidence must be conditioned on a laboratory error rate. Other measures, discussed below, could improve laboratory performance without colliding with the basic evidentiary principles discussed in the next section.

B. Burdening the Prosecution with an Error Rate Contravenes Evidentiary Principles

1.   Pooled Proficiency-Based Error Rates Prevent Jurors from Analyzing the Evidence

If we return one last time to the baseball hypothetical, we will see why it furnishes an inappropriate analogy for a trial. As Professor Koehler states, the infielder's far better error rate for throwing than fielding the ball is insignificant if we are sitting in the stadium placing bets on how likely he is to make the next

---

[68] *See* Paul C. Giannelli, *The Abuse of Scientific Evidence in Criminal Cases: The Need for Independent Crime Laboratories*, 4 VA. J. SOC. P. & L. 439 (1997) [hereinafter *Abuse of Scientific Evidence*] (discussing recent cases in which forensic scientists perjured themselves, faked laboratory reports, or both); Paul C. Giannelli, *The Admissibility of Laboratory Reports in Criminal Trials: The Reliability of Scientific Proof*, 49 OHIO ST. L.J. 671, 688-692 (1988) (discussing laboratory errors made in connection with variety of forensic techniques in pre-DNA era); Randolph N. Jonakait, *Real Science and Forensic Science*, 1 SHEPARD'S EXPERT & SCI. EVIDENCE Q. 435, 448 (1994) ("Error rates are generally unknown, but the available data indicate large error rates."); Andre A. Moenssens, *Novel Scientific Evidence in Criminal Cases: Some Words of Caution*, 84 J. CRIM. L. & CRIMINOLOGY 1, 1-2 (1993) (analogizing error rates in voiceprint identification and identification by individual bite marks).

[69] *See* U.S. SENTENCING GUIDELINES MANUAL § 2D1.1(c) (1995) (establishing base offense levels for different quantities of variety of controlled substances). The permissible length of sentence determined by the offense level is set out in Chapter 5, Part A (ranging from 0 to 293 months).

play correctly. But is it appropriate to evaluate the probative value of evidence by asking how much a juror would be willing to bet beforehand on the likelihood of what will occur?[70] Although gambling analogies are often used in analyzing probabilities, they do not capture how we wish jurors to function. The fallacy of the Koehler analogy becomes clear if we switch the fielder's error rates. Suppose a shortstop makes fielding errors once in a thousand attempts, but makes throwing errors two percent of the time. If the next ball is hit in the air, his bad throwing arm is relevant only if he drops the ball. Jurors should not be asked to assess the likelihood of laboratory error independently of whether an event occurred. Their function should be to determine what happened with regard to a completed play, not to consider in the abstract the probability of error in the next case in which DNA will be analyzed.[71]

This conclusion does not rest on a criticism of pooled proficiency rates as relying solely on probabilities. I would agree with those who argue that "[a]ll inferential processes are probabilistic."[72] Further, the academic debate on whether "naked statistical evidence" is admissible is also beside the point.[73]

---

[70] *See* Glen Shafer, *Diverse Models of Evidence and Inference: The Construction of Probability Arguments*, 66 B.U. L. REV. 799, 808 (1986) (explaining why the gambling metaphor is inappropriate for the courtroom).

[71] *Cf.* Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357, 1360-62 (1985) (arguing that perception of verdict as statement about event (rather that as statement about evidence) is more powerful). Professor Nesson argues that there is a difference between a statement by someone who has briefly glimpsed a playing card, and on the basis of that glimpse concludes, "the card flashed was not a picture card," and a statement, made before blindly drawing a card from a well-shuffled deck, that "the card is probably not a king." *See id.* at 1361. This latter statement, which he labels a statement about the evidence, "is based on our knowledge about the make-up of the deck, the fairness of the draw, and the laws of probability — not about the unseen card." *See id.* Professor Nesson further argues that verdicts will have a deterrent value if "the public views the verdict as a statement about a past event" and "not merely a statement about the evidence." *See id.* at 1360-61.

[72] *See* United States v. Chaidez, 919 F.2d 1193, 1200 (7th Cir. 1990); *see also* Ronald J. Allen, *On the Significance of Batting Averages and Strikeout Totals: A Clarification of the "Naked Statistical Evidence" Debate, the Meaning of "Evidence," and the Requirement of Proof Beyond a Reasonable Doubt*, 65 TUL. L. REV. 1093, 1098 (1991) (arguing that assumption that statistical and non-statistical evidence are distinct types of evidence is false).

[73] *See, e.g.*, Allen, *supra* note 72, at 1097-99 (naked statistical evidence does not exist); Craig R. Callen, *Adjudication and the Appearance of Statistical Evidence*, 65 TUL. L. REV. 457, 466-75 (1991) (statistics alone are meaningless unless they are placed in an empirical context); Daniel Shaviro, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV. L.

The probability of laboratory error in DNA cases is not one of the small category of issues in which "naked" — or almost naked — statistical evidence has proven essential.[74] Case-specific evidence will be available when DNA evidence is introduced to show that the likelihood of laboratory error is lower or higher than the betting odds on what laboratories might do wrong in the future. Consequently, a base error rate, derived from industry-wide proficiency testing, is not needed as part of the prosecution's case to ensure that relevant, otherwise unobtainable information, reaches the factfinder.[75] Burdening the prosecution with a pooled error rate would interfere with the jury drawing those inferences that the evidence permits, the function our legal system expects a jury to perform.[76]

2. The Adversary System Can Detect Case-Specific Error

Interpretative errors, which are a significant component of laboratory error, are not invisible when defendants are provided with adequate documentation of the laboratory's operation.[77] Mistakes in applying proper scientific principles can then be attacked with the assistance of existing evidentiary and procedural rules.[78] Counsel can cross-examine the prosecution's experts

---

REV. 530, 535-42 (1989) (discussing whether verdicts based entirely on statistical evidence create unjust punishment of innocent person); Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 B.U. L. REV. 439 (1986) (Bayes Theorem and other forms of statistical reasoning may be better suited for certain problems of proof than others).

[74] *See, e.g.*, Steve Gold, *Causation in Toxic Torts: Burdens of Proof, Standards of Persuasion, and Statistical Evidence*, 96 YALE L.J. 376 (1986) (describing epidemiological proof); RAMONA L. PAETZOLD, THE STATISTICS OF DISCRIMINATION (1994) (discussing use of statistical evidence in discrimination cases).

[75] *Cf.* United States v. Hannigan, 27 F.3d 890, 899 n.5 (3d Cir. 1994) (Becker, J., concurring) ("The advantage or disadvantage . . . of the statistical approach is that in proper circumstances it provides a more precise guide to the factfinder than would an inference based solely upon his or her experience, and in other circumstances provides a superior guide than mere inference based on the jury's collective knowledge *where no other evidence is available.*") (emphasis added).

[76] *See* Allen, *supra* note 72, at 1098 (*"The real question* . . . is how one draws inferences from evidence.") (emphasis added).

[77] *See* State v. Schwartz, 447 N.W.2d 422, 427 (Minn. 1989) ("[A]ccess to the data, methodology, and actual results is crucial so a defendant has at least an opportunity for independent expert review."); NRC1, *supra* note 5, at 150 ("All data and laboratory records generated by analysis of DNA samples should be made freely available to all parties.").

[78] *See* Imwinkelried, *supra* note 34, at 898 (analyzing how mistakes in applying scientific

about how they reached their conclusions, dispute their opinions with the help of the learned treatise rule,[79] and call experts of their own. Prosecution experts can also be cross-examined about the possibility of performance errors[80] or the procedures that the laboratory employs to minimize their occurrence. In addition, courts have the power to enlist the assistance of court-appointed experts.[81] Procedures such as duplicate testing by the prosecution,[82] retesting by the defense,[83] and testing all crime scene samples[84] eliminate the possibility of many laboratory performance errors,[85] which may also be detectable through scrutiny of the laboratory's internal documentation.[86]

---

principles can be attacked with assistance of evidentiary and procedural rules). For an account of how the defense attacked alleged errors in the collection and handling of DNA evidence in the O.J. Simpson case, see Thompson, *supra* note 33, at 831-40.

[79] *See* FED. R. EVID. 803(18).

[80] *See* Thompson, *supra* note 38, at 419 (concluding that proficiency tests can be useful in assessing risk of error by giving example of noticing, while reviewing data from particular laboratory, that the analysts sometimes attributed extra alleles to artifact rather than existence of second contributor to mixed DNA sample). Certainly, the expert can be cross-examined about the possibility of such an error and how the laboratory guards against it. It is not clear to what extent present law would permit prosecution experts to be cross-examined about statistics on laboratory error rather than the possibility of error. *See* Imwinkelried, *supra* note 51.

[81] *See* FED. R. EVID. 706.

[82] *See* Thompson, *supra* note 33, at 844 (stating that O.J. Simpson defense did not challenge any test results that second laboratory replicated; defense instead focused on potential for error before samples were split). Errors may also be detected when a laboratory uses more than one technology. *See* Koehler et al., *supra* note 18, at 218 n.27 (discussing how reversal of band sizes on RFLP analysis did not occur on PCR analysis).

[83] *See* NRC2, *supra* note 3, at 88 ("The best protection that an innocent suspect has against an error that could lead to a false conviction is the opportunity for an independent retest."); *see also* Recommendation 3.3, *in id.*

[84] *See, e.g.*, State v. Hammond, 604 A.2d 793, 798-808 (Conn. 1992) (stating that jury had convicted defendant despite DNA test on semen stains found on victim's clothing that excluded him; court remanded for failure of prosecution to submit vaginal swabs taken from rape victim for DNA testing).

[85] Some laboratory performance errors may remain undetectable if the mix-ups or contaminations occurred before the samples were split. But as was said before, these errors are not unique to DNA evidence. The police deliberately or negligently may have lifted the defendant's fingerprints at one location and mistakenly have labeled them as coming from the murder scene. Important business records may have been accidently or intentionally destroyed.

[86] *See* Thompson, *supra* note 33, at 845 (commenting on case in which laboratory's switch of victim's and defendant's reference samples was discovered when defense expert noted discrepancies in laboratory's chain-of-custody documents).

The call by some commentators for a special error rule for DNA cases may reflect their fear that defendants will not obtain adequate discovery, competent defense counsel, or the assistance of qualified experts.[87] But this is hardly a concern limited to cases in which DNA evidence is introduced. We have but to look at death penalty cases having nothing to do with DNA evidence to see that the direst outcome possible may result from defense counsel's lack of resources, bad lawyering, or both. The cases reveal abundant failures in investigating adequately and in obtaining needed expert assistance, as well as ineptitude in presenting evidence, cross-examining effectively, and making appropriate evidentiary and constitutional objections.[88] Before we carve out a special prophylactic rule for DNA cases that undercuts the jury's role in assessing evidence, we should look at whether something specific to DNA evidence restrains defense counsel from raising questions about laboratory error, or whether counsel desist for the same reasons they fail to make arguments in other cases.

If the adversary system does not adequately protect poor criminal defendants, the cure is not a special admissibility rule for DNA evidence. This would distract from the need for reform in all cases. A special rule is inconsistent with basic evidentiary principles, and sidesteps the procedural safeguards that our system should provide. Liberal discovery, expert assistance for indigent defendants, the greater utilization of court-appointed experts, and a higher standard for the effective assistance of counsel would make many trials more fair.[89]

---

[87] *Cf. id.* at 846 (stating that defense needs resources to adequately question DNA evidence).

> If subjective interpretation is an issue going merely to weight, then questionable subjective calls are likely to be detected only in cases where the defense has the resources to hire an expert and gain access to the underlying laboratory work through discovery. Although this sometimes happens, the majority of criminal cases involving DNA evidence are resolved without any independent check on the work of the forensic laboratory.

Id. Professor Thompson calls for more objective procedures and duplicate testing, when possible, as a condition of admissibility of DNA evidence. *See id.*

[88] *See generally* Stephen Bright, *Counsel for the Poor: The Death Sentence Not for the Worst Crime but for the Worst Lawyer,* 103 YALE L.J. 1835 (1994) (describing how inadequate representation leads to harsher sentences against indigents).

[89] *See* Berger, *supra* note 13, at 1352-63 (discussing problems with scientific evidence in

Requiring an error rate as part of the prosecution's case for fear the jury will not properly assess the evidence against the defendant is reminiscent of a now generally outmoded common law approach to evidence. Burdens, such as the requirement of corroboration, were placed on the prosecution to protect the innocent instead of trusting jurors to weigh the evidence in the case.[90] We have abandoned this distrust of jurors in recent years, and should not now readopt fixed evaluation of evidence requirements without some demonstration that defendants are at greater risk in DNA cases than in criminal cases in general. The demand for incorporating an error lower bound into presentations of DNA evidence obscures the fact that, in the great majority of cases, no laboratory error will have occurred.[91]

3. *Daubert* and the Calculation of Error

If, despite the arguments above, pooled proficiency test error rates are deemed useful to estimate the likelihood that a particular laboratory erred in the case at hand, the error rate would have to be adjusted by case-specific factors. The Supreme Court set out a framework for the admissibility of scientific evidence in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*[92] It explained that in addition to passing on the scientific validity of proffered expert testimony, the trial court must also determine whether the expert's testimony "fits" the facts of the case.[93] The Court endorsed Judge Becker's opinion in *United States v. Downing*,[94] as "aptly describ[ing]" the consideration of "fit," and quoted Judge

---

criminal trials and suggesting solutions).

[90] *See* JACK B. WEINSTEIN ET AL., CASES AND MATERIALS ON EVIDENCE 1182-84 (7th ed. 1982) (indicating areas of common law where prosecution is required to corroborate various types of evidence).

[91] *See* Koehler et al., *supra* note 18.

[92] 509 U.S. 579 (1993). The "scientific validity" prong of *Daubert* is discussed *supra* at notes 15-52 and accompanying text.

[93] The Court explains that "'[e]xpert testimony which does not relate to any issue in the case is not relevant and, ergo, non-helpful.'" *Daubert*, 509 U.S. at 591 (quoting JACK WEINSTEIN & MARGARET BERGER, 3 WEINSTEIN'S EVIDENCE ¶ 702[02], at 702-18). Because fit is about relevancy, not about scientific validity, the other prong of the *Daubert* test, *see supra* note 13, the analysis that follows should apply as well in jurisdictions that still employ the "general acceptance" test articulated in *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923), for determining scientific validity.

[94] 753 F.2d 1224 (3d Cir. 1985).

Becker's language: "[a]n additional consideration under Rule 702 — and another aspect of relevancy — is whether expert testimony proffered in the case is sufficiently tied to the facts of the case that it will aid the jury in resolving a factual dispute."[95]

In *Downing*, the court was reviewing a ruling that precluded the defense from calling an expert to testify to the unreliability of eyewitness identifications. In remanding, Judge Becker's opinion did not rule out the possibility that such testimony might be helpful, but it cautioned that:

> a defendant who seeks the admission of expert testimony must make an on-the-record detailed proffer to the court, including an explanation of precisely how the expert's testimony is relevant to the eyewitness identifications under consideration. . . . Failure to make such a detailed proffer is sufficient grounds to exclude the expert's testimony.[96]

For instance, "fit" is lacking if an expert on eyewitness identifications seeks to testify about the "weapon focus phenomenon" — that witness identifications are less accurate when the witness is facing a weapon — in a case in which no weapons were displayed,[97] or about the increased inaccuracies in cross-racial identifications relative to same-race identifications when the eyewitnesses and the defendant come from the same racial group.[98]

Daubert's insistence on "fit" means that unadjusted laboratory error rates cannot be admitted. If expert testimony on the possibility of eyewitness error must be geared to the facts of the case when offered to impeach a prosecution witness, then *a fortiori*, fit is essential when an error rate is a condition of admissibility. The laboratory error rate would have to be compatible with what actually occurred in the case. For instance, certain errors calculated in connection with VNTR analyses cannot arise with PCR- or STR-based systems.[99] The laboratory's sample-handling

---

[95] *Daubert*, 509 U.S. at 591.

[96] *Downing*, 753 F.2d at 1242.

[97] *See, e.g.*, United States v. Dowling, 855 F.2d 114, 119 (3d Cir. 1988), *cert. granted*, 489 U.S. 1051 (1989), *aff'd*, 493 U.S. 342 (1990).

[98] *See, e.g.*, United States v. Norwood, 939 F. Supp. 1132, 1137 (D.N.J. 1996) (discussing showing defense must make to establish that proffered expert testimony fits facts of case).

[99] *See supra* note 44 and accompanying text (stating that if laboratories abandoned

practices,[100] or its use of controls[101] or techniques such as duplicate testing[102] in the case at hand will have eliminated the possibility of certain errors. Most importantly, DNA evidence will often indicate on its face that certain kinds of mistakes that were scored as errors on proficiency tests, and thus would enter into an industry-wide calculation, did not occur in this case.[103]

In addition to the difficulties these case-specific technical details pose in quantifying a laboratory's error rate, one can easily imagine other problems. Must the defense inform the prosecution that it retested so that certain kinds of errors can be subtracted from the rate?[104] Should the error rate be adjusted if the defendant was offered the opportunity to retest but failed to do so? Will this depend on how likely it was that the state would have paid for the retesting? Will the defense need additional discovery in order to deal with the threshold error rate issue?

Assuming that the problems in quantifying adjusted laboratory error could be overcome, a very questionable supposition, is it worthwhile to burden the prosecution with this issue? The consequence would be that, at least potentially, a laboratory error rate would have to be calculated for each case in which DNA evidence is offered.[105] Such a case-specific estimate threshold to the admissibility of DNA evidence might engulf courts in costly and protracted in limine hearings if the defense disagreed with

---

VNTR-based testing, certain errors would not occur).

[100] *See, e.g.,* NRC2, *supra* note 3, at 82 ("[I]n the loading of an electrophoresis gel, a sample loaded in one lane might leak into an adjacent lane, which might then appear to contain a mixed sample. Confusion resulting from lane-leakage problems is typically avoided by leaving alternate lanes empty or by putting critical samples in nonadjacent lanes.").

[101] *See supra* notes 82-86 and accompanying text (arguing that certain procedures help avoid laboratory errors).

[102] *See supra* note 82 and accompanying text (describing duplicate testing in O.J. Simpson case).

[103] *See* NRC2, *supra* note 3, at 80-85 (discussing various measures that laboratories can use to eliminate various kinds of mistakes).

[104] *Cf.* United States v. Lowe, 954 F. Supp. 401, 403 n.3 (D.Mass. 1997) (acknowledging that samples had been split, but "Lowe has declined to inform the court of the results, if any, of his independent analysis").

[105] Of course, defense counsel who now do not challenge DNA evidence effectively or at all would be unlikely to act markedly differently with respect to the prosecution's estimate of the applicable error rate. *See supra* note 66 and accompanying text. It is difficult to assess, therefore, the degree to which there would be any impact if the prosecution's choice of a statistic were not challenged.

the prosecution's calculation. Furthermore, if the available research is correct in concluding that laboratory error rates by themselves are ineffectual,[106] then the time and effort devoted to establishing an acceptable case-specific estimate would be non-productive unless, perhaps, it were combined with the RMP. A combined statistic, however, so underestimates the probative value of a DNA match in the great majority of cases that, as a matter of policy, we ought not deprive ourselves of this evidence — at least until we have a more sound basis for concluding that more mistaken convictions occur because of inaccurate DNA results than occur with all the other forensic techniques for which no error rate is required.

4. Misleading the Jury

The Supreme Court acknowledged in *Daubert* that courts may exclude scientific testimony under Rule 403 of the Federal Rules of Evidence.[107] The Court then quoted a statement by Judge Weinstein: "Expert evidence can be both powerful and quite misleading because of the difficulty in evaluating it. Because of this risk, the judge in weighing possible prejudice against probative force under Rule 403 of the present rules exercises more control over experts than over lay witnesses."[108] Rule 403 needs to be kept in mind because the addition of a mandatory error rate might make expert testimony about DNA evidence more misleading.

Two different fallacies need to be avoided in presenting statistical evidence. The "prosecutor's fallacy" occurs when a prosecutor suggests that the evidence indicates the likelihood of the defendant's guilt, rather than the odds of a match, given that defendant is not the source.[109] Furthermore, even in the ab-

---

[106] *See supra* note 56 and accompanying text.

[107] *See* Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 595 (1993). In *Daubert*, the Court stated that Rule 403 provides for the exclusion of evidence "if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury." *See id.* (quoting Jack B. Weinstein, *Rule 702 of the Federal Rules of Evidence is Sound; It Should Not be Amended,* 138 F.R.D. 631, 632 (1991)).

[108] *See id.*

[109] *See* NRC2, *supra* note 3, at 31-32, 133, 198. *See generally* William Thompson & Edward Schuman, *Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the Defense Attorney's Fallacy,* 11 L. & HUM. BEHAV. 167 (1987) (comparing prosecutor's fallacy

sence of prosecution insinuation, the danger persists that jurors will transpose the conditional probability, despite the court's or experts' instructions explaining the RMP.[110]

If the RMP estimate were limited by the laboratory error rate, however, explanations by the court and experts of the meaning of the single adjusted statistic would reveal that this is a RMP estimate modified to account for the probability of laboratory error in this case.[111] Is it not perhaps possible that a reference to error *in this case* would make jurors even more prone to assume that the adjusted figure must represent the odds of guilt *in this case?*[112] Even if the probability were much higher than a pure random match probability, a juror who thinks that the statistic means that there is only a one in five thousand chance that defendant is not guilty may conclude that the reasonable doubt standard has been satisfied.

The "defendant's fallacy" occurs when defense counsel suggests, or jurors conclude, that anyone with the same profile as the defendant is as likely as the defendant to be the source of the crime scene sample.[113] In other words, if the error-adjusted probability is one in five thousand and the crime occurred in a city of five million, jurors might erroneously conclude that one thousand other equally likely suspects exist. This fallacy, which

---

with defense attorney's fallacy).

[110] *But see* United States v. Shea, 957 F. Supp. 331, 344-45 (D.N.H. 1997).

> I am unconvinced by Shea's claim that a jury cannot properly assess the potential of a false match unless a false match error rate is calculated and combined with the random match probability estimate. . . . In a real trial setting, the parties are given an opportunity to explain the significance of statistical evidence through expert testimony. Further, if a trial judge concludes that jurors could be confused by statistical evidence, the judge can deliver carefully crafted instructions to insure that the evidence is properly understood.

Id.

[111] *See supra* notes 92-106 and accompanying text (discussing why error rates must be case specific).

[112] *Cf.* Koehler, *Proving the Case, supra* note 18, at 876-83 (discussing how likelihood ratios could give jury false sense that DNA evidence answers question directly). Professor Koehler conducted a small experiment on the impact of likelihood ratios versus frequency statistics. He found that despite the fact that likelihood ratios have scientific merit and are the choice of some commentators who believe they make better use of DNA data, jurors do not understand them and get confused. *See id.* at 878; *see also* NRC2, *supra* note 3, at 200.

[113] *See* Thompson & Schuman, *supra* note 109, at 171 (defining defense attorney's fallacy as when people conclude that associative evidence is irrelevant). ·

current research indicates is more likely to sway jurors than the "prosecutor's fallacy,"[114] results from the jury's failure to evaluate the non-DNA evidence that tends to incriminate the defendant, but not a random member of the population. Larger probabilities such as would result with an error-adjusted statistic are likely to fuel this type of erroneous juror reasoning — a probability of one in one thousand might lead some to conclude that there are five thousand potential culprits in our hypothetical city.

5. Jury Nullification

A final question that should be considered is whether making laboratory error an issue for the prosecution in every DNA case would cause more jurors to acquit persons whom they believe to be guilty. By raising this question I do not wish to get embroiled in the heated debate about when and whether jurors are ever justified in refusing to convict.[115] I suggest only that even if we assume that jury nullification may be appropriate under certain conditions, that does not mean that jurors should exercise this power without regard to the law, the evidence, or the circumstances of the case before them. There is a difference between a refusal to convict because of qualms about the applicable law or doubts about the prosecution's behavior,[116] and an unwillingness to convict regardless of the situation. Even Professor Butler, who favors broad jury nullification as a response to racism and states as his goal "the subversion of American criminal justice,"[117] draws the line at jury nullification in crimes of violence like murder, rape, and assault. In those kinds of cases, he

---

[114] *See* NRC2, *supra* note 3, at 198.

[115] *See, e.g.*, Robert F. Schopp, *Verdicts of Conscience: Nullification and Necessity as Jury Responses to Crimes of Conscience*, 69 S. CAL. L. REV. 2039 (1996) (examining normative basis for jury nullification and its relationship to necessity defense); Steven M. Warshawsky, Note, *Opposing Jury Nullification: Law, Policy, and Prosecutorial Strategy*, 85 GEO. L.J. 191 (1996) (arguing that jury nullification threatens to undermine American legal system); Jack B. Weinstein, *Considering Jury "Nullification": When May and Should a Jury Reject the Law to do Justice?*, 30 AM. CRIM. L. REV. 239 (1993) (arguing that juries really disregard law, but that jury nullification is appropriate in some instances to promote justice).

[116] *See* Paul Butler, *Racially Based Nullification: Black Power in the Criminal Justice System*, 105 YALE L.J. 677 (1995).

[117] *See id.* at 680.

agrees that "jurors should consider the case strictly on the evidence presented."[118] These, of course, are precisely the kinds of cases in which DNA evidence is introduced.

We have but to shift the discussion from laboratory error to police error and from DNA profiles to fingerprinting to understand the possible consequences. We have had a number of extremely disturbing incidents in the last few years, in which police in Chicago, Philadelphia, and New York were found to have planted fingerprint evidence. The result is the exclusion of evidence and the reversal of convictions when there is evidence that such misconduct occurred. We certainly do not, however, require prosecutors to inform jurors about the possibility of police error, or make them advise jurors in surrounding police precincts or communities about a possible error rate. Before we embrace the notion that scientific evidence must be accompanied by an error rate, we should consider the impact on our criminal law system.

It may not be in the best interests of society to convey the message that the ever-present possibility of laboratory error means that there is always reasonable doubt.[119] The danger is that some jurors might convince themselves that they are following instructions to deliberate even if they acquit in the face of overwhelming evidence of guilt.

CONCLUSION

DNA evidence is extremely powerful and we rightly fear that innocent persons might be convicted if DNA matches result from laboratory error. We also know that it is no more possible to eliminate all sources of error in connection with DNA testing

---

[118] *Id.* at 715.

[119] *Cf.* State v. Ragland, 519 A.2d 1361 (N.J. 1986). The defendant objected when the court instructed the jury that if it finds a, b, and c, "you *must* find defendant guilty," claiming that instruction deprived him of a fair trial. The court declined to change its instruction to "may," a change it viewed as "mak[ing] it more likely that juries will nullify the law, more likely, in other words, that no matter how overwhelming the proof of guilt, no matter how convinced the jury is beyond any reasonable doubt of defendant's guilt, despite the law, it will acquit." *Id.* at 136. The court explained: "Jury nullification is an unfortunate but unavoidable power. It should not be advertised, and to the extent constitutionally permitted, it should be limited. Efforts to protect and expand it are inconsistent with the real values of our system of criminal justice." *Id.*

than with any other kind of evidence, including fingerprints. Before we overreact by restricting the probative value of DNA evidence in every case, we need to take two steps.

First, we need to decrease the possibility of error in laboratories as much as possible. As discussed above, a principal argument for conditioning the admissibility of DNA evidence on a laboratory error rate is that such a rule will have a prophylactic effect in minimizing error. There are more direct ways of achieving the same result, however, without skewing evidentiary principles. Forensic laboratories could be made independent of police control.[120] Courts could pressure laboratories into seeking accreditation by refusing to qualify experts from unaccredited laboratories.[121] Courts could also insist, as a minority of courts do, that DNA evidence will not be admitted if the laboratory involved failed to follow proper protocols.[122] Courts could instruct jurors about laboratory quality control and assurance programs to permit them to draw a negative inference about the quality of the DNA evidence if such procedures had not been followed.[123]

Second, we need to make sure that our traditional mechanisms are functioning effectively. Defense counsel must be given adequate procedural tools, such as discovery and access to experts, in order to deal with DNA evidence. We also need more information about how to make DNA evidence — and other statistical data — more comprehensible to jurors.

We should take these steps before we engage in an expensive and time-consuming quest for an elusive statistic about error that is unlikely to improve our fact-finding system. The very

---

[120] *See* Giannelli, *Abuse of Scientific Evidence, supra* note 68, at 469-73 (urging that control be transferred to Medical Examiners' (ME) offices and that the ME offices should be shielded from political pressure).

[121] *See* NRC2, *supra* note 3, at 4 ("Recommendation 3.1. Laboratories should adhere to high quality standards (such as those defined by TWGDAM and the DNA Advisory Board) and make every effort to be accredited for DNA work (by such organizations as ASCLD-LAB)."). For a discussion of accreditation programs, see *id.* at 77-78.

[122] *See* Imwinkelried, *supra* note 32, at 29-31, 33-34 (indicating potential for error in scientific tests of evidence, and suggesting that admissibility of DNA evidence should be conditioned on use of proper test protocol, but placing burden of raising issue on opponent); *see also* People v. Castro, 545 N.Y.S.2d 985, 987 (Sup. Ct. 1989) (laboratory must conform to generally accepted protocol in order for DNA evidence to be admissible under *Frye*).

[123] *See* NRC2, *supra* note 3, at 180 n.42.

limited available research suggests that error rates by themselves will not have any impact. A combined error and RMP statistic may confuse jurors further or perhaps make them more cynical about the operation of the judicial system. Given these choices, neither science nor policy requires a special burden-shifting rule for DNA cases to handle the possibility of laboratory error.