1991

Forensic Science: The Need for Regulation

Randolph N. Jonakait
New York Law School, farrah.nagrampa@nyls.edu

Recommended Citation
Harvard Journal of Law & Technology , Vol. 4, Issue 1 (Spring 1991), pp. 109-192

Volume 4, Spring Issue, 1991

FORENSIC SCIENCE: THE NEED FOR REGULATION

Randolph N. Jonakait *

INTRODUCTION

Justice and liberty often depend on the reliability of forensic laboratories. Analytical errors can mean freedom for the guilty and incarceration for the innocent. As this Article demonstrates, however, poor quality pervades forensic science. It is time to stop the miscarriages of justice which result. Forensic laboratories can improve, and a mandatory regulatory system will help bring needed improvement.

Section I of this Article surveys information revealing the widespread, poor quality of work done in forensic science laboratories. Sections II and III explore the underlying reasons for the current inadequacies. Section IV describes a regulatory system that has produced better quality in clinical labs and shows how similar regulations would improve forensic science laboratories. Section V discusses differences between clinical and forensic facilities that might make the imposition of the clinical laboratory regulation model difficult. An alternative regulatory system that only requires proficiency testing is proposed as an initial step towards the full regulatory model.

I. FORENSIC LABORATORY PERFORMANCE

Since forensic labs have never allowed a detailed look at the caliber of their work, only fragmentary information is available. Those fragments, however, reveal a consistent pattern of unacceptable errors and inaccuracies.

A. LEAA Proficiency Testing

A 1978 study funded by the Law Enforcement Assistance Administration ("LEAA") provides the most detailed picture available of the quality of our crime laboratories.[1] Although some sporadic proficiency

---

* Professor of Law, New York Law School; A.B. 1967, Princeton University; J.D. 1970, University of Chicago Law School; LL.M. 1971, New York University Law School.

1. PETERSON, FABRICANT & FIELD, CRIME LABORATORY PROFICIENCY TESTING RESEARCH PROGRAM 26 (1978) [hereinafter LEAA]. A test of this magnitude has never been repeated.

testing of crime labs had been done previously, this three-year investigation was the first broad, nationwide proficiency examination. Known samples in a wide range of forensic specialties were sent to labs for analyses.[2] Participation in the testing was high[3] since laboratories were guaranteed anonymity.[4]

A disturbing pattern emerged. A surprisingly large number of reports contained erroneous results.[5] Indeed, only one quarter of the participating labs provided entirely acceptable responses in all cases.[6] Two-thirds

2. This approach typified the entire proficiency program. "To test the accuracy of these crime laboratories, the Project Advisory Committee prepared samples and submitted them to the laboratories for evaluation. The Committee knew the data that a proper forensic analysis of the samples would generate, and the Committee judged the laboratories' reports against the known data." Imwinkelried, *The Constitutionality of Introducing Evaluative Laboratory Reports Against Criminal Defendants*, 30 HASTINGS L.J. 621, 636 (1979).

3. "Most crime laboratories in the country elected to participate. . . . Participation for the initial three test samples, for example, was 90%, 78% and 81% . . . ." LEAA, *supra* note 1, at 1. "Overall, 219 laboratories in the United States and Canada participated with an average of 188 laboratories for each of the 23 tests." Lucas, Leete & Field, *An American Proficiency Testing Program*, 27 FORENSIC SCI. INT'L 71, 73 (1985).

4. "Laboratory directors were initially somewhat reluctant to participate based upon concerns about their independence, that standardized methods might be required, and particularly about confidentiality and anonymity of results." *Id. Cf.* LEAA, *supra* note 1, at 1. "Because this constituted a pilot study utilizing untried manufacturing and sample distribution techniques, untested questionnaires and completely new methods for analyzing responses from the crime laboratories, the Law Enforcement Assistance Administration and the Forensic Sciences Foundation assured all participating laboratories that they would remain anonymous and that all research and statistical data would be considered confidential." *Id.* Of course, these reasons would not support a continuing need for anonymity in proficiency testing once the pilot study was completed, but interestingly, the report's findings, which recommend continued proficiency testing, conclude, "Laboratory anonymity and the confidentiality of the submitted data are the key factors to insure a high participation rate in a voluntary program such as this one." *Id.* at 257.

5. A. MOENSSENS, F. INBAU & J. STARRS, SCIENTIFIC EVIDENCE IN CRIMINAL CASES 5–6 n.10 (3d ed. 1986). *See also* Imwinkelried, *supra* note 2, at 623: "[The] report revealed a startlingly high incidence of errors in analyses performed by police laboratories."

The study itself concluded that a "wide range of proficiency levels among the participating laboratories exists, and in general, there are several evidence types with which the laboratories are having serious difficulties." LEAA, *supra* note 1, at 257. An author of that report later stated, "I acknowledge that a disturbingly high percentage of laboratories are not performing routine tests competently, as shown by our proficiency testing." *Symposium on Science and the Rules of Legal Procedure*, 101 F.R.D. 599, 645 (1983) [hereinafter *Symposium*] (remarks of Professor Joseph Peterson).

6. LEAA, *supra* note 1, at 255 (Table 91), summarizes:

Percentage of Total Responses Considered Acceptable	Percentage of All Participating Labs With This Rating
100%	25.3%
90.0–99.9%	8.6%
80.0–89.9%	31.8%
70.0–79.9%	19.3%

of the labs rendered unacceptable responses for at least 10% of the analyses.[7]

The problems, moreover, were not confined to any particular kind of test or evidence. Thus, the percentage of unsuitable conclusions reached 71% in one blood test, 51% in a paint test, and 67% in a hair test.[8] An

---

60.0—69.9%	9.4%
50.0–59.9%	3.0%
Below 50%	2.6%

7. Unacceptable responses included not only wrong outcomes, but also correct answers given for the wrong reasons, unnecessarily equivocal results, and inconclusive results unsupported by the analytical work or based on improper or inedquate methods. *Id.* at 239. The report noted that "there is nothing inherent in an inconsistent opinion that demonstrates proficiency." *Id.* at 238.

All such results, as the report recognized, could pose dangers to criminal justice. For example, an inconclusive response could seem inculpatory although it should have provided exculpatory evidence. The report gave many such illustrations. A particular paint comparison provides an example. In that test, a fifth of the labs reported that two different paint samples could have shared a common origin. The results of their analyses were not wrong, but if the labs had done all the testing they should have, they would have concluded that the paints did not come from the same source. Thus, instead of presenting what should have been exculpatory information, these facilities would have presented an inculpatory report. This error, moreover, was inexcusable:

> The Committee does not condemn in any way the reporting of inconclusive results, where appropriate. Situations in which such a response would be appropriate might include an inadequate amount of evidence, a contaminated sample, or where the sample possesses few inherent characterizing features. This is not the case in this test sample. The state of the art in criminalistics is certainly advanced to the point that these samples of paint should be easily distinguished by techniques available to any laboratory attempting to conduct paint examinations. The Project Advisory Committee believes that an inconclusive report in this sample is not supportable.

Id. at 203. *See also* Starrs, *In the Land of Agog: An Allegory for the Expert Witness*, 30 J. FORENSIC SCI. 289, 303–06 (1985) (discussion of imprecise forensic testing contributing to the conviction of an innocent man).

8. The summary of the unacceptable responses indicated:

Drugs		7.8%
Firearms		28.2%
Blood		3.8%
Glass		4.8%
Paint		20.5%
Drugs		1.7%
Firearms		5.3%
Blood		71.2%
Glass		31.3%
Paint		51.4%
Soil		35.5%
Fibers		1.7%
Physiological	(A)	2.3%
Fluids	(B)	1.6%
Arson		28.8%
Drugs		18.2%

analysis of some representative errors indicates the dangers posed to criminal justice.

The validity of firearms comparisons is well accepted, and laboratories should, therefore, provide conclusive and reliable results.[9] The straightforward firearms test revealed otherwise. Each lab was sent three .25 caliber bullets that had been fired from two distinct guns.

> [Of the laboratories, 5.7%] misidentified one projectile, incorrectly reporting that all three projectiles had been fired through a single weapon.... 3.4% of all laboratories responding, incorrectly reported that none of the three projectiles could have been fired through the same weapon. [Therefore,] 9.1% of the total responding, reported results that are clearly in error.[10]

Such errors have disastrous implications for justice, since a criminal prosecution may hinge on a single bullet comparison.[11] Since firearms

---

Paint		34.0%
Metal		22.1%
Hair	(A)	50.0%
	(B)	27.8%
	(C)	54.4%
	(D)	67.8%
	(E)	35.6%
Wood		21.5%
Questioned	(A)	5.4%
Documents	(B)	18.9%
Firearms		13.6%

LEAA, *supra* note 1, at 251 (Table 89). Several types of tests were performed in more than one version with different levels of complexity. For example, two blood tests were performed, one simple (3.8% unacceptable answers) and one difficult (71.2% unacceptable answers). *See infra* notes 14–16 and accompanying text.

9. Professor Giannelli, who contends that the prosecution should have to establish the validity of a scientific technique beyond a reasonable doubt before it is admitted, concludes that firearm comparisons would satisfy this burden. Giannelli, *The Admissibility of Novel Scientific Evidence:* Frye v. United States, *a Half-Century Later*, 80 COLUM. L. REV. 1197, 1248 (1980).

10. LEAA, *supra* note 1, at 238. In addition, 4.6% of the labs reported inconclusive results. *Id.*

This test was not an aberration. In a previous test done in the study, each lab was sent three projectiles and two cartridge cases and asked if the samples had been fired from the same weapon: 3.8% of the laboratories "misidentified a projectile, reporting that one of the projectiles actually fired through [one of the guns] had been fired through the other weapon .... [3.8%] misidentified a cartridge case, reporting that one of the cartridge cases actually fired through [one weapon] had been fired in the other weapon ...." *Id.* at 207.

11. LEAA, *supra* note 1, at 238. *See also* Bradford, *Barriers to Quality Achievement in Crime Laboratory Operations*, 25 J. FORENSIC SCI. 902, 904 (1980). Bradford asserts, "This kind of error has been documented in several cases at trial in the United States in recent years through the discovery or reexamination process." *Id.*

analyses are one of the most frequent forensic procedures,[12] a one in eleven error rate has a widespread impact.

Soil comparison may not always be as crucial as a firearm examination, but its proficiency test produced an even higher error rate. Three samples were sent, two of which were duplicates. The labs were asked whether either of the two could have shared a common origin with the third specimen, which in fact differed from the other two. While about two-thirds of the labs provided correct analyses, over thirty percent were incorrect.[13]

In one physiological fluids test, participating labs were asked whether two bloodstains could have shared a common origin.[14] Only 52 out of 132 responding facilities correctly reported that the stains were different. But, fourteen of these made errors in typing. Therefore, only 28.8% (38 out of 132) did the analysis completely and correctly.[15] Fifty other labs incorrectly reported that the stains could have shared a common origin, while another twenty-six reported inconclusive results. The study labelled those inconclusives as unacceptable. Even if they are disregarded because they would not lead to court testimony, only thirty-eight labs, or 37.2%, got correct results, while sixty-four had it wrong.[16]

In another test, the labs were given a sample of paint removed from a

---

12. *See* Peterson, Ryan, Houlden & Mahajlovic, *The Uses and Effects of Forensic Science in the Adjudication of Felony Cases*, 32 J. FORENSIC SCI. 1730, 1733 (1987) (firearms analysis is one of five categories of scientific evidence found most often in prosecutorial files).

13. LEAA, *supra* note 1, at 218.

14. A second physiological fluids examination was simpler. Laboratories only had to test for the ABO factor in an uncontaminated bloodstain on clean cloth. Three percent got that factor wrong. A number of labs tried to do more sophisticated analyses to identify other factors. Many of those labs did not do well. For example, only 60% of the labs attempting to type MN in blood got it right. LEAA, *supra* note 1, at 200–01.

15. The study concluded, "The Project Advisory Committee wishes to point out that a correct answer which is only coincidental still constitutes an error." *Id.* at 210.

16. *Id.* at 209. Many of the labs reached the wrong result because they only did ABO typing and more tests were necessary to distinguish the stains. Even so, the level of incorrect typing was high. The report stated:

> 1.6% of the laboratories reporting [the ABO] system [made errors]. Six laboratories . . . of the 30 . . . reporting [the MN] system reported incorrect results. . . . Five of the 20 laboratories reporting results for the Rh system reported incorrect results. . . . 6.1% of the 33 laboratories attempting the PGM system reported incorrect results. One laboratory of the 8 . . . reporting Esterase D results reported an incorrect type. One laboratory of the 7 attempting [the] AK system reported incorrect results, and 1 of the 15 labs reporting the Hemoglobin type reported an incorrect result.

Id. As discussed below, *infra* note 38, an error in typing any one factor makes the entire result erroneous. In other words, the typing errors are cumulative so that the more systems typed, the higher the error rate.

burglarized building's door jamb, and two samples representing paint found on separate suspects. The paints on the suspects differed from the building's paint. Only 48.9% of the labs, however, reported this conclusion. All the others concluded that one or both paint samples recovered from the suspects could have come from the building. In other words, more than half of the labs would have presented inculpatory information when the evidence actually showed the contrary.[17]

Incorrect analyses can lead to miscarriage of justice not only by condemning the innocent, but also by helping to free the guilty. In an arson examination, one piece of cloth was cut in two, with one half soaked in a sample of leaded gasoline. The labs were asked whether the gasoline sample and the gasoline in the soaked cloth could have shared a common origin and, furthermore, whether the two pieces of cloth could have come from the same source. Ten percent of the labs wrongly concluded that the gasoline sample and the gasoline in the cloth were different, and 3.4% incorrectly found that the two pieces of cloth had separate origins.[18] In other words, one-in-seven labs would have reported falsely exculpatory information that might have helped free a guilty person.

Such illustrations are numerous, but the point ought to be clear — the most thorough study of crime labs ever done proves crime lab performance is dangerously poor.

Moreover, error rates in actual casework can be expected to be higher than in the proficiency tests for several reasons. First, only volunteer laboratories were tested, and each participated only in the areas of its choice.[19] Proficiency testing in other areas has established that the better labs voluntarily undergo such examination.[20] Consequently, the error

---

17. *Id.* at 213. An automobile paint comparison produced better data, but still a large percentage of wrong results. While 97 out of 121 labs gave a correct response, "[t]wenty-four laboratories reported results at variance with the manufacturers' statement and the results of the referee laboratories." *Id.* at 203.

18. *Id.* at 224.

19. Some of the labs only participated in one of the 21 tests. Decker, *Access to Expertise*, 101 F.R.D. 636, 637 (1983).

20. *See Health Care Financing Administration's Management of Medical Laboratories: Hearings on S. 765 Before the Subcomm. on Oversight of Government Management of the Sen. Comm. on Governmental Affairs,* 100th Cong., 2d Sess. 46 (1988) [hereinafter *Hearings*] (comments of Dr. Paul Fischer, Associate Professor, Medical College of Georgia): "A laboratory [in a physician's office] which voluntarily participates in PT [proficiency testing] is usually a very good lab.... The fact that most office laboratories do not participate in PT is therefore indirect evidence, in my opinion, that there may be some problems with office testing." *Cf.* LaMotte & Robinson, *Federal Implementation*, in FEDERAL LEGISLATION AND THE CLINICAL LABORATORY 71 (M. Schaeffer ed. 1981) (discussing proficiency tests for toxicology labs).

> Nonparticipation, of course, does not necessarily mean that a laboratory could not perform adequately. Experience has suggested, however, that these laboratories have so little confidence in their test results that they do not report their results ...

rates in the LEAA study would probably increase if all labs had been tested in all subject areas in which they routinely do work. Second, the proficiency analyses were not necessarily done in a blind fashion. The analyst performing the work was not shielded from learning that he was being tested. The analyst will likely be more careful with specimens known to be test samples than with routine work.[21] Empirical data indeed confirm that non-blind proficiency testing leads to more accurate results than when the samples are treated as regular laboratory work.[22] Third, the test samples were much simpler than those a forensic scientist faces in actual casework.[23] Obviously, more difficult and complex samples of actual casework will lead to more errors.

This study of practical forensic science, then, showed most starkly that crime laboratory performance is routinely unreliable[24] and that the quality of forensic science needs drastic improvement.[25] These

---

> for evaluation. One can speculate that the laboratories that do not participate need improvement at least as much as those that do participate.

Id. See also Risinger, Denbeaux & Saks, *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting "Expertise,"* 137 U. PA. L. REV. 731, 749 (1989), who state about the handwriting proficiency tests:

> [C]onsider the possible effect on any aggregate conclusions of the fact that, of the more than 250 police laboratories that perform handwriting examination . . . , only a fraction even ordered test materials in the first place. It is at least arguable that, by self-selection, the sample is inherently biased in favor of the more conscientious and capable practitioners to begin with. If this is true, the reported results would overstate the accuracy of the handwriting examination field generally.

21. Investigators examining the continuing proficiency program, *see infra* text accompanying notes 27–32, have found that in those non-blind tests most labs "do more work on each sample than may be required or than would be done on an actual case." Lucas, Leete & Field, *supra* note 3, at 75–76.

22. *See* LaMotte, Guerrant, Lewis & Hall, *Comparison of Laboratory Performance with Blind and Mail-Distributed Proficiency Testing Samples*, 92 PUB. HEALTH REP. 554 (1977). These investigators sent duplicate specimens disguised as patient samples to drug laboratories that were participating in non-blind proficiency testing. The labs were more accurate on the specimens known to be proficiency testing samples than with the samples thought to be routine patient specimens. *See also* LaMotte & Robinson, *supra* note 20, at 72–73 (summarizing study).

23. *LEAA, supra* note 1, at 257: "Many of the evidence types that were selected for sample manufacturing were not fully exploited and were often presented in their simplest or most unchallenging form. . . . The samples can become more realistic by incorporating contaminants and by minimizing sample size and quantity."

24. *Cf.* Imwinkelried, *supra* note 2, at 637: "[T]he report unquestionably documents a very real possibility of error in the forensic analyses conducted by police laboratories in the United States."

25. The authors of the LEAA report were not surprised by their findings:

> During the course of the proficiency testing program, it was quickly recognized that many of the laboratories were experiencing difficulty in the examination and

conclusions seem to be confirmed by the experiences of independent forensic scientists.[26] Data produced since the LEAA study continue to document the pervasive problems in forensic lab analyses.

B. The Continuing Proficiency Program

A continuing proficiency testing program administered under the auspices of the Forensic Science Foundation provides more recent data on crime lab performance.[27] These data are less complete than that furnished by the LEAA study partly because fewer laboratories participate than did in the earlier government-funded program.[28] No lab, of course,

---

> analysis of various physical evidence types. To be perfectly candid, this could be expected. All of the previous reports which have addressed the issue have inferred the likelihood of such a finding.

LEAA, *supra* note 1, at 261. *See also* M. SAKS & R. VAN DUIZEND, THE USE OF SCIENTIFIC EVIDENCE IN LITIGATION 10 (1983): "Given the low pay and slight emphasis on education and professional qualifications in many forensic science laboratories, [the LEAA] result ought not to be as surprising as it seems." *But cf.* A. MOENSSENS, F. INBAU & J. STARRS, *supra* note 5, at 5 n.10 ("It was believed that such a testing program would establish that crime laboratories performed at high levels of professionalism and obtained accurate results in their examinations . . . .").

26. Some in forensic science stated that their own experiences affirmed the report's findings. *See, e.g.*, Bradford, *supra* note 11, at 904–05:

> These proficiency test findings are reinforced by the work of independent forensic scientists who have the opportunity, through the discovery process in criminal law, to examine and evaluate prosecution evidence from many different crime laboratories. Through the process of discovery, fundamental errors in identification and faulty methods have been found in numerous cases at the trial or trial preparation phase.

27. *See* Lucas, Leete & Field, *supra* note 3, at 74.

> In spite of the strong feelings existing in 1978, many laboratory directors . . . expressed a desire for a continuing proficiency testing program, albeit one without governmental involvement. Collaborative Testing Services Inc. with the affiliation of [the Forensic Sciences Foundation] developed and offered such a program on a subscription basis.

Id. See also Bashinski, *Laboratory Standards: Accreditation, Training, and Certification of Staff in Forensic Context*, in BANBURY REPORT 32: DNA TECHNOLOGY AND FORENSIC SCIENCE 165 (J. Ballantyne, G. Sensabaugh & J. Witkowski eds. 1989).

28. *Cf.* Lucas, Leete & Field, *supra* note 3, at 74:

> While the laboratories participating initially were understandably fewer than in the earlier "free" program, the numbers have continued to increase. In 1984, approximately 142 laboratories subscribed to one or more tests.

Labs must now pay subscription costs for the tests. In 1989, those costs were either $104 or $52 per year for each area of testing. COLLABORATIVE TESTING SERVICES, INC., must take part.[29] A participating laboratory need not be tested in all areas in which it does casework; instead, it can select the areas in which it will be tested.[30] A lab need not even submit responses to the test specimens it obtains.[31] In fact, only a fraction of the labs subscribing to any proficiency test turn in data.[32]

The continuing proficiency program also provides less information than the LEAA study because the test results are not published or otherwise made available to the public.[33] For unknown reasons, this valuable information about a vital part of our criminal justice system largely remains hidden.

The incompleteness and secretiveness of these results make it difficult to draw firm conclusions,[34] but the data that have surfaced confirm the

---

FORENSIC PROFICIENCY TESTING PROGRAM: 1989 [hereinafter BROCHURE].

The participation varied from a high of 114 labs for a drugs test to a low of 40 for a questioned documents test, with an average participation of 82 labs for each of eight tests conducted in 1984. By contrast, for the LEAA study, "219 laboratories in the United States and Canada participated with an average of 188 laboratories for each of the 23 tests." Lucas, Leete & Field, *supra* note 3, at 73. A description by the Collaborative Testing Services, Inc. of the 1989 program indicates proficiency tests in the following areas: drugs, physiological fluids, latent prints, questioned documents, firearms, tool marks, fibers, flammables, footwear impressions, hair, glass, and paint. BROCHURE, *supra.*

29. Bashinksi, *supra* note 27, at 165.

30. BROCHURE, supra note 28.

31. *Id.*

32. In 1985 it was reported, "The response rate is now ranging between 44% and 73% and is gradually increasing . . . ." Lucas, Leete & Field, *supra* note 3, at 74. The average response rate for the eight 1983 tests was 57%. *Id.*

33. The Collaborative Testing Services refused to provide copies of the reports to this author.

Much effort is spent to ensure that only the participating lab has access to its performance results on any test. "To ensure the confidentiality of data and anonymity of participating laboratories, no lab code numbers are assigned and no laboratory can be identified." BROCHURE, *supra* note 28. *See also* Lucas, Leete & Field, *supra* note 3, at 75:

> Confidentiality is assured by providing two envelopes with each test to each participant. A responding laboratory seals its answer sheet inside a plain envelope which is placed inside another envelope and mailed to a member of the [Project Advisory Committee]. The inner envelope is removed and forwarded to [Collaborative Testing Services]. Neither the answer sheet nor the inner envelope has any marks. As a result, no one knows to what laboratory any result can be attributed or, in fact, what laboratories have responded at all.

Of course, summaries of the data could be made available with lab identities still kept confidential. Although that was done in the LEAA study, such data is not available from the continuing program.

34. *See* Lucas, Leete & Field, *supra* note 3, at 76: "Since the response rate is limited and since it is not known whether a laboratory routinely examines a particular type of evidence or is simply 'testing the waters', the data in the summary reports may or may not represent practices and results of the forensic science community as a whole. Extreme care must therefore be exercised in drawing inferences from the results."

findings of the LEAA study—that forensic lab performance is dangerously unreliable. The two proficiency programs for which results are known—handwriting and physiological fluids—support this conclusion.

Professors Risinger, Denbeaux, and Saks obtained the handwriting identification proficiency tests from 1984–1987. Their summary of those tests discloses an astonishing lack of proficiency on the part of crime labs.[35]

> A rather generous reading of the data would be that in 45% of the reports forensic document examiners reached the correct finding, in 36% they erred partially or completely, and in 19% they were unable to draw a conclusion. If we assume that inconclusive examinations do not wind up as testimony in court, and omit the inconclusive reports, and remain as generous as possible within the bounds of reason, then the most we can conclude is this: Document examiners were correct 57% of the time and incorrect 43% of the time.
>
> But let us turn to more meaningful readings of the aggregate data. The pilot test in 1975 may have been unrealistically easy. . . . Omitting the 1975 data, the examiners were correct 36% of the time, incorrect 42%, and unable to reach a conclusion 22% of the time. Even these results are biased in favor of accuracy because of the intentional ease of the 1987 test. Disguised handwriting fooled them all and forged printing fooled two-thirds of those who hazarded an opinion about it.[36]

The authors concluded that "the kindest statement we can make is that no available evidence demonstrates the existence of handwriting identification expertise."[37]

The results of physiological fluids proficiency tests support similar conclusions. A summary of the continuing physiological fluids proficiency program concludes, "From 1978 to 1989, the number of laboratories making one or more errors has ranged from 7% . . . to 78%

35. Risinger, Denbeaux & Saks, *supra* note 20, at 743–51.

36. *Id.* at 747–48. The authors report, "Because of complaints from document examiners that prior tests were too difficult, the Proficiency Advisory Committee decided to make the 1987 test easy." *Id.* at 747.

37. *Id.* at 750–51. The authors also noted that "the studies failed to reveal that certification or experience enhanced accuracy." *Id.* at 749.

... with an average of 25% ...."[38] Certainly, an average error rate of one in four represents an unacceptable threat to life, liberty, and safety.

A closer look at one of the tests where the labs did better than usual underscores this point. A brief scenario accompanying the 1988 test indicated that a one month old child was missing. Several weeks after the disappearance, a blood-stained apron was found in a pile of rags. The adoptive mother said that the blood on the apron was hers. Three bloodstains were submitted for analysis: one representing the apron stain; another from the mother; and the third from the adoptive father.[39] The labs were asked whether the blood on the apron could have come from mother, father, or child.

The samples were sent to 169 labs, but only 67 returned results. Seven of those labs reported at least one result that differed from the correct result. This alone constitutes an error rate of over 10%, but the real rate was even higher. The supplier had difficulty obtaining enough infant or fetal blood from one source, and instead fifteen fetal samples were pooled to make the apron stain. Because of this mixture, the stain

---

38. B. GRUNBAUM, COMMENTS ON THE ADMISSIBILITY OF DNA TYPING EVIDENCE IN AMERICAN COURTS OF LAW 12 (Unpublished, Presented in part, Arbeitsgemeinschaft für Gen-Diognositc E.V., Tübingen, W. Ger., Sept. 11–13, 1989). A survey of the proficiency tests from 1979 through 1983

> shows a 6.3 percent rate of error in ABO determinations of bloodstains. The rate of error in ABO typing of secretion stains (semen, saliva, mixtures of body fluids) is 8.4 percent. The rates of error in the electrophoretic determination of genetic markers in bloodstains are as follows: PGM (2.8 percent), EsD (3.2), EAP (2.1), ADA (O.3), AK (0.7), GLO (5.2), and HP (3.4).

Grunbaum, *Physiological Stain Evidence—Assuring Quality Analysis*, CAL. DEFENDER, Spring, 1985, at 20, 20. The true error rate for these tests is not the number of times a particular factor was mistyped divided by the total number of times all the tests were done. Instead, on any particular sample a number of tests are done. ABO typing may reveal a factor shared by a third of the population. PGM typing may do the same, but together the two tests show that only one in nine have that particular combination of ABO and PGM types. Consequently, the more systems typed, the more the blood can be individualized. That also means, however, that the error rates are cumulative. A typing error in any one system makes the entire result wrong. "Each successive genetic marker determination increases the probability of exclusion but also increases the probability of error." *Id. See also* Bretz, *Scientific Evidence and the Frye Rule: The Case for a Cautious Approach*, 4 COOLEY L. REV. 506, 518 (1987) (expert stated that the proficiency tests showed that error rate for electrophoretic typing in blood was 1.9%, but then admitted that the error rate was 10.9% if computed per sample of blood). Moreover, casework submitted to a lab does not consist of just one blood or bloodstain sample, but of three or more—a sample from the victim, one from the suspect, and one or more crime-scene samples.

39. COLLABORATIVE TESTING SERVICES, INC, PHYSIOLOGICAL FLUIDS ANALYSIS: REPORT NO. 88–14 at 45. (Proficiency reports will be referred to as "CTS" followed by the number of the report. The first two digits of that number indicate the year in which the report was done.)

could not be typed for anything but the presence of fetal hemoglobin. Many of the labs, however, did attempt more. At least three in this group did an erroneous analysis. As the report concluded, "[T]he three labs which reported group O for sample 'a' . . . should review their technique since this is an impossible result for this sample."[40] The report also noted that another lab "reported spermatozoa in sample 'a.' The manufacturer has assured the [Project Advisory Committee] that such contamination could not have occurred in the production process."[41]

All told, of the sixty-seven labs submitting results, at least ten made mistakes, for an error rate of about fifteen percent. Moreover, this was a relatively simple test. No special problems of aged or degraded samples were presented. The stains, unlike real case work, appeared on clean cloth so contamination did not exist.[42] Even so, more than one in seven labs obtained an erroneous result.[43] Even with error rates of this

---

40. *Id.* at 2. The Report continued, "The 15 sources for the pooled sample 'a' consisted of 8 A, 2 B, 1 AB and 4 O. Since ABO isoagglutination production normally begins at about 3 months, no levels of ABO antibodies were expected or reported. The control person found that, by absorption elution, the stain sample 'a' gave a strong A reaction and a weak B reaction." *Id.* One of the labs that erroneously reported O also made other typing errors.

41. *Id.* at 2.

42. While the narrative indicated that the stain must have aged for several weeks and was on an apron, which probably would have been contaminated with many biological products in addition to blood, the actual stain was made on white unwashed cotton cloth, and sent two days after it was made. *Id.*

43. Even though these reports are not intended to be made public, the report did not give a true error rate. *Cf. supra* note 38. Instead, it stated, "Of the approximately 1,000 typing results reported in total, there were 15 incorrect." *Id.* at 2. But 15 incorrect typing results in a thousand translates to 10 erroneous reports out of only 67.

Other reports which I have obtained paint a similar proficiency picture. Thus, in one 1987 test, two samples were sent to 123 labs. Sixty-eight responded, with 16, or 23.5%, giving wrong results. CTS 87–2. In that year's companion examination, 15 out of 54 labs, or 28%, were in error.

In 1986, a hit and run scenario was presented with blood-stained samples, including clothing from the victim, scrapings from the suspect's car, and cloth from the suspect's garbage, and blood from the suspect. Sixty laboratories out of 123 receiving the samples returned information. Fifteen gave incorrect results. CTS 86–11, at 13. In its summary, the Report states, "Of the approximately 1300 results reported, there were 13 which were incorrect." *Id.* at 1. Table 4, *id.* at 13, indicates, however, that 30 errors were made by fifteen different labs.

The other 1986 test was easier since only two stains were sent to the labs. The labs were asked to type the stains for all the systems used by the laboratory. Sixty-six labs returned data; seven reported wrong results for the first sample, and fourteen were in error on the second. There was an overlap in these wrong results, and sixteen labs in all gave wrong results, for an error rate of 24%. CTS 86–2, at 25–26.

The Project Advisory Committee said about one of its 1985 tests, "The samples were adequate and uncomplicated but the results disappointing." CTS 85–11, at 1. Seven stains were sent out: one semen, three blood, and one saliva. Fifty-one labs responded. "Twenty-two labs (43%) reported at least one result different from the manufacturer's data." The other test that year was once again a simpler one with only two samples to be proportion, and the summary of the proficiency reports indicates that on most tests the error rates were significantly higher,[44] crime labs must be making thousands upon thousands of mistaken physiological fluid analyses each year.

C. Other Recent Proficiency Studies

A test of forensic toxicologists not part of the continuing proficiency program also confirms that problems in forensic labs are widespread.[45] The one-year research program "was designed to simulate case samples seen in typical forensic toxicology laboratories."[46] Over one hundred labs participated, and their identities were kept confidential. Blind testing was not required. Thus, performance was probably better than on real casework.

The study found that fewer than seventy-five percent of the labs detected a drug present in seven of the twenty samples sent.[47] By the study's generous definition, problems were apparent in thirty-five percent of the samples. Indeed, although the study did not choose to report

---

typed using all relevant systems in the lab. The Report "noted that 50 of the 62 reporting laboratories got all of the results they reported correct." In other words, the error rate was 12 out of 62, or 19%.

44. As with the LEAA tests, *see supra* text accompanying notes 19–23, the high error rates reported in these two tests indicate an even higher error rate for real casework. The samples in the continuing program are not as complex as those in actual cases. The tests "are not 'real world' by their very nature. The samples must be available in sufficient quantity for identical samples to be sent to up to 115 or more laboratories, they must impose only reasonable demands on time and equipment for the laboratory, they must be amenable to a range of analytical techniques and there must be no tricks." Lucas, Leete & Field, *supra* note 3, at 77. Blind testing is rare: "No instructions are given as to how to examine the sample. Some laboratories assign their most experienced examiners while others use the samples for training new examiners. In a few instances, laboratory directors have repackaged the samples and assigned them as real cases." *Id.* at 76. The voluntariness of the proram, of course, points toward a higher error rate if all labs participated, as does the fact that only a fraction of the subscribing labs return results. *Cf.* Risinger, Denbeaux & Saks, *supra* note 20, at 748–49: "Assuming that an examiner who has worked on an answer and then decides not to return it has serious doubts about its accuracy, then the sample of respondents is composed of an unrepresentatively large proportion of those who obtained—or at least think they obtained—correct answers."

45. *Cf.* Imwinkelried, *The Standard for Admitting Scientific Evidence: A Critique From the Perspective of Juror Psychology*, 28 VILL. L. REV. 554, 561 (1983) ("The results of [the toxicology] survey are consistent with the findings of the [LEAA] Laboratory Proficiency Testing program.").

46. Peat, Finnigan & Finkle, *Proficiency Testing in Forensic Toxicology: A Feasibility Study*, 28 J. FORENSIC SCI. 139, 140 (1983).

47. *Id.* at 149. The tested laboratories were also asked for quantitative analyses of the detected drugs. These results also revealed problems in the analyses. The report concluded, "These particular examples demonstrate the considerable interlaboratory variation in quantitation." *Id.* at 157.

its data this way, the total percentage of correct positive results for the entire proficiency test was only seventy-one percent.[48] This study, therefore, revealed a pervasive problem for forensic toxicology.[49]

DNA analysis proficiency testing has found problems consistent with the studies from other areas. A test conducted by the California Association of Crime Laboratory Directors reported that "two of the three private laboratories made an error in analyzing samples. One company was wrong in one of the forty-four matches it identified, another was wrong in one of fifty matches, and only the third company was correct in all of its matches."[50] While this level of proficiency seems much better than that reported on other tests, the New York State Panel assessing forensic DNA analysis labelled these outcomes "disturbing." The panel noted that "These results fall far short of the private laboratories' claims of absolute certainty of forensic DNA testing."[51]

Information which surfaced later indicates that the results of that testing were more disturbing than first thought. Documents obtained through discovery in a criminal case indicated that one of the three tested labs had actually produced two reports. The first response "contained an

48. Calculated from *id.* at 145 (Table 3) which indicates that the total number of potentially correct positive analyses was 4,142, but that the laboratories only performed 2,924 correctly. Unfortunately, the study did not present complete information about false positives. It gave details only about the percentage of positive responses for each sample without giving a full listing of all the falsely "detected" drugs. *See id.* at 145 (Table 3). The report merely concluded, "The rate of false positive results was particularly low with one notable exception." *Id.* at 144. For that sample, 13% of the labs falsely "detected" one drug and 11% incorrectly reported another. No further information about false positives was presented. The overall rate was not given, nor can it be calculated from the data published.

49. These toxicology findings confirmed results from an earlier proficiency report. *See* Dinovo & Gottschalk, *Results of a Nine-Laboratory Survey of Forensic Toxicology Proficiency*, 22 CLINICAL CHEM. 843 (1976). *See* Kelly & Sunshine, *Proficiency Testing in Forensic Toxicology: Criteria for Experimental Design*, 22 CLINICAL CHEM. 1413 (1976); McCloskey & Finkle, *Proficiency Testing in Forensic Toxicology*, 22 J. FORENSIC SCI. 675 (1977); Dinovo & Gottschalk, *More on Proficiency Testing in Forensic Toxicology*, 22 CLINICAL CHEM. 2056 (1976). "The major finding of this proficiency study was that the toxicological laboratories ... varied considerably in the precision and accuracy with which they performed drug assays." Dinovo & Gottschalk, *Results of a Nine-Laboratory Survey of Forensic Toxicology Proficiency, supra*. The authors noted that this variation could have significant consequences: "[S]ince there exists a wide interlaboratory variance in the detection, accuracy, and precision of toxicological analysis, this wide variance is probably reflected in the drug-death statistics and drug-death certification in these and other cities in the United States." *Id.*

50. J. POKLEMBA, DNA: REPORT OF THE NEW YORK STATE FORENSIC DNA ANALYSIS PANEL 28 (1989). *See also* Note, *The Dark Side of DNA Profiling: Unreliable Scientific Evidence Meets the Criminal Defendant*, 42 STANFORD L. REV. 465, 493 (1990) (two out of three labs "did not complete fifty tests without inculpating an innocent person—surely an unproficient report").

51. J. POKLEMBA, *supra* note 50, at 28–29.

extraordinary number of misclassifications, including one false positive, at least three false negatives ... and at least two incorrect reports of mixed stains."[52] A commentator, after reporting this further information, concluded that this "study is the only meaningful blind trial of the proficiency of DNA laboratories" and it "raises serious concerns" about the quality of the work being done.[53]

In sum, a review of the data revealed by proficiency studies indicates that lab performance is inadequate and unreliable.[54] The most thorough

52. Thompson, *Cellmark's Errors on the CACLD Proficiency Study*, in 2 DNA: UNDERSTANDING, CHALLENGING AND CONTROLLING THE NEW EVIDENCE OF THE 90'S 2 (1990). *See also* Thompson, *Letter*, CAL. LAW., July, 1989, at 11, 11. "[T]he President of the California Association of Crime Laboratory Directors ... helped cover up errors made by a commercial DNA laboratory in a 'blind' test of its accuracy.... She contacted Cellmark officials, met with them to review the problems, and allowed them to submit cleaned up conclusions three months later." *Id.* This conclusion was hotly denied by the DNA Committee of the California Association of Crime Laboratory Directors:

> Our committee found Cellmark's initial blind trial report incomprehensible and impossible to score. Since Cellmark's Dan Garner was also unable to interpret this report, we asked him to submit another one in an understandable format. We also discussed with him the one error we noted in the autorads which we had reviewed out of his presence. No other information regarding the "correct answers" was revealed to or discussed with him. The blind aspect of the test was never compromised.

Bashinski, Kuo & Hartstrom, *More on DNA*, CAL. LAW., Sept., 1989, at 17, 17. Thompson, however, has stuck by his conclusions. "The claim that Cellmark's initial blind trial report was 'incomprehensible and impossible to score' is nonsense. Anyone who compares Cellmark's report with the CACLD records showing the true identity of the samples will readily comprehend that Cellmark made a number of errors." Thompson, *Letter*, CAL. LAW., Sept., 1989, at 18, 18.

53. Thompson, *Cellmark's Errors on the CACLD Proficiency Study*, *supra*, note 52, at 1. Other reports indicate that this study was not truly a blind test. "[T]he laboratories made the mistakes knowing that their results would be scrutinized carefully." J. POKLEMBA, *supra* note 50, at 29. *See also* Note, *supra* note 50, at 493 ("Although Lifecodes called all fifty correctly, its researchers, rather than the technicians who usually perform it, completed the test.").

54. One article does indicate excellent forensic science proficiency, but it does not present enough information to begin to rebut the rest of the data about crime lab quality. Scientists from the national laboratory of the Bureau of Alcohol, Tobacco and Firearms ("ATF") described a quality assurance program there for arson and explosives cases, which includes blind proficiency testing. Brunelle, Garner & Wiseman, *A Quality Assurance Program for the Laboratory Examination of Arson and Explosives Cases*, 27 J. FORENSIC SCI. 774 (1982). The authors reported, "All blind tests so far have resulted in 100% correct identifications by the examiners tested." *Id.* at 780. The authors also state that in the test's final step "results are evaluated and necessary corrective action is identified and implemented as necessary." *Id.* The authors do not explain what corrective actions could be necessary when the examiners have only produced all correct identifications.

Other than the barest outline of the program, however, no additional, useful data about the testing was published, and the bureau has refused to make such information public. This author requested the results of ATF's proficiency testing. By a letter dated September 21, 1989, Richard E. Tontarski, Chief, Forensic Science Laboratory of the ATF replied that of the tests, the LEAA study, showed abysmal performances, and all subsequent testing indicates that problems persist.

II. FORENSIC SCIENCE EDUCATION AND RESEARCH

A. Forensic Science Education and Training

The generally poor training and education of forensic scientists is one cause of substandard performance. There is no defined or organized preparation for the profession of forensic scientist.[55]

> [There] is no uniform or core curriculum or internship that leads to the practice of criminalistics[;] there are no minimum course requirements in terms of a structured program[;] there is not even a consensus of what the educational requirements should be in the specialized forensic science subjects[;] and there are no codified standards of practice, either formal or informal, in the identification aspects of criminalistics toward which an educational program can be planned.[56]

While forensic science has gained importance in the criminal justice arena, it has not enjoyed a similar rise in stature in the academic com-

---

the testing was an "internal review process," was used only for internal quality control, and could not be made available to outside sources. Basically, then, one has to take it on faith that this blind proficiency testing was well-designed and run. Such a proficiency program does not support a general conclusion that forensic laboratories perform well.

55. The title of this law-science profession is not certain. It is referred to by a variety of appellations including criminalistics and forensic science. *See generally* Blanke, *Discussion of "The Desirability of a Ph.D. Program in Forensic Science,"* 30 J. FORENSIC SCI. 309 (1985); Kirk, *The Interrelationship of Law and Science*, 13 BUFFALO L. REV. 393 (1964); Kirk, *The Ontogeny of Criminalistics*, 54 J. CRIM. L. CRIMINOLOGY & POLICE SCI. 235 (1963); Peterson & DeForest, *The Status of Forensic Science Degree Programs in the United States*, 22 J. FORENSIC SCI. 17 (1977).

56. Bradford, *supra* note 11, at 905. For similar discussions by others noting the lack of specific requirements in some forensic science subspecialties see Risinger, Denbeaux & Saks, *supra* note 20, at 772 n.183 (handwriting analysis); Peterson & DeForest, *supra* note 55, at 18 (examination of questioned documents); Cowger, *Moving Towards Professionalization of Latent Print Examiners*, 24 J. FORENSIC SCI. 591, 591 (1979) (fingerprinting identification); Koenig, *Spectographic Voice Identification*, 13 CRIME LAB. DIG. 105, 115–16 (1986) (voiceprint examination); Winek, *The Adversary System: Role of the Forensic Toxicologist*, 18 J. FORENSIC SCI. 178, 179 (1973) (forensic toxicology); Wilkinson & Gerughty, *Bite Mark Evidence: Its Admissibility is Hard to Swallow*, 12 CAL. W. L. REV. 519, 538–39 (1985) (forensic odontology).

munity.[57] Instead, the number of forensic science academic programs has been declining.[58] The remaining group has done little to define educational requirements.

> The Council of Forensic Science educators has recently been formed with the specific goal of developing standards for forensic science education.... However, a uniform core curriculum has not yet been adopted in the field, nor are internships widely available to enrich academic programs.[59]

Furthermore, the few academic forensic science programs that do exist are not particularly strong.[60] Since they have modest enrollments,

---

57. The relationship between academic and practicing forensic science has been haphazard. *Cf.* Peterson & DeForest, *supra* note 55, at 25 ("Although at the local level colleges and universities do interact with neighboring crime laboratories, there is no systematic or regular exchange of information among schools and laboratories at a national level.").

58. "The number of college and university-based programs in forensic sciences ... probably reach[ed] a high watermark of about 25 around 1978. [Since then] the number of academic programs has been declining. At present, there are about 15." Gaensslen & Lee, *Regional Cooperation and Regional Centers Among Forensic Science Programs in the United States*, 33 J. FORENSIC SCI. 1069, 1069 (1988). Many of the about 600 criminal justice programs offer some courses in forensic science without offering a forensic science degree. *See* Peterson & DeForest, *supra* note 55, at 18. *See also* Bashinski, *supra* note 27, at 161: "Formal academic course work in forensic science is available in a few institutions at both the undergraduate and graduate levels, the latter most notably in the Department of Public Health at the University of California at Berkeley and in John Jay College of Criminal Justice at City College of New York."

59. Bashinski, *supra* note 27, at 161.

60. Crime laboratory managers and other forensic scientists have regarded degrees in forensic science with suspicion. One survey concluded:

> Another result ... which stands out is the lack of preference for the B.S. in Criminalistics/Forensic Science.... In the case of the crime lab directors, [it] ranked last among the seven alternatives presented. Those who commented stated that they felt that too many programs passing themselves off as forensic science programs were actually little more than criminal justice programs with a forensic science internship and a smattering of "hard" science.... [T]here is apparently little uniformity among programs which call themselves forensic science.

Siegel, *The Appropriate Educational Background for Entry Level Forensic Scientists: A Survey of Practitioners*, 33 J. FORENSIC SCI. 1065, 1068–69 (1988). *See also* Higgins & Selavka, *Do Forensic Science Graduate Programs Fulfill the Needs of the Forensic Science Community?*, 33 J. FORENSIC SCI. 1015, 1017 (1988). (Summarizing another survey which reached similar conclusions and elaborated on the preference for chemistry degrees over Forensic Science Degrees); Peterson & DeForest, *supra* note 55, at 31 ("the capabilities of graduates from the respective institutions are not uniform. Laboratories are forced to evaluate each graduate individually ....") Others have advocated that a forensic science degree is better because

> the forensic science major has selected his field and has demonstrated the motivation and ability to handle such a curriculum: he may have completed course work in

university support is limited. Consequently, "few if any of the programs can cost-effectively support a full-time faculty and staff appropriately representative of the many subspecialties that make up forensic sciences . . . ."[61]

As a result, forensic science faculty have heavy teaching loads, are disproportionately composed of adjuncts,[62] and are not highly regarded by academics in other areas.[63] Students, often not the strongest in the institution,[64] do not get the best possible education.[65] Because entry-

---

> such areas as law, evidence, ethics, and criminal investigation which will speed his transition into the field; he has acquired basic scientific concepts with an emphasis on forensic science applications; and perhaps most important, he has developed a forensic science ethic or way of thinking which is unique to this particular profession.

Id. *Cf.* Blanke, *supra* note 55, at 309 ("[O]ne can no more offer a Ph.D. degree in forensic science than a Ph.D. degree in general science. Is it really possible for one individual to master forensic pathology, forensic psychiatry, and all the other forensic specialties, to the extent implied by a Ph.D. degree?").

61. Gaensslen & Lee, *supra* note 58, at 1069 ("there tends to be limited support for space, equipment and supplies").

62. One survey found that forensic science specialists are much more likely than criminal justice faculty generally to be at a two year college than a four year institution and are two-and-a-half times more likely to be an adjunct than the criminal justice faculty generally. These and other factors indicate that the forensic science faculty's "affiliation with the university is temporary and not like likely to result in a stable, long-term relationship." Peterson & Angelos, *Characteristics of Forensic Science Faculty Within Criminal Justice Higher Education Programs*, 28 J. FORENSIC SCI. 552, 555 (1983). *See also* Higgins & Selavka, *supra* note 60, at 1018.

63. Forensic science faculty do little research and produce few publications. *See infra* text accompanying notes 90–93. "One cannot escape the fact that publications constitute the primary means by which faculty are evaluated and in this regard the record of the forensic science faculty . . . is not impressive." Peterson & Angelos, *supra* note 62, at 558.

64. "When questioned, institutions reported that most students [in a forensic science degree program] fell into the 'above average' category but that very few could be rated as 'excellent.'" Peterson & DeForest, *supra* note 55, at 32.

65. Peterson and DeForest report:

> While many of the programs possess faculties with both specialized training and experience in the forensic science field, others do not and depend entirely on adjunct faculty for professional expertise. . . . The use of adjunct lecturers is quite acceptable for the teaching of specialized courses but cannot take the place of dedicated, qualified full-time faculty. . . .
>
> Teaching environments should be avoided where experienced forensic scientists serving as adjunct faculty are teaching courses where their own employees make up the majority of students. . . . Ideally, educational institutions should introduce students to a variety of alternative procedures so that the graduate can possess a broader perspective upon entering the field.

Id. at 31–32.

level salaries in forensic science are low, the top graduates often pick other careers.[66]

Since few colleges and universities provide adequate preparation for potential forensic scientists, the real training and education occurs during employment. "At present, professional training is achieved primarily through on-the-job apprenticeship, workshops, and seminars. . . . The major burden falls on the crime laboratories themselves to ensure the professional development of their staff by providing in-house training programs . . . ."[67] This on-the-job training is often insufficient. "The quality of training will be dependent on continued budgetary support. Inasmuch as education is not the mission of a public service laboratory, it will necessarily take on secondary importance when funding becomes restricted or when services are in great demand."[68]

Reliance on on-the-job training is especially troublesome for the numerous small forensic labs[69] and for "many so-called 'full-service' laboratories [where forensic scientists] are expected to be generalists with expertise in several diverse areas such as drug chemistry, arson and explosives, ballistics, toxicology, and trace evidence analysis."[70] Although some labs continue a generalist approach, even when unneces-

---

66. Salary data indicate:

> While the median salary for new M.S. graduates in chemistry was near $27,000 in 1985, that is, a 13% increase over the 1984 figure of $24,000, a quick survey of the offers expressed in the want ads for forensic science positions in the past six months revealed that the median starting salary would fall between $22,000 and $25,000 for similarly trained individuals in forensic science.

Higgins & Selavka, *supra* note 60, at 1016.

67. Bashinski, *supra* note 27, at 161–62. Many others have noted this fact. "Without a standardized and relevant university education, a substantial amount of fundamental instruction must be provided within the individual forensic science laboratories." Stoney, *A Medical Model for Criminalistics Education,* 33 J. FORENSIC SCI. 1086, 1087 (1988). at 1087. *See* Bretz, *supra* note 38, at 511. *See also* Grunbaum, *Professional Responsibility: Development and Enforcement of Standards with United States Crime Laboratories for the Genetic Marker Typing of Physiological Stain Evidence* in ADVANCES IN FORENSIC HAEMOGENETICS 501, 502 (B. Brinkmann & K. Henningsen eds. 1985) ("Most crime laboratory analysts learn these techniques on the job or [in] workshops that are oriented toward the transfer of technical skills."); Winek, *supra* note 56, at 179 ("The remainder of [the forensic toxicologist's] training comes from the less formal 'on the job' or apprentice type training.'").

68. Stoney, *supra* note 67, at 1088.

69. A recent survey found that half of the forensic science laboratories employed six or fewer scientific personnel. Peterson & Angelos, *supra* note 62, at 13.

70. Grunbaum, *supra* note 67, at 502. *See also* Peninger v. State, 721 P.2d 1338 (Okla. Crim. App. 1986); Rogers v. State, 721 P.2d 805 (Okla. Crim. App. 1986) (forensic scientist who had typed blood and semen also did hair comparisons).

sary,[71] training anyone adequately for such positions seems impossible. As a distinguished British forensic scientist recognized a generation ago, "If forensic science is to grow as it should . . . or is even to use all the scientific knowledge which even now it could, it just has to be a collaborative affair. No one can now be an expert in all its branches."[72] Thus, a person does not become a forensic scientist through formal education or a prescribed apprenticeship, but by convincing the right agency to employ him and by having his work accepted in court.

> What are the entry requirements [to forensic science]? Employment and function. One joins the profession when one is hired by a crime laboratory and when one begins to write reports and to testify in court. In effect, the entry gate of our profession has been delegated to individual civil service boards and to individual judges. Educational requirements, such as they are, are determined by individual laboratories.[73]

One forensic scientist maintains, "The fundamental requirements of achieving a high standard of performance and accuracy in forensic science are a good general scientific education coupled with specialised training in the subject and a constant monitoring of performance."[74] If accurate and reliable work demands these components, the quality of our crime laboratories is in trouble. American forensic science has no guarantee that its practitioners have that necessary education or specialized training.[75]

---

71. "In many crime laboratories, forensic scientists are specialists in one of these subdisciplines of criminalistics. In other laboratories, a 'generalist' approach is employed, with analysts rotating through various areas of the laboratory." Bashinski, *supra* note 27, at 160.

72. Walls, *Whither Forensic Science?*, 6 MED. SCI. & L. 183, 183 (1966).

73. Stoney, *supra* note 67, at 1088. The standards for consultant forensic scientists are even less. Criminalists who work as private consultants are neither licensed by the state nor certified by any professional group. They are self-appointed experts." Grunbaum, *Problems Inherent in the Analysis of Rape Evidence—State of the Art*, FORUM: CAL. ATT'YS FOR CRIM. JUST., Sept.–Oct., 1986, at 30, 31.

74. Pereira, *Quality Assurance in Forensic Science*, 28 FORENSIC SCI. INT'L 1, 1 (1985).

75. A reliance upon on-the-job training has proved detrimental to clinical laboratory quality.

> In its development, this complex scientific activity has been beset with a variety of problems relating to quality assurance. Initially, [clinical] tests were performed by a limited number of skilled persons and by individuals receiving varying amounts of on-the-job training. As the technology developed, these marginally trained individuals could not keep pace with the introduction of newer and more complicated methods. The need for scientific education was recognized early by some workers in the field; while others erroneously held to the position that on-the-job training was sufficient if certain external controls were applied.

Furthermore, forensic scientists are not required to establish competence by obtaining a license or certification. Although states require licenses for many occupations,[76] no jurisdiction requires a forensic scientist to be licensed.[77] Even when the forensic scientist performs the same analytical procedures that require a licensed professional outside the crime laboratory, he is not required to be licensed. For example, to determine whether a bloodstain at the crime scene came from the defendant, the forensic scientist must type the defendant's blood. This typing uses the same procedures used in a blood bank or a hospital. States that are concerned about the accuracy of clinical typing, require clinical analysts to be licensed. However, they make no such demands on the forensic serologist. "To reduce error [in medical laboratories], you have to prove and demonstrate competency in that lab or office where the tests are being done."[78] This concept is simply not applied by the states to forensic science laboratories.

Forensic scientists have failed to rectify the lack of governmental oversight with a self-regulation system. A few voluntary certification programs do exist,[79] but forensic science has not developed general, national standards to assure competency.[80] Indeed, the largest segment of forensic science has fiercely resisted even a voluntary certification program.

Bodily, *Foreword* to FEDERAL LEGISLATION AND THE CLINICAL LABORATORY at x (M. Schaeffer ed. 1981).

76. "License" means a mandatory permit required to practice a profession legally. "[P]rofessionals, including lawyers, doctors, nurses, clinical analysts, pharmacists, teachers, even hairdressers and barbers, must meet certain educational standards and qualify to be licensed to practice their profession." Grunbaum, *Genetic Typing of Physiological Stain Evidence: A Proposal for Quality Assurance Guidelines*, FORUM: CAL. ATT'YS FOR CRIM. JUST., Nov.–Dec., 1984, at 11, 14.

77. "Thus far, no state has enacted licensure for crime laboratories, although several have requirements in restricted areas such as forensic alcohol analysis." Bashinski, *supra* note 27, at 166.

78. *Hearings, supra* note 20, at 43 (statement of Dr. George M. Hoffman, American Society of Clinical Pathologists).

79. *See, e.g.*, Bashinski, *supra* note 27, at 166 (noting the uneven pace at which certification programs have developed in various areas of the forensic sciences). "Certification" is used in the sense of a profession regulating itself: "a voluntary process of peer review whereby a practitioner is recognized as having attained the minimum qualifications necessary to practice in one or more particular disciplines of criminalistics." *Id.* "Certification" is, thus, being used differently from its use in other fields where certification only comes after a license to enter the profession has been obtained.

80. *E.g.*, Bretz, *supra* note 38, at 511. *See also* Giannelli, *The Admissibility of Laboratory Reports in Criminal Trials: The Reliability of Scientific Proof*, 49 OHIO ST. L. J. 671, 690 (1988) ("National standards to ensure the competency of [forensic] examiners have not been developed.").

> A proposal for national certification [of criminalists] was made in 1979 by the national [Criminalistics Certification Study Committee] on the basis of a 3-year study of the issues conducted under the auspices of the [Law Enforcement Assistance Administration.] This proposal did not receive support of the majority of the profession at the time, and the profession has yet to adopt a national certification program.[81]

That the proposal did not receive support is an understatement. It was "resoundingly rejected by the criminalists."[82] It "was defeated by a 2-to-1 margin."[83]

Voluntary certification, even if widely prevalent, would have only a limited effect on forensic science quality even if widely prevalent. As clinical laboratory studies establish, nonmandatory regulatory systems are not as effective in producing quality control as are mandatory programs.[84]

Moreover, certification can be meaningful only if it has important consequences. "The value of certification is in the prestige and economic advantage afforded to those certified."[85] For example, physicians certified as specialists "benefit economically and [are] eligible for more hospital affiliations."[86] However, no consequences flow from the lack of a forensic science certificate. Courts do not deny forensic scientists expert status because they lack certification.[87] The forensic scientist continues to work with or without participation in one of the few existing certification programs. The following statement concerning certification for fingerprint analysts applies generally to all forensic scientists:

> In a 1982 report on a certification program given to "existing fingerprint experts," fifty-three percent failed the exam. Because fingerprint "experts" need not be certified to qualify

81. Bashinski, *supra* note 27, at 166–67.

82. Grunbaum, *supra* note 76, at 14.

83. Peterson, *Ethical Issues in the Collection, Examination, and Use of Physical Evidence*, in FORENSIC SCI. 34, 43 (G. Davies ed. 1986).

84. *See supra* text accompanying notes 6–11. Consequently, as one health worker concluded, "Delegation of authority to peer review organizations is clear abdication of the regulatory role." *Hearings*, *supra* note 20, at 59 (statement of Herbert W. Dickerman, New York State Department of Health).

85. M. BARAM, ALTERNATIVES TO REGULATION: MANAGING RISKS TO HEALTH, SAFETY AND THE ENVIRONMENT 62 (1982).

86. *Id.* at 64.

87. Or at least I have found no such cases. *Cf.* M. SAKS & R. VAN DUIZEND, *supra* note 25, at 61 ("We found [forensic science certification programs] virtually unknown to the bar and bench.").

> as experts, those that failed the test can be expected to testify in court.[88]

Not surprisingly, no published study has shown that the few certified forensic scientists perform any better than their uncertified colleagues.[89]

Since forensic scientists as a group are not well educated or trained, since they are not required to demonstrate their capabilities, and since most resist voluntary programs requiring demonstrations of their abilities, good reasons exist to distrust the quality of their work. There are, however, further causes of poor crime lab performance.

B. University and Crime Laboratory Research

While one might expect universities to be a major source of forensic science research,[90] they actually do very little of it.

> As a matter of practice, academic institutions are heavily involved in basic scientific research, so it seems only logical that graduate programs in forensic science should be actively

88. Bretz, *supra* note 38, at 511 (citing Bonebreak, *International Association for Identification*, Newsl. to Chesapeake Division (1982)).

89. Certification can only be effective if the certifying bodies have an effective disciplinary mechanism in the event of incompetence. At the present time, courts make some attempt to control forensic experts. *See* McCarty v. State, 765 P.2d 1215 (Okla. Crim. App. 1988). The court reversed a capital murder conviction because a forensic scientist, Joyce Gilchrist, gave personal opinions on scientific matters beyond her scientific capabilities. Even though the court concluded that Ms. Gilchrist had violated the ethical code, she was not reprimanded. *See also* Fox v. State, 779 P.2d 562 (Okla. Crim. App. 1989) (court found that Gilchrist again gave an opinion outside of her scientific knowledge). *But cf.* M. SAKS & R. VAN DUIZEND, *supra* note 25, at 78 (reporting the statement of a former chair of the ethics committee of the American Academy of Forensic Sciences that the committee had never concluded that an expert witness had asserted an opinion unwarranted by the available data).

90. All forensic science practitioners, however, do not welcome more research from academic institutions. A survey found that many forensic scientists

> felt that an atmosphere should exist where the research conducted in this country is the result of a combined effort by academia, state, federal, and local laboratories and national research laboratories. Many others felt that the state laboratories know the problems, have more experience with them on a "grass-roots" level, and therefore should do more research, but that these laboratories lack time and funding. Academic laboratories were posited to have time, and the federal laboratories to have funding and time, but the respondents felt that these labs conduct too much esoteric research that is useless for practitioners who lack the instrumentation or skills to utilize such research.

Higgins & Selavka, *supra* note 60, at 1019.

> involved in such research. However, the fact remains that far less than what is expected is actually performed in the academic atmosphere.[91]

A survey of forensic science academies showed that "95% spent less than 25% of their time on research . . . , and none of the respondents reported that they devoted more than 50% of their time to research activities . . . ."[92] Furthermore, those few with a research orientation have difficulty finding money to support their efforts. "[O]pportunities for external grants and contracts to which many parent institutions expect programs to look for additional support of both programmatic and research activities are very limited."[93]

The absence of university research harms the quality of the forensic sciences[94] and forces the research on to the crime laboratories. These facilities, however, cannot produce much significant research. Those who go into crime labs typically lack the training and skill required to be good research scientists. As noted above, the academic programs that produce many forensic scientists often have low status and do not attract the best students.[95] Research-inclined students are unlikely to enter such programs. Forensic science instructors, not themselves researchers, are unlikely and probably unable to teach their students the fundamentals of good research.

Moreover, well-educated forensic scientists with good research skills have little time and few resources for research. "[H]igh case backlogs are the rule, not the exception, in forensic science, and there is generally little time or money available to allow for research in this environment."[96] Surveys find that the mean amount of time spent in crime labs researching new laboratory techniques is 4.4%[97] and that those

91. *Id.* at 1016; *but see* Smith, Liu & Lindquist, *Research Experience and Future Criminalists*, 33 J. FORENSIC SCI. 1074, 1075 (1988) (supporting the assertion that "a substantial portion of forensic science research does originate from universities").

92. Peterson & Angelos, *supra* note 62, at 555.

93. Gaensslen & Lee, *supra* note 58, at 1069.

94. "The lack of [research and writing] hinders the faculty member's long-range professional development and learning. It also retards the development of new scientific procedures and knowledge that are crucial to the growth of the forensic science profession." Peterson & Angelos, *supra* note 62, at 558. *See also* Higgins & Selavka, *supra* note 60, at 1018 (noting that the heavy reliance on adjuncts for forensic science education bodes ill for research).

95. *See supra* text accompanying notes 60–66.

96. Higgins & Selavka, *supra* note 60, at 1016.

97. *See* Peterson, Mihajlovic & Bedrosian, *The Capabilities, Uses, and Effects of the Nation's Criminalistic Laboratories*, 30 J. FORENSIC SCI. 10, 19 (1985); *but see id.* at 20 ("In terms of published journal articles, 36% of the laboratories surveyed reported they had at least one article published in a scientific journal in 1982.").

laboratories not involved in research outnumber those that are.[98] Commentators have concluded that "American crime laboratories are not research laboratories, and many crime laboratory analysts have neither the education, training, time nor resources to conduct extensive reliability experiments."[99]

C. Peer Review and Forensic Science

This research picture indicates that good forensic studies are few, and that forensic scientists are generally outside of the peer review system. Science usually takes place in a setting where the scientific community judges the work of fellow scientists. As one commentator notes, "[S]cience is in some respects a self-governing republic, with scientists deciding what is good work and what is not."[100] This self-governance takes place through peer review.[101] "Peer review has . . . been labeled 'institutional skepticism' and is best understood as a process that allows the scientific community to police itself."[102] Although peer review may encompass many formal and informal procedures,[103] one of the two

---

98. *See* Higgins & Selavka, *supra* note 60, at 1019.

99. Grunbaum, *supra* note 76, at 13. Since few people in universities and practicing crime labs are doing research, a small number of federal laboratories are the major source of good, extensive research. However, they cannot truly fill the need.

> Thus, the onus is placed on others to perform the basic research required for the improvement of the discipline. Three large national laboratories, the Drug Enforcement Administration (DEA), the Bureau of Alcohol, Tobacco and Firearms (ATF), and the Federal Bureau of Investigation (FBI), have the funding and facilities available for the basic research in forensic sciences. However, it is clear that these three organizations cannot be solely responsible for fulfilling all the research needs of the entire forensic science community. The DEA laboratory reported a decline in the amount of time devoted to research from 4 to 1.5% for the 1978 to 1982 interval, and the ATF's decrease ranged between 5 and 7% for the same period.

Higgins & Selavka, *supra* note 60, at 1016.

100. Goldberg, *The Reluctant Embrace: Law and Science in America*, 75 GEO. L.J. 1341, 1361 (1987).

101. *Cf.* Anderson, *The Federal Government's Role in Regulating Misconduct in Scientific and Technological Research*, 3 J.L. & TECH. 121, 125 (1988) ("The three principal self-correcting mechanisms of science are replication of experimental work, peer review of proposals for funding, and merit review of articles before publication in scientific journals.").

102. Note, *Of Reliable Science: Scientific Peer Review, Federal Regulatory Agencies, and the Courts*, 7 VA. J. NAT. RESOURCES L. 27, 32 (1987) (quoting W. LOWRANCE, MODERN SCIENCE AND HUMAN VALUES 85 (1985)).

103. *See, e.g.*, Woolf, *Deception in Scientific Research*, 29 JURIMETRICS J. 67, 81–82 (1988) ("[P]eer review is often quite narrowly construed as consisting only of *editorial* peer review. . . . [T]he definition [should include] all the evaluative processes of collegial interaction—from lab seminars to public professional meetings, from reading a paper for a friend to assessing it for grant support.") (emphasis in original).

major peer review mechanisms is the critical review that takes place when a scientist seeks funding for research. Most granting agencies rely on panels of experts to assess the worth of a proposal.[104] "Research funded by such agencies, therefore, has often survived intense and highly competitive screenings."[105]

Scientific publishing provides the other major peer review mechanism. In determining whether a manuscript should be published, the editors of most respected scientific journals seek the comments and analyses of the author's peers.[106] Thus, published articles must also have survived an intense and highly competitive screening.

The peer review process in funding and journal decisions requires scientists whose careers are dependent on grants and publications to continually propose and carry out projects that will be carefully assessed for their scientific rigor. This provides a powerful inducement to do quality work.[107] Since forensic scientists do not spend much effort on research, rarely publishing or seeking grants, they do not experience the forces of the peer review system that push science towards quality.[108]

---

104. The funding system for the National Institutes of Health ("NIH") provides a typical example. The NIH

> use advisory panels, known as "study sections," which consist of fourteen to twenty independent nongovernmental scientists that review the scientific and technical merits of all research grant proposals. These study sections make recommendations to a separate advisory council, which in turn makes funding recommendations on individual grant proposals to NIH.

Note, *supra* note 102, at 33–34. *See also* Monahan & Walker, *Social Authority: Obtaining, Evaluating, and Establishing Social Science in Law*, 134 U. PA. L. REV. 477, 500 (1986) ("Federal funding agencies generally rely on an expert review panel to evaluate the methodological adequacy of research proposals.").

105. Monahan & Walker, *supra* note 104, at 500.

106. *Cf. id.* at 500 (discussing peer review in the context of research publications within the social sciences).

107. These peer review processes serve not only as inducements to do quality research, but also have an educational function. The reviewers for the funding agencies and the publishers usually give written critiques of the proposals. Scientists who make submissions learn how some other scientists believe better quality science can be done. Indeed, many proposals that are rejected are improved as a result of peer review evaluation, resubmitted, and then accepted for funding or publication. Thus, this peer review process is an active process impelling scientists towards better science as well as a mechanism for quality control.

108. Forensic science operates more like an industrial science than an academic science. Because industrial science is subjected to little peer review, it tends to be less reliable than science produced in other areas. Furrow, *Governing Science: Public Risks and Private Remedies*, 131 U. PA. L. REV. 1403, 1411 (1983).

Compare the view expressed in J. POKLEMBA, *supra* note 50, at 44, which argues that courts should not look to publications in peer review journals to determine whether DNA profiling should be admitted, but instead suggests the creation of a Scientific Review Board to act as an adviser to the courts by assuming "some of the traditional functions tradition-

D. *Forensic Scientists and Scientific Thinking*

Forensic scientists are neither trained nor forced to think like "scientists." Scientists seeking to advance knowledge about the empirical world do not rely on assertions, convictions, or mere logic. Instead, they analyze and test hypotheses.[109] The process is always dependent on skepticism and doubt.[110] Skepticism promotes inquiry, experimentation, and validation that may remove that doubt.[111] The scientist looks at hypotheses and determines how they can be proved wrong by reproducible experiments.[112] Only after the possible shortcomings of an assertion are tested and the experiments fail to discount the hypothesis does a new, accepted scientific thesis emerge.[113]

---

ally performed by publications and peer reviews."

109. Sir Karl Popper, the distinguished philosopher of science, has explained:

> A scientist ... puts forward statements, or systems of statements, and tests them step by step. In the field of the empirical sciences, ... he constructs hypotheses, or systems of theories and tests them against experience by observations and experiment.

K. POPPER, THE LOGIC OF SCIENTIFIC DISCOVERY 27 (1968). *See also* Black, *A Unified Theory of Scientific Evidence*, 56 FORDHAM L. REV. 595, 622 (1988), (quoting *Afterword* to THE STRUCTURE OF SCIENTIFIC THEORIES at 650 (F. Suppe 2d ed. 1977)).

110. "Of course, there is no *one* scientific method. The techniques used to develop and test hypotheses necessarily vary for different disciplines." Note, *The Scientific Model in Law*, 75 GEO. L.J. 1967, 1970 n.16 (1987) (emphasis in original).

111. *Cf.* Note, *Different Standards and Conflicting Results: A Re-Evaluation of the Test for Admitting Novel Scientific Evidence in Light of Decisions Involving Spectrographic Evidence Introduction*, 5 REV. LITIGATION 327, 364 (1986) ("Science thrives on doubt. Scientific skepticism promotes inquiry and experimentation, and, hence, advances our understanding of the world around us.").

112. Sir Karl Popper states:

> Theories are ... *never* empirically verifiable.... [But a system is] empirical or scientific only if it is capable of being *tested* by experience. These considerations suggest not the *verifiability* but the *falsifiability* of a system is to be taken as [the] criterion of demarcation.... *[I]t must be possible for an empirical scientific system to be refuted by experience.*

K. POPPER, *supra* note 109, at 40–41 (emphasis in original). Popper points out that experiments must be reproducible to be part of scientific proof. *See also* Walls, *supra* note 72, at 187 (explaining that scientific hypotheses can only be disproved).

113. The emergence of a thesis has been compared to the process of evolution:

> [W]hat characterizes the empirical method is its manner of exposing to falsification, in every conceivable way, the system to be tested. Its aim is not save the lives of untenable systems but, on the contrary, to select the one which is by comparison the fittest, by exposing them all to the fiercest struggle for survival.

K. POPPER, *supra* note 109, at 42. *See also* Goldberg, *supra* note 100, at 1342–43; Black, *supra* note 109, at 623. *Cf.* Monahan & Walker, *supra* note 104, at 502 (referring to con-

The good researcher must either have or develop a probing, skeptical mind that can design experiments to test scientific assertions. Most forensic scientists have not been placed into a crucible that is likely to forge that kind of analytical thinking. They are unlikely to have had the opportunity to learn how to design experiments which answer rigorous questions. Even if they have the abilities, most forensic scientists have not been educated to ask and answer probing questions; they have not been taught to develop skepticism and curiosity; they have not been trained to do experiments.[114] As a consequence, the research that they do is frequently of low quality.[115] Forensic scientists often lack the capability to do rigorous research validating their analytical techniques.[116] And techniques without thorough validation can only produce suspect results.

---

sensus among social scientists that research methods must justify the investigator's conclusions). One court rejected the testimony of a graphologist who formed opinions about personality and mental states from handwriting. The court concluded, "[G]raphology is not a science because its results are neither verifiable nor repeatable." State v. Davis, 154 Ariz. 370, 375, 742 P.2d 1356, 1361 (1987).

114. *Cf.* Grunbaum, *supra* note 76, at 13 ("There is a general uncritical belief that acceptance of a method within the community of crime laboratory analysts is proof of its reliability.").

115. *See* Shapiro & Charrow, *Scientific Misconduct in Investigational Drug Trials*, 312 NEW ENG. J. MED. 731, 736 (1985) (concluding that one of the chief reasons for the incompetent scientific work presented to the Food and Drug Administration is that "some investigators have neither the research training nor the inclination to undertake serious, rigorous scientific trials").

116. These patterns also raise the question of how many "forensic scientists" are "scientists." If "scientist" means one who seeks to advance knowledge about the empirical world by testing ideas in a scientifically rigorous manner, then few in forensic science are scientists. Instead, the profession is heavily peopled with those that merely apply techniques that they have been taught; it is populated with scientific technicians. *But cf.* Kirk, *The Interrelationship of Law and Science*, *supra* note 55, at 394.

Others seek to distinguish forensic technicians from those who can interpret the results of a procedure.

> By definition, the role of the criminalist includes the interpretation of physical evidence in written reports and in court testimony. Although crime laboratories may use technicians for some restricted routine analytical tasks, most forensic scientists practice at the professional level and are directly responsible for both the analytical work and the interpretation, reporting, and testimony derived from that work.

Bashinski, *supra* note 27, at 162. Merely interpreting and reporting results, however, does not make one a scientist. A person may be able to interpret a sphygmomanometer and report blood pressure. That ability does not make a scientist.

> Lawyers tend to group all technically trained persons under the generic title scientist. We have mentioned laboratory technicians who are skilled in a number of routine analyses.... At another level are scientists who make inferences from the results of those tests. For example, how reliable is the presumption that a person recently fired a handgun if a swab of the palms reveals a certain concentration of barium or antimony?

E. Lack of Empirical Validation

Forensic science is supported by almost no research.[117] The laboratory practices are based on intuitions and deductions, not on empirical proof. For example, "the forensic profession has not undertaken ... research to determine the optimal conditions in which to do their work and maximize their accuracy. One can only be astounded at the volume of research on eyewitness accuracy and the paucity of parallel work on forensic science accuracy."[118]

Fiber comparisons provide an illustration. Wilkaan Fong, a California forensic scientist, noting a three-decade absence of literature on the methodology of textile fiber evidence, set out to detail better analytical methods for fiber identifications.[119] These procedures included "[c]omparisons against fibers of similar description, but from another source, ... done through the rapid interposition of the two slides."[120] In a conclusion crucial to the validity of such a procedure, he contended, "The ability to compare the directly viewed fiber with the retained mental image of another fiber is sufficient in virtually all instances."[121]

Strong criticism followed: "The available scientific literature on successive color matches does not support Fong's ... belief that reliable color matches can be made with retained mental images.... The procedure suggested by Fong is therefore likely to introduce a systematic

---

Symposium, supra note 5, at 633 (comments of Ann Hurley). The scientist is the one who answers that reliability question not by deduction or assertion, but by trying to find alternative explanations for the presence of those substances and by subsequently examining the empirical proof about those possibilities.

117. Little basic research of any kind gets done in forensic science.

> [Forensic science] progress has been technical rather than fundamental, practical rather than theoretical, transient rather than permanent. Many persons can identify the particular weapon that fired a bullet, but few if any can state a single fundamental principle of identification of firearms. Document examiners constantly identify handwriting, but a class of beginners studying under these same persons, would find it difficult indeed to distinguish the basic principles used. In short, there exists in the field of criminalistics a serious deficiency in basic theory and principles, as contrasted with the large assortment of effective technical procedures.

Kirk, *The Ontogeny of Criminalistics, supra* note 55, at 235. *See also* Koblinsky & Sheehan, *The Desirability of a Ph.D. Program in Forensic Science*, 29 J. FORENSIC SCI. 706, 707 (1984) (stressing the unrecognized need for forensically oriented basic research).

118. Risinger, Denbeaux & Saks, *supra* note 20, at 775 n.189.

119. Fong, *Fiber Evidence: Laboratory Methods and Observations from Casework*, 29 J. FORENSIC SCI. 55, 55 (1984).

120. *Id.* at 58.

121. *Id.* at 97.

bias into color comparisons."[122] Fong responded by questioning the relevancy of the other scientific literature and concluded, "Whether or not a comparison microscope is an essential need is to be left to the individual worker as determined by his or her work situation. However, no one can deny that an essential need in the practice of the criminalistics art is *wisdom*."[123]

The response, however, misses the point. Wisdom may be valuable in science, just as wisdom may be valuable elsewhere, but wisdom has little use without a scientific answer to the question. Science requires that skeptical, probing questions be met with testable answers verified in a reproducible manner. Only empirical work can provide an answer to a scientific dispute. Fong did not supply the necessary empiricism; he just made an assertion. That is not science.

Fiber comparisons comprise only a tiny segment of the forensic sciences. Fong's dispute with his critics would be insignificant if it did not typify a frequent pattern in forensic science. Forensic science is not an abstract, theoretical undertaking. It is a practical, real life endeavor which purports to employ the workings of the empirical world.[124] No matter how intelligent the forensic scientist, he cannot prove that a forensic science technique works as claimed with nothing more than a clever explanation. He can only establish efficacy through empirical validation. The electrician can give a perfectly sensible explication of why her wiring job is perfect, but the homeowner should be satisfied only after turning on the lights and verifying that everything works as claimed. The same should be true for forensic science,[125] but the basic maxim of

---

122. Rowe, *Discussion of "Fiber Evidence: Laboratory Methods and Observations from Casework,"* 30 J. FORENSIC SCI. 602, 602–03 (1985) (citing Newhall, Burnham & Clark, *Comparisons of Successive with Simultaneous Color Matching*, 47 J. OPTICAL SOC. AM. 43 (1957)). *See also Preliminary Report—Committee on Forensic Hair Comparison*, CRIME LAB. DIG., July, 1985, at 50, 53 ("The use of a high quality comparison microscope is mandatory when comparing the microscopic characteristics of hairs to distinguish between subtle or slight differences that may exist between hairs from different individuals.").

123. Fong, *Author's Response*, 30 J. FORENSIC SCI. 603, 603 (1985) (emphasis in original).

124. *Cf.* S. GOULD, HEN'S TEETH AND HORSES TOES 254–55 (1983) ("The final proofs of logic and mathematics flow deductively from stated premises and achieve certainty only because they are *not* about the empirical world.") (emphasis in original).

125. In fact, forensic scientists should take particular care:

> The question whether someone or some technique can do what it purports to do is fundamental, not only for handwriting identification, but for a great variety of endeavors. Can astrology predict the course of a person's life? Can a biomedical test detect the presence of a particular disorder? . . . The answers to these and similar questions are generally pursued by conducting empirical studies to evaluate the extent to which the claims are fulfilled. . . . These kinds of empirical evaluations enable us to separate the more effective from the less effective techniques and the

science often gets ignored in forensic science, where claims without empirical validation are frequently presented as fact.[126]

Identification by foot impressions provides another example. Footprints with details like fingerprints—friction ridges, scars, and wrinkle patterns—may be as unique as fingerprints.[127] But forensic scientists have claimed that foot impressions without such delineations are also unique. Dr. L. M. Robbins, writing in 1978, asserted that "[t]he information on footprint (and foot) morphology is especially significant because it elucidates the individuality of each person's footprints.... The combined effects of heredity and life experiences are operative in determining the size and shape of our feet; and for each of us, the manifestation of those effects is uniquely our own."[128] That assertion, however, has not been validated by scientific experimentation. As two forensic scientists recently noted, "There is ... little or no published data that support

---

> valid from the invalid theories.... There is no other way to determine which is which.

Risinger, Denbeaux & Saks, *supra* note 20, at 735–37.

126. *See* Giannelli, *supra* note 9, at 1238 ("Overstatements by experts about the conclusions that can be drawn from various scientific techniques are not uncommon."). *See also* Starrs, *supra* note 7, at 289 (describing many unsupported opinions by forensic scientists.) Edward T. Blake, a forensic scientist associated with a DNA typing lab, has noted:

> We should be quick to debunk the untenable claims for DNA technology. For example, we have heard claims of [DNA] analysis on vaginal swabs collected more than 24 hours postcoitus.... We know from routine casework that 24-hour vaginal swabs do not ... provide the requisite amount of DNA. These uncritical claims damage the credibility of DNA technology.

Blake, *Scientific and Legal Issues Raised by DNA Analysis* in BANBURY REPORT 32: DNA TECHNOLOGY AND FORENSIC SCIENCE 115 (J. Ballantyne, G. Sensabaugh & J. Witkowski eds. 1989).

127. *See* Grubb & Morton, *Discussion of "The Individuality of Human Footprints,"* 25 J. FORENSIC SCI. 271, 271 (1980); *see also* Commonwealth v. Bartolini, 299 Mass. 503, 13 N.E.2d 382, *cert. denied*, 304 U.S. 365 (1938). The Court noted:

> There was no error in permitting the witness Wentworth to testify that [bloody] footprints of a naked foot on the linoleum floor of the bathroom at the house of the deceased were made by the same person who had made prints at the police station identified as those of the defendant.... There was also ample evidence that footprints, like finger prints, remain constant throughout life and furnish an adequate and reliable means of identification.

Id. at 513, 13 N.E.2d at 388.

128. Robbins, *The Individuality of Human Footprints*, 23 J. FORENSIC SCI. 778, 784 (1978). *See also* Laskowski & Kyle, *Barefoot Impressions—A Preliminary Study of Identification Characteristics and Population Frequency of Their Morphological Features*, 33 J. FORENSIC SCI. 378, 379 (1988) (listing similar assertions from other sources).

these conclusions. These statements have either gone unnoticed, or have been accepted without equivocation by a large part of the forensic science community."[129]

Bite mark evidence provides yet another illustration of untested assertions. Although courts have admitted such evidence for decades,[130] they have not required forensic odontologists to demonstrate empirically the validity of their proposition that bite marks are unique. For example, in the Missouri case of *State v. Sager*,[131] a dentist, after examining photographs of a bite mark on the victim's breast and a cast of the defendant's teeth, concluded, beyond a reasonable doubt, that the defendant had made the bite mark.[132] The appellate court, after reviewing the trial court record and the literature about bite mark identifications, concluded "that forensic odontology, inclusive of bite mark identification, is an exact science. It is exact in the sense that through acceptable scientific procedures, an expert can form an opinion useful to the courts in their quest for the truth."[133] Perhaps the odontologist can give a useful opinion for the courts, but that does not make the opinion scientific. Science is not the acceptance of assertions that seem to make sense; science is the search for empirical data that confirms or disproves the assertions.

The *Sager* court did not point to any studies that validated the assertion that bite marks could be used to make unique identifications.

---

129. Laskowski & Kyle, *supra* note 128, at 379. Grubb and Morton assert:

> Dr. Robbins offers the idea that *in the absence of the individualizing detail of this type* footprints are unique with respect to their morphology, but she leaves us without the pertinent information to accept the idea.... [Before accepting the validity of Dr. Robbins' idea, we need to answer important questions such as] how much variation can occur between consecutive footprints of an individual, and ... how much variation occurs among many different individuals [with] approximately the same foot size.... How much variation is allowed before one can form an opinion of exclusion? ... How closely do two impressions have to correspond before they can be identified as originating from the same individual?

Grubb & Morton, *supra* note 127, at 271. Laskowski and Kyle do provide some empirical data, but as the title of their article indicates, their work can be seen as preliminary at best. For example, they studied footprint impressions from 107 adult men made in a wooden tray containing a mixture of fine clay, fingerprint powder, and bathroom scouring powder. How would conclusions reached from such research translate into the actual forensic practices of impressions left in the soft soil of a garden bed or the marks left by bloody feet? No published study appears to give answers. Despite this, defendants often do not challenge this evidence. *See, e.g.*, People v. Farmer, 47 Cal.3d 888, 254 Cal. Rptr. 508, 765 P.2d 940 (1989) (defendant did not dispute the identification of his footprints).

130. *See* Wilkinson & Gerughty, *supra* note 56, at 529–35 (discussing cases which have admitted bite mark evidence).

131. 600 S.W.2d 541 (Mo. Ct. App. 1980).

132. *See id.* at 563.

133. *Id.* at 569.

Indeed, almost no such research exists. Forensic scientists writing after *Sager* have concluded, "Few experimental studies have been carried out to determine the reliability of comparative techniques of bite marks in food or skin.... [T]here is no generally accepted approach to the evaluation of bite marks...."[134] There is no science of bite mark analysis.[135]

While fiber, bite mark, and footprint analyses may seem to be on the fringe of the forensic sciences, the patterns they represent are not.[136] Conclusions are frequently presented by people claiming to be scientists,

---

134. Rawson, Vale, Sperber, Herschaft & Yfantis, *Reliability of the Scoring System of the American Board of Forensic Odontology for Human Bite Marks*, 31 J. FORENSIC SCI. 1235, 1236 (1986). *See also* Rawson, Ommen, Kinard, Johns & Yfantis, *Statistical Evidence for the Individuality of Human Dentition*, 29 J. FORENSIC SCI. 245, 245 (1984); Wilkinson & Gerughty, *supra* note 56, at 536–37 (noting the failure of bite mark studies to address the question of specificity of bite mark analysis).

135. While a group of scientists has attempted to fill the research gap, they could only conclude, "The scoring guide evaluated here is the *beginning* of a truly scientific approach to bite mark analysis." Rawson, Vale, Sperber, Herschaft & Yfantis, *supra* note 134, at 1259. *See also* Wilkinson & Gerughty, *supra* note 56, at 547.

> [T]hat a standard technique for analyzing bite marks has not been established speaks strongly for the proposition that bite mark analysis remains in its incipient stages of development.... [T]here also does not appear to be any established plan or procedure for advancing forensic bite mark knowledge along the scientific lines necessary for testing the validity of the fundamental hypothesis underlying the principle of bite mark analysis, namely, its uniqueness.... Exhaustive research has failed to reveal any published data concerning the classification of bite mark characteristics on any segment of the population.

Id.

136. Human hair comparisons provide yet another area where assertions have been presented without research. *See* Barnett & Ogle, *Probabilities and Human Hair Comparison*, 27 J. FORENSIC SCI. 272, 272 (1982) (reviewing the literature on hair comparisons).

> Beginning in 1974, [two forensic scientists] published a series of papers concerning the probabilities of human hair comparison. Their conclusions are cited by expert witnesses when asked about the certainty of a hair comparison. Nonetheless, in the seven years that have elapsed since the publication of the first of these articles, there has been no attempt reported in the literature to confirm [the] work or criticize [the] treatment of [the] data.

Id., but see Gaudette, *A Supplementary Discussion of Probabilities and Human Hair Comparisons*, 27 J. FORENSIC SCI. 279, 279 (1982) (responding that Barnett and Ogle do not understand well the "underlying principles and concepts" of the work on hair comparisons).

Writing more recently, Aitken and Robertson conclude, "The benefit of following the [published] method of evaluating hair evidence ... would be that it would encourage the use of standardized nomenclature and a uniform approach to hair examination [that] should improve the overall ability of hair examiners to discriminate hairs. However, ... [the necessary] data base lies some way in the future, and to date, there has not been the willingness on the part of funding agencies to support such a scheme." Aitken & Robertson, *A Contribution to the Discussion of Probabilities and Human Hair Comparisons*, 32 but who have neither analyzed the basis of their own assertions nor undertaken experiments to test those hypotheses.[137] When they do test their claims, they too often do it only after having already relied upon a new technique. Research patterns indicate that testing is frequently motivated not by the forensic scientists' scientific impulses but because of probing questions about the accuracy and reliability of forensic procedures raised by those outside of the field.[138] Voiceprints, identification of protein markers in bloodstains, and the recent DNA "fingerprinting" all illustrate this pattern.

The high speed sound spectrograph is a device for the study of acoustics. Some researchers, believing that every person's voice patterns are unique, concluded that the spectrograph could be used forensically to make "a graphic representation of the frequency spectrum of a segment

---

J. FORENSIC SCI. 684, 689 (1987). *See also*, Note, *Splitting Hairs in Criminal Trials: Admissibility of Hair Comparison Probability Estimates*, 1984 ARIZ. ST. L.J. 521, 530 (reviewing the scientific research and concluding that "[t]here are no available studies that have determined a forensically significant level of dissimilarity between hairs").

137. Sometimes even when a forensic scientist has collected data that would be useful to forensic science, the scientist does not analyze and present it. For example, an FBI review article on gunshot residue ("GSR") explains that tests for GSR examine hands for traces of barium and antimony. The article notes that this test is not foolproof because hands may have traces of these elements without having fired a gun, and people can fire guns without leaving barium and antimony on their hands. *See also* Kilty, *A Review of the FBI Laboratory's Gunshot Primer Residue Program*, CRIME LABORATORY DIG., April, 1986, at 54, 57. The article concludes that there can be several explanations for the lack of antimony and barium deposits on the hand. The FBI in its test firings surely has generated data that sheds light on how often a gun is fired without leaving gunshot residue, but even so, the FBI does not present that information. It leaves completely speculative the question of how exculpatory it might be for a suspect to have hands free of these elements when the criminal has fired a gun. *See* Reed, McGuire & Boehm, *Analysis of Gunshot Residue Test Results in 112 Suicides*, 35 J. FORENSIC SCI. 62 (1990) (GSR tests in suicides by firearms only found residues consistent with firing a gun 38% of the time). *See also* Stone, *Capabilities of Modern Forensic Laboratories*, 25 WM. & MARY L. REV. 659, 666 (1984) ("Over the last ten years in Dallas, the forensic laboratory has reported positive evidence of gunshot residue metals in about forty percent of all suicides with handguns. In analyzing handwipings from live suspects, fewer than ten percent revealed positive evidence of residue.").

This data makes the testimony reported by Saks and Van Duizend interesting. The prosecution produced a forensic expert to prove that defendant's companion had not fired a gun. He said that he found no gunshot residue on the companion's hand, "that this was indicative of his not having fired a gun . . . [, and] that residue is found through this test in 75 to 80 percent of the cases." M. SAKS & R. VAN DUIZAND, *supra* note 25, at 41.

138. For example, many probing questions about bite mark identifications seemed to have been raised first in an influential law student article written after People v. Marx, 54 Cal. App. 3d 100, 126 Cal. Rptr. 350 (1975) (admitting bite mark evidence). Note, *The Admissibility of Bite Mark Evidence*, 51 S. CAL. L. REV. 309 (1978).

of recorded speech."[139] The resulting "voiceprint" could absolutely identify a voice, it was declared. With some research undertaken, the advocates of voiceprints began to use the spectrograph forensically. A number of courts allowed this evidence to be used in criminal cases.[140]

The forensic science community did not skeptically examine the claims made for voiceprints. A scientific controversy erupted only when scientists in other fields questioned the adequacy of the spectrograph research and produced studies indicating much less accuracy than that asserted by the proponents.[141] Meanwhile, voiceprint identifications continued.

In response to this debate, the Federal Bureau of Investigation asked the National Academy of Sciences to review voiceprint technology. A Committee on Evaluation of Sound Spectrograms analyzed the research and concluded that the underlying theory of voiceprints had not been validated and that existing data did not support the proponents' claims of high accuracy.[142]

Even after the report, many courts continued to allow the introduction of such evidence.[143] Nevertheless, one group reported that as of 1985

> no further respected scientific studies have been published which alter or invalidate the findings of the 1979 National Academy of Sciences study. There still is no scientific data that proves the fundamental premise upon which the technique is based, namely: voice uniqueness. Nor is there available authoritative data on the effect on voice spectrograms of nasal or oral surgical operations, muffling of the voice, mimicking [and so on]. Such data would appear to be a necessary prerequisite to scientific acceptance based on proven reliability.[144]

Indeed, only in 1986 did a forensic scientist, in a study with important

---

139. Note, *supra* note 111, at 333.

140. Not all courts, however, admitted the voiceprint evidence. *See* A. MOENSSENS, F. INBAU & J. STARRS, *supra* note 5, at 665–69 (surveying cases).

141. The history of the scientific research is presented in Note, *supra* note 111, at 332–40.

142. Committee on Evaluation of Speech Spectrograms, *On the Theory and Practice of Voice Identification* (1979). The report is summarized in A. MOENSSENS, F. INBAU & J. STARRS, *supra* note 5, at 663–64. *See also* Note, *supra* note 111, at 339–40.

143. *See* A. MOENSSENS, F. INBAU & J. STARRS, *supra* note 5, at 669 ("[E]arlier court trends, some going away from admissibility, others going towards it, continued").

144. *Id.* at 684.

limitations,[145] produce research to fill any of these gaps.[146] Even with his work, the scientist could still only conclude that the future of forensic spectrographic analysis remained unsettled.[147]

Another example of this research pattern can be found in the area of conventional genetic marker analysis. In the late 1960's, forensic scientists began to use electrophoresis, a procedure with many scientific applications, to type various proteins in bloodstains. By the late 1970s, the technique was widely used,[148] and in 1981 the first appellate court to consider the issue upheld the admissibility of this form of blood typing.[149]

The next year, however, a law review article by this author surveyed

---

145. The researcher issued an important caveat about his study: "[T]he error rates were determined from direct feedback from field investigators which may not always be correct. . . ." Koenig, *supra* note 56, at 115. Indeed, his method of determining voiceprint mistakes hardly seems scientific:

> After each [FBI] examination . . . , a written report of findings is mailed to the contributor. . . . If an identification or elimination is made, the contributor is contacted by telephone and asked if the results are consistent with interview and other evidence in the investigation. If other information strongly supports the voice comparison result, the contributor is told to contact the FBI if later developed evidence contradicts the finding. If the voice comparison results contradict the other evidence, the matter is closely followed until legally adjudicated or investigatively closed. In the few occurrences where no final determination was possible, the voice comparison result was considered a "no decision."

Koenig, *Spectrographic Voice Identification: A Forensic Survey*, 79 J. ACOUSTICAL SOC. AM. 2088, 2089 (1986).

Koenig also noted that "the educational and training requirements of the FBI examiners, the criteria for quality and quantity of comparable words, and the sophisticated scientific procedures available to the FBI exceed the standards normally used and/or available to most other practitioners in the field." Koenig, *supra* note 56, at 115.

146. *See* Koenig, *Spectrographic Voice Identification*, *supra* note 145, at 2089.

147.

> [The future for forensic spectrographic comparisons will] probably lead in one of two directions: either to providing testimony using the FBI standards and the admission in court of a realistic error rate of approximately 1% . . . or to severely limiting testimony and using the results principally as an investigative aid for law enforcement.

Koenig, *supra* note 56, at 117.

148. "[Protein gel electrophoresis] was used in thousands of cases in the late 1970's and was accepted by courts in several states." Note, *supra* note 50, at 510 n.273. *See also* BUDOWLE & MURCH, ELECTROPHORESIS RELIABILITY. II. HISTORICAL USE IN THE FBI LABORATORY, VALIDITY, AND SCIENTIFIC SCRUTINY (unpublished manuscript) (history of electrophoresis development in FBI Laboratory).

149. State v. Washington, 229 Kan. 47, 622 P.2d 986 (1981).

the scientific literature on bloodstain genetic marker analysis.[150] It found that almost no research had been published demonstrating that such analysis could be done reliably on the kinds of bloodstains actually encountered in forensic work—bloodstains that are aged, degraded, or contaminated. Further, some of the sparse research actually indicated the likelihood of producing inaccurate typing results on such samples.[151]

In 1986, the Michigan Supreme Court held electrophoretic protein typing inadmissible because the technique's reliability on real crime samples had not been established:

> Reliability remains in dispute and unresolved because of the questions unanswered. The questions are not likely to be answered and the reliability of electrophoresis of evidentiary bloodstains established until independently conducted validation studies on the thin-gel multisystem analysis are undertaken and comprehensive control tests evaluating the effects of different contaminants are run, and the results have been subjected to the scrutiny of the scientific community. The evidence produced by electrophoresis should, therefore, not have been admitted.[152]

Only after these events did forensic scientists produce and publish research attempting to prove that they could accurately type the proteins in the kinds of bloodstains associated with real crimes. Even if that research could answer the previously raised questions about the forensic application of the technique,[153] the unscientific pattern is important.

---

150. Jonakait, *Will Blood Tell? Genetic Markers in Criminal Cases*, 31 EMORY L. REV. 833 (1982).

151. *See id.* at 875–912.

152. People v. Young, 425 Mich. 470, 476, 391 N.W.2d 270, 272 (1986).

153. *See* Wraxall & Stolorow, *The Simultaneous Separation of the Enzymes Glyoxalase I, Esterase D, and Phosphoglucomutase*, 31 J. FORENSIC SCI. 1439 (1986). This "Technical Note" was criticized by FBI scientists:

> The majority of the space in their paper is devoted to reporting work that was completed before 1978.... It appears that the paper was submitted and published only in an attempt to satisfy criticism which has been leveled at the authors by certain factions. One of the criticisms is that the method was never properly subjected to the scrutiny of the forensic science community through its publication in a reputable scientific journal.

Murch, Kearney & Budowle, *Discussion of "The Simultaneous Separation of the Enzymes Glyoxalase I, Esterase D, and Phosphoglucomutase,"* 32 J. FORENSIC SCI. 1498, 1498 (1987). After further analysis of the article, the authors concluded that "the paper by Wraxall and Stolorow did little to strike at the heart of the criticisms confronting its use. The paper did not support their scientific claims...." *Id.* at 1499. Wraxall and Stolorow responded:

Conclusions preceded research. Untested assertions remained accepted until they were challenged from outside forensic science.[154]

The rapidly unfolding DNA profiling saga also fits this pattern. DNA evidence has been used in numerous cases[155] even though many commentators have pointed out that almost no research exists to support the reliability of the technique on forensic samples.[156] Only now, after forensic scientists have attested to reliability many times in court, has some of this necessary research started to appear.[157] Once again, the

---

> We are gratified to see that the FBI used and taught the Group I method from 1979 to 1986 and found the method to be a reliable and valid electrophoretic approach for the analysis of serological materials. We note that the FBI has accumulated a large amount of data on the validation of the Group I system and would encourage them to contribute this data in published form which, by their own account, is long overdue.

Wraxall & Stolorow, *Author's Response,* 32 J. FORENSIC SCI. 1500, 1500 (1987). Black summarizes these exchanges and concludes, "To some extent, even the developers of the multi-system test have acknowledged the need for more data to establish its validity and reliability." Black, *supra* note 109, at 634 n.201.

154. *See* Budowle & Allen, *Electrophoresis Reliability: 1. The Contaminant Issue,* 32 J. FORENSIC SCI. 1537 (1987) (written in response to *People v. Young).* The article attempts to establish that contamination of forensic samples is not a problem for protein marker typing because the effects of contamination can be predicted and recognized by a competent analyst.

155. Note, *supra* note 50, at 466. ("Forensic DNA testing has been employed in hundreds of criminal cases nationwide.").

156. *See, e.g.,* Comment, *DNA Printing: The Unexamined "Witness" in Criminal Trials,* 77 CALIF. L. REV. 665, 671–72 (1989).

> Government scientists, primarily those of the British Home Office, have published a few studies of DNA printing under forensic conditions. In 1988, however, the head of the [FBI's] serology unit admitted: "We've only done a limited number of tests yet. So we can't say with absolute certainty that [the technique] works on different sizes of stains, aged stains, [or] putrefied stains."
>
> Scientists from Lifecodes Corporation . . . have published articles discussing trials of their technique under modest simulations of forensic conditions. . . .
>
> The other two commercial laboratories have offered even less information about the reliability of their methods of forensic testing. . . . Independent verification of forensic DNA printing's reliability is clearly necessary.

Id. See also Thompson & Ford, *DNA Typing: Acceptance and Weight of the New Genetic Identification Tests,* 75 VA. L. REV. 45, 107 (1989) ("To meet a more rigorous application of *Frye,* . . . proponents of the test will need to make a better showing of reliability on forensic samples"); Note, *supra* note 50, at 482–83 ("Not enough research has been done on contaminated samples at this point to understand exactly how contamination affects the results.").

157. *See, e.g.,* McNally, Shaler, Baird, Balazs, DeForest & Koblinsky, *Evaluation of Deoxyribonucleic Acid (DNA) Isolated from Human Bloodstains Exposed to Ultraviolet Light, Heat, Humidity, and Soil Contamination,* 34 J. FORENSIC SCI. 1059, 1059 (1989); McNally, Shaler, Baird, Balazs, Koblinsky & DeForest, *The Effects of Environment and Substrata on Deoxyribonucleic Acid (DNA): The Use of Casework Samples from New York City,* 34 J. FORENSIC SCI. 1070, 1070 (1989) (noting the deficiency of published reports

point is not whether this research answers the critical questions,[158] but whether the conclusions given by forensic scientists are supported by accepted research.[159]

---

addressing the environmental and substrata effects on DNA). These two papers were the product of the Lifecodes Corporation. Scientists from that company had testified many times to the reliability of DNA typing before these studies were done.

Certainly all the needed research has not been produced. *See, e.g.*, Neufeld & Colman, *When Science Takes the Witness Stand*, SCI. AM. May, 1990, at 46.

> Two years ago, when DNA evidence was first introduced in U.S. courtrooms, most forensic DNA scientists rejected the existence of band shifting. But now some experts think band shifting occurs in perhaps 30 percent of forensic DNA tets. There are now many theories about the cause, but as of this writing not one refereed article on the subject has been published.

Id. at 51.

158. A news article written after some of the recent research concludes:

> Leading molecular biologists say [that DNA profiling] is too unreliable to be used in court.... Many say they are leery enough of the method that they would not allow their DNA fingerprints to be taken if they were innocent suspects in a criminal case.
>
> The problem, the scientists say, is that the DNA fingerprints can stretch and shift, like a design printed on rubber, making them difficult, if not impossible, to interpret. They say that even without these shifts, DNA patterns can be almost impossible to compare. And they say the scientific underpinnings have not been established for determining just how unlikely it would be for DNA fingerprints from two people to match by accident.

Kolata, *Some Scientists Doubt the Value of "Genetic Fingerprint" Evidence*, N.Y. Times, Jan. 29, 1990, at 1.

159. The FBI has indicated that they will continue to follow this pattern. "John Hicks, the assistant director of the laboratory division at the F.B.I, said his unit had brought DNA evidence into court 35 times since December 1988." *Id.* at 1, 18. *Cf.* Note, *supra* note 50, at 471 n.31 ("[T]he FBI has conducted over 500 DNA tests and already has a backlog of 150 cases."). Earlier at a conference on the forensic applications of DNA profiling, Hicks stated that the FBI has

> tried to address the questions that we know are of concern to the courts. But we are going to [keep] moving forward with the technology without having it in publication. The papers are being prepared and will be published; but knowing the publication process, it will be months from now, a year perhaps, before the studies that we have carried out will be in print. In the meantime, we are going to be doing the work.

BANBURY REPORT 32: DNA TECHNOLOGY AND FORENSIC SCIENCE 101 (J. Ballantyne, G. Sensabaugh & J. Witkowski eds. 1989) [hereinafter BANBURY REPORT] (comments of John W. Hicks).

The willingness to present untested assertions as established fact is not limited to those forensic experts who work in crime labs. *See, e.g.*, Levy, *Using "Scientific" Testimony to Prove Child Sexual Abuse*, 23 FAM. L.Q. 383 (1989). Levy discusses how a California psychiatrist, Roland Summit, formulated the "child sexual abuse accommodation syndrome" not as a test of whether abuse had occurred, but to improve therapy. The formulation was based on impressionistic findings, not on empirical data. *Id.* at 393–95. Even so, These examples covering diverse areas of forensic science—from fiber comparisons to DNA profiling, from bite marks to voiceprints—indicate a widespread pattern of poor research practices. Forensic scientists present opinions and conclusions without research. They fail to test the accuracy and reliability of their work until questions are raised by others.

These practices raise concerns about the quality of the forensic sciences. Certainly, the results of procedures and techniques supported only by the assertions or deductions of some forensic scientists, rather than by rigorous research, are suspect. But the research patterns raise a more fundamental concern. They indicate that substandard scientific thinking pervades these fields. Since the lack of scientific rigor is so widespread, even forensic science research which does occur may not be done adequately. Techniques and procedures based on inferior research should be as suspect as those based on no research. Even though inferior scientific techniques may not exist in all areas of the forensic sciences, enough substandard practices exist to suggest severe quality problems for crime laboratories.

F. The Special Needs of Forensic Science Research

Forensic science requires good research perhaps more than any other area of scientific inquiry. Unlike many other endeavors, forensic science is often unable to adopt scientific knowledge and techniques from related areas. Frequently, there is no related discipline to draw upon. Only the forensic scientist is interested in studying for identification purposes such things as bullets, toolmarks, and fingerprints.

In addition, even when a crime lab field relates to other scientific areas, the work done by forensic scientists is fundamentally different from work done in the other areas. Their mission is not to identify a sample in the way other scientists might, but instead to distinguish the sample from other specimens.[160] Crime scene samples are often unique,

---

experts now rely on the syndrome in their testimony to establish that abuse took place. *Id.*

160. "The real aim of forensic science is to establish individuality, or to approach it as closely as the present state of science allows. Criminalistics is the science of individualization." Kirk, *supra* note 117, at 236. "Above all, the criminalist is a specialist in identification and determination of the source of evidence items." Kirk, *supra* note 116, at 394. *Cf.* Comment, *supra* note 156, at 667 ("All forensic identification methods aim to isolate physical qualities or traits that vary among individuals. The value of any particular method lies in its ability to individualize a subject.").

The forensic scientist does not just try to catalog the components or the characteristics of an object, but instead tries to determine if two objects had a common origin. This makes forensic science identifications unique. *See also* Barnett & Ogle, *supra* note 136, at 272: "Most of the evidence examined by criminalists falls into the category of associative evidence, that is, evidence that tends to associate one person, place or thing with another per-of unknown origin and unascertainable history. In contrast, most scientific researchers know what they are working with and its history. Whereas most scientists would not work with samples that have been dropped on the floor, the samples for a DNA identification quite literally may have been scraped off the linoleum.[161] It cannot be assumed that research on pristine, known samples applies to the contaminated, degraded, aged samples of unknown history confronted by the forensic scientist. Validation studies must be done specifically on samples approximating forensic condition. This issue, which has been raised with respect to DNA profiling[162] and blood tests,[163] is relevant to many other areas of forensic science as well.[164] It is only conjectural to

---

son, place or thing." Since the goal of the work is different from the work of other scientists, the procedures and knowledge from those other areas can not always be simply transferred to forensic science. "Common origin determination is often what is sought, and this too sets the crime lab apart from other types of testing laboratories. No other proficiency testing program concerns itself with the possibility of common sources of test samples. These different approaches do not lend themselves to the type of testing that is carried out by most other types of 'testing' laboratories." LEAA, *supra* note 1, at 26.

161. Biochemist Richard Roberts of the Cold Spring Harbor Laboratory explained that:

> samples in the lab are sterile; if a sample fell on the floor, he would throw it out and start over, "but in forensics, all of the samples have been on the floor, so you don't really know what you have got."

Note, *supra* note 50, at 481 (quoting Cooke, *Standards Sought for DNA Testing*, Newsday, June 15, 1989, at 27). Significantly, an expert in medical laboratory testing indicated that the main reason for analytical inaccuracies in medical laboratories is "[s]pecimen contamination or deterioration." *Hearings, supra* note 20, at 356 (statement of Mark S. Birenbaum, American Association of Bioanalysts).

162. Comment, *supra* note 156, at 670–71 (comparing forensic DNA analysis unfavorably with clinical DNA analysis).

163. *See, e.g.*, Jonakait, *supra* note 150, at 841 (stressing that the forensic serologist works with dried blood which is obviously different than the fresh blood with which other scientists normally work).

164. The issue of whether a scientific technique can be validly applied to forensic situations is a frequent one.

> For many applications of science to forensics, the underlying theory is well established, and legal debate rages mainly over whether one must prove only that a technique is generally accepted for scientific research or, more strictly, that the technique is reliable when applied to forensics.
>
> Why the distinction between nonforensic and forensic applications? Scientists commonly accept that when any technology is tried in a different application, such as forensics, it must be tested thoroughly to ensure an empirical understanding of the technique's usefulness and limitations. Indeed, many a technique that has proved reliable for research . . . has turned out to be of questionable reliability when applied to forensic casework.

Neufeld & Colman, *supra* note 157, at 49.

Bite marks provide yet another example. While it may be that each set of teeth is unique, it does not necessarily follow that uniqueness can be established in the materials conclude that the techniques and tools developed for related disciplines can be accurately applied to crime lab work without first validating those scientific procedures under forensic conditions.[165] The assumption that a method proved valid in one area can be applied in crime laboratories has frequently led to faulty results.[166]

---

with which a forensic scientist has to work.

> While dental chart analysis allows comparison of inlay substances, root abnormalities, and other factors not directly related to anatomical configuration, bite mark analysis relies solely on the comparison of anatomical features. Bite mark analysis thus rests on two additional postulates: first, that the anatomical configuration of every mouth is unique, and second, that this asserted uniqueness is graphically displayed in the bite mark.

Note, *supra* note 138, at 323. *See also* Rawson, Ommen, Kinard, Johns & Yfantis, *supra* note 134, at 252 ("This mathematical evaluation of a general population sample demonstrates the uniqueness of the human dentition beyond any reasonable doubt.... The real concern now is the determination of the match between the dentition and the impression or bruising of the skin or other material.").

165. *Cf.* Frank & Gunn, *Standardization of Forensic Drug Analyses*, 19 J. FORENSIC SCI. 163, 163–65 (1974) (comparing the work of inspectors for the Food and Drug Administration and those doing forensic drug analyses).

> [FDA] evidence is usually in a bottle or container labeled with the manufacturer's name, active drug components, etc. The composition of the material is known because the product must meet FDA standards....
>
> To determine whether or not a drug meets the requirements, it is subjected to physical and chemical tests that have been established to specifically demonstrate that the substance is, or is not, the same as that identified on the label of its container. Since the composition of the material is on file, specific tests, which have been set up for the particular product and subjected to scientific scrutiny, can be used for the examination....
>
> [On the other hand, forensic drug] evidence is frequently obtained in undercover buys in unmarked containers; it may be a pure drug or may be cut with diluents or excipients. The illicit manufacturer is not a member of the Pharmaceutical Manufacturers Association ..., and the illicit distributor ... is not concerned with labeling his product....
>
> In short, because there is a great variance in the type of material being submitted for the [forensic drug analyst's] examination, it is necessary to use a wide range of methodology in analyzing the material.

Id.

166. *See* Strong, *Questions Affecting the Admissibility of Scientific Evidence*, 1970 U. ILL. L. F. 1, 18. *See also* Starrs, *supra* note 7, at 296 (examples of valid techniques inappropriately applied to novel areas). *Cf.* Starrs, *"A Still-Life Watercolor": Frye v. United States*, 27 J. FORENSIC SCI. 684, 686 (1982) ("The distinction between the validity of a principle and the validity of a technique applying it in concrete circumstances is crucial in determining whether expert testimony should be admitted or not."). *Cf.* Giannelli, *supra* note 9, at 1212.

> A novel forensic technique, however, may involve either the new application of a well-established theory or the application of a new theory. In the latter case, the theory can be validated only empirically or inferentially, not deductively. In other

Since a cautious forensic scientist cannot readily use research techniques from other areas, forensic science has a greater need for research than other areas. This characteristic, however, is not the only one indicating that forensic science has exceptional needs for research. Special needs arise in forensic science because the use of a technique or test will not validate or expose the flaws in those procedures as it does in other scientific and technical areas.

> [A] new scientific procedure will normally go through a three-step process. . . . Suppose, for example, that some types of blindness are caused by tiny blood vessels in the eyes wearing out and bursting. A scientist . . . [w]orking on animals . . . develops a laser surgery technique that he believes can stop the bleeding without doing damage to the rest of the eyeball. He publishes his results.
>
> Other eye specialists reading this report may be impressed by it, agree with the theory behind it, and understand how the new procedure should alleviate the condition. They would not yet, however, accept the procedure as reliable. Instead, verification through the use of controlled studies would next be required . . . to see if the procedure worked the cure as claimed; and . . . to see if any limitations, such as side-effects, could be discovered.
>
> If the new procedure passed the controlled studies, . . . [t]he ultimate test would be the actual use of the technique in general practice. Any limitations that were not unearthed in the studies would surface as the procedure was used under various circumstances. It might be eventually learned that while the surgery at first seem to be successful with diabetics, after a

---

> words, the successful application of the technique proves the validity of the underlying theory or principle.

Id. Some courts do recognize this critical issue. *Compare* State v. Huynh, 49 Wash. App. 192, 742 P.2d 160 (1987) (noting that gas chromatography was validly accepted as a method of determining the type of accelerant used in an arson and for comparing for identification purposes liquid gasoline samples. The technique, however, was used to identify a weathered and burned accelerant at the fire scene. The court held that the resulting evidence was inadmissible because of insufficient empirical validation of gas chromatography's ability to identify this kind of forensic specimen) *with* Santiago v. State, 510 A.2d 488, 489 (Del. 1986) (allowing the evidence to be admitted after concluding "that the gas chromatograph is commonly relied upon by forensic scientists throughout the country"). It may be that this result is correct, but the court did not ask the right question. The court should have asked whether the technique was reliable or relied upon for the particular forensic purpose that generated the evidence.

> year or so the blood vessels burst again causing greater problems than before. Certainly, only when a procedure has gone through these three steps—development of methodology, verification of the claimed results, and actual employment of the new technique can the community of concerned scientists know both the procedure's reliability and its limitations.[167]

The last step, however, is often absent in forensic science because the use of a forensic technique will not necessarily show its flaws. For example, the use of a hospital blood test will reveal its limitations, but the use of the seemingly related forensic method will not:

> Unlike hospitals that test undeteriorated, uncontaminated blood using established techniques designed to assure safe transfusions, crime labs have no way of knowing if they are right or wrong in finding a match based on deteriorated, contaminated blood samples examined with a technique developed by police scientists. In the hospital, patients would die if the blood test did not work, but complaints from innocent defendants cannot be distinguished from the protestations of the guilty.[168]

---

167. Jonakait, *supra* note 150, at 846–52. *Cf.* M. SAKS & R. VAN DUIZEND, *supra* note 25, at 74. Saks and Van Duizend assert: "A scientific or professional field has a powerful defense against unsubstantiated or exaggerated conclusions: replication. Professional/technical fields that apply such knowledge have a 'defense': erroneous principles will be found out because patients will not get well, planes will not fly, or chemicals will not [be] synthesized." *Id.* For further comments on replication and the peer review system, see Neufeld, *Admissibility of New or Novel Scientific Evidence in Criminal Cases* in BANBURY REPORT 32: DNA TECHNOLOGY AND FORENSIC SCIENCE 80 (J. Ballantyne, G. Sensabaugh & J. Witkowski eds. 1989); Woolf, *supra* note 103, at 82; *Symposium on Science and the Rules of Evidence*, 99 F.R.D. 188 (1983) (comments of Renee Fox); Note, *Of Reliable Science: Scientific Peer Review, Federal Regulatory Agencies, and the Courts*, 7 VA. J. NAT. RESOURCES L. 27, 31–32 (1987) (quoting Harnad, *Creative Disagreement*, 19 SCIENCES 18–20 (1979)).

168. Black, *supra* note 109, at 634. If the procedure has erroneously exculpated someone, no protest whatsoever may be heard. *See also* Jonakait, *supra* note 150, at 851.

> [T]he forensic scientist does not put his results to a scientific use that invariably show the limitations and liabilities, if any, of the procedures. When the forensic laboratory classifies a blood sample's PGM into one of its genetic variants, this classification is not merely a research step upon which that forensic scientist or others will continue to build. The scientist's job conversely, ends with that categorization rather than putting these findings to any further use. Consequently, if for some reason the procedures to detect genetic markers in dried and aged blood falsely detect PGM "1" as "2–1," this incorrect result will not necessarily be discovered by any subsequent scientific use or practice.

Id. Neufeld, *supra* note 167, at 80:

This pattern, of course, is not restricted to blood tests. It applies to many forensic methods including the ballistics procedure, fingerprint comparison, and drug analysis.

In addition, the possible effects of crime laboratory errors should compel extra research. Liberty and even lives are at stake in forensic work. While this may be true for the medical lab, crime laboratory errors can be even more significant. Medical procedures often have inherent quality checks that tend to disclose errors.[169] Moreover, absolute precision is not required for many medical tests. A test that reports a cholesterol level of 150 when the result should have been 165 is probably precise enough. Ninety percent accuracy may still mean that a medical test is performing adequately in all cases.[170]

The same would not be true for many forensic procedures. They often are not quantitative. Their function is to present a clearcut "yes or

---

> Regrettably, in most forensic science—although the initial stages of validation are generally complied with—once a new technique is developed, it is put into immediate use in crime laboratories and no further nonforensic research is conducted. The results are not carried forward in ways that would detect any limitations.

169. For example, only four bands are possible in DNA profiling for diagnostic purposes, and this limitation provides a consistency control on the work. By contrast, the probes used in forensic DNA can produce a hundred bands, many of which will occur close to, or even overlap with, each other. Note, *supra* note 50, at 484. Moreover, in clinical testing, a physician may recognize a lab result that is inconsistent with a physical examination or with the patient's history. A re-test may be ordered, and lab errors may consequently be discovered. *See* Bogdanich, *Medical Labs, Trusted as Error Free, Are Far From Infallible*, Wall St. J., Feb. 2, 1987, at. 1, col. 6 ("Sloppy lab work isn't as obviously dangerous as incompetent surgery [since] errors are often caught by a patient's physician."). *See also* Caplan, *Drug Testing in Urine*, 34 J. FORENSIC SCI. 1417, 1420–21 (1989) (commenting that incorrect forensic reports have serious ramifications because they do not come with the information normally included in a personal history).

170. Additional medical procedures make absolute accuracy unnecessary:

> The results of *clinical laboratory* procedures are provided to the patient's physician who combines this information with the patient's medical history and clinical symptoms to either make a diagnosis, prescribe treatment, or monitor a patient's health or medical treatment.... In most instances, the accuracy of *clinical laboratory* procedures is much greater than that required by the clinician who utilizes the information.

Hearings, *supra* note 20, at 354 (statement of Mark S. Birenbaum, American Association of Bioanalysts).

The consequences of a mistaken medical report, while important, will often not be as irreversible as the mistaken forensic analyses. For example, a clinical test that reports a cholesterol level at 260 when it should have read 190 may cause medicines, diet, and exercise to be prescribed, but the competent physician will continue to monitor the cholesterol level, and soon the doctor will conclude that the elevated reading was an aberration. Once a case is closed, however, forensic evidence is unlikely to be questioned further.

no" answer. Are these the defendant's fingerprints? Was the bullet fired from this gun? Does the bloodstain on the defendant's clothes match the victim's blood? Answers to such questions are either right or wrong. Ninety percent precision will not lead to a procedure that works well in all cases, but to a procedure that is wrong in one out of ten cases. Since forensic tests may require more precision than tests in other areas, forensic procedures need extra experimental development.

Moreover, poor training and education create special needs for research and validation. Few of the needed empirical studies, however, are done. As a consequence, the quality of crime lab work suffers.

III. QUALITY CONTROL PRACTICES

A. *Lack of General Quality Control Programs*

The effect of inadequate education, training, and research on reliable and accurate forensic results might be mitigated if these facilities had meaningful quality control programs—formally instituted procedures that, if followed carefully, assure the best possible results.[171] Crime labs, however, are not required to follow quality control plans.[172] Thus,

---

171. *Cf. Hearings, supra* note 20, at 225 (statement of Virginia L. Worrest, Medical Consultant) ("Quality control is a series of regularly followed procedures that assure a certain level of accuracy to a test result.").

The FBI draws a distinction between quality control and quality assurance programs:

> Although often used interchangeably, QA [quality assurance] and QC [quality control] refer to specific quality functions. The function of QA is to provide evidence that the QC function is being performed adequately. This is accomplished through the use of external QA requirements ..., such as analysis of blind and open proficiency samples and audits. The QC functions are internal activities ..., which ensure that the quality of the product ... is maintained and will meet and satisfy specified criteria.

Mudd & Presley, *Quality Control in DNA Typing: A Proposed Protocol*, 15 CRIME. LABORATORY DIG. 109, 111 (1988). *See also* Kearney, *Guidelines for a Quality Assurance Program for DNA Restriction Fragment Length Polymorphism Analysis*, 16 CRIME LABORATORY DIG. 40, 41 (1989).

172. Just as individuals do not have to be certified to work in crime laboratories, *see supra* text accompanying notes 76–84, laboratories as entities do not have to be licensed or certified to operate, and thus, no mechanism now exists to ensure that laboratories follow good quality control measures. *See* Comment, *supra* note 156, at 673–74.

> [U]nlike clinical laboratories and blood banks, forensic laboratories need no certification to operate.... A private forensic laboratory typically is not subject to government regulation, inspection, or licensing with respect to performance standards. Government-operated crime laboratories set their own standards; there are accrediting organizations, but accreditation is voluntary.

"[q]uality control in American crime laboratories remains on the honor system."[173] This system, however, has not produced widespread quality control programs.[174] Instead, forensic science has often treated the topic of quality control with hostility.[175]

Id. The American Society of Crime Laboratory Directors formed a Laboratory Accreditation Board in 1981, and in the next year accredited the forensic labs in the Illinois state system. Moenssens, *Admissibility of Scientific Evidence—An Alternative to the* Frye *Rule*, 25 WM. & MARY L. REV. 545, 561 n.64 (1984). This accreditation is a totally voluntary program, and thus, "[a]t present, forensic science is virtually unregulated—with the paradoxical result that clinical laboratories must meet higher standards to be allowed to diagnose strep throat than forensic labs must meet to put a defendant on death row." Lander, *Population Genetic Considerations in the Forensic Use of DNA Typing*, in BANBURY REPORT 32: DNA TECHNOLOGY AND FORENSIC SCIENCE 505 (J. Ballantyne, G. Sensabaugh & J. Witkowski eds. 1989).

173. Comment, *supra* note 156, at 674.

174. "[T]he application of formal [quality assurance and quality control] programs to forensic laboratories is not widespread. . . ." Mudd & Presley, *supra* note 171, at 109. *See also* Kearney, *supra* note 171, at 41; Peterson, *The Crime Lab* in THINKING ABOUT POLICE 184, 196 (C. Klockars ed. 1983) ("Crime laboratories are unique among publicly supported scientific operations in that few participate in external quality assurance programs."). *Cf.* BANBURY REPORT, *supra* note 159, at 244 (comments of C. Thomas Caskey, Institute for Molecular Genetics, Baylor College of Medicine) ("[W]hile the mechanisms for quality control are being put in place in forensic laboratories, it is clear that they are not very far along that pathway compared to other types of diagnostic laboratories."); Mudd & Presley, *supra* note 171, at 109 ("Although the application of formal [quality assurance and quality control] programs to forensic laboratories is not widespread and little information has appeared in the forensic science literature . . . , a great deal has been written on the application of [such] programs to clinical and federally operated laboratories.").

Some laboratories, of course, have instituted quality control programs on their own. For descriptions of several such attempts, *see* Brunelle, Garner, & Wineman, *supra* note 54, at 774; Kearney, *supra* note 171.

Forensic science has not even developed widely accepted criteria for quality control. *See* J. POKLEMBA, *supra* note 50, at 15 ("There are no widely accepted criteria for quality control or proficiency testing for forensic laboratories at a state or national level."). *See also* Pereira, *supra* note 74, at 1.

Part of the reason for the lack of consensus on quality control, as well as for other forensic science issues, is the fierce independence many labs maintain. *See* Lucas, Leete & Field, *supra* note 3, at 72.

> Forensic science laboratories . . . developed for the most part independently of each other. . . . While there is considerable professional cooperation between laboratories particularly through the various scientific societies, there is also some jealously guarded independence particularly between local, State and Federal authorities. As a result, a national blind proficiency testing program is beyond comprehension.

Id.

175. "The topic of quality control is one that is regarded with suspicion and contempt by most scientists, including those in the forensic laboratory." Thomson, *Bias and Quality Control in Forensic Science: A Cause for Concern*, 19 J. FORENSIC SCI. 504, 510 (1974). 510.

Assuring quality, however, is crucial for forensic laboratories. Experience and data have shown that quality control programs are necessary for guaranteeing proper clinical lab performance.[176] The absence of such plans from crime laboratories only assures unreliable testing by forensic facilities.[177]

Instead, of instituting and following good, formal quality control procedures, forensic laboratories too often cling to methods and routines that only decrease accuracy and reliability. A few of these practices illustrate the depth of the problem.

B. Lack of Protocols

Scientists in other areas live by protocols. Protocols, the lists of instructions for performing scientific procedures, are the recipes of science. Like good recipes, good protocols are tested procedures that, if followed, assure that the desired results are most likely to occur.[178]

Scientific procedure and common sense dictate the use of such protocols. If the scientist and others are to have confidence in the results, tested procedures must be followed to get those results. Procedures that do not follow an established protocol can only produce experimental outcomes. Of course, a crime lab should not be reporting these experimental results as scientific fact.[179] The importance good science places on adherence to protocols is illustrated by a study of scientists disciplined by the Food and Drug Administration. Eighty percent of those disci-

---

176. "The farther a lab operates from a quality assurance system that includes personnel standards, proficiency testing, and inspection, the less accurate its results are likely to be." *Hearings, supra* note 20, at 392 (statement of the American Association of Clinical Chemistry). "[E]very laboratory needs an internal quality control program to assure 1) that the staff understands and properly uses procedures, 2) that reagents, media, solutions, equipment, and procedures perform as expected, and 3) that only reasonably accurate and reliable information is reported to the clinician." LaMotte & Robinson, *supra* note 20, at 48.

177. *Cf.* Brunelle, Garner & Wineman, *supra* note 54, at 774 ("Quality assurance in forensic science laboratories is no less important than quality assurance or control in government regulatory laboratories, clinical laboratories, or in industry. The prosecutor, defense attorney, judge or jury, and, of course, the defendant have every right to expect accurate and reliable results from forensic science laboratories."); Pereira, *supra* note 74, at 1; Caskey & Hammond, *DNA-based Identification: Disease and Criminals*, in BANBURY REPORT 32: DNA TECHNOLOGY AND FORENSIC SCIENCE 128 (J. Ballantyne, G. Sensabaugh & J. Witkowski eds. 1989).

178. Even when tested procedures are followed, results will not always be accurate. As biologist Franklin Berger notes about DNA profiling: "Just like baking a cake, results are not always ideal. Day-to-day variability in the quality of test results may occur." Berger, *Chromosomes in the Courtroom*, Legal Times, Nov. 20, 1989, at 19.

179. *Cf.* Grunbaum, *supra* note 67, at 502 ("The bench analyst does not and should not experiment with evidence materials or offer untested hypotheses to explain ambiguous or inconclusive results.").

plined had failed to follow protocols.[180]

Crime laboratories, however, frequently perform analyses without adhering to established procedures. "In many crime laboratories in the United States, the analyst is not required to possess or to follow a printed protocol, nor is he required to heed any instructions or warnings."[181] In some areas of forensic science, established protocols do not even exist.[182]

An absence of tested protocols means that crime lab analysts are left to determine for themselves what modifications in established procedures or what new procedures will best fit their abilities, their equip-

---

180. Shapiro & Charrow, *supra* note 115, at 733. *See also* Woolf, *supra* note 103, at 74 (reviews scientific misconduct and concludes that failures to adhere to protocols "speak directly to concerns about authenticity of the data.").

181. Grunbaum, *supra* note 76, at 13. *Cf.* Bradford, *Quality Assurance—Considerations in the Expression of Professional Evaluations From Examinations*, 30 J. FORENSIC SCI. 981, 981 (1985).

> It was only after I began to review and study the work of other laboratories that I gained an insight into the extent of a lack of a standard of practice which exists in crime laboratory operations. . . . In the capacity of an independent consultant, one is in a position to review physical evidence examination results from a broad base of crime laboratory sources. The scope of variation in the form and method of expression of evaluations is surprising and in some instances dismaying.

Id.

> Frequently, forensic laboratories have inadequate standards for performing particular scientific analyses. Laboratory examiners also have substantial discretion in selecting the methods of analysis they will use. . . . These examiners may choose inappropriate methods instead of more accurate methods.

Moenssens, *supra* note 172, at 570.

Even within the well-regarded California Department of Justice laboratory system, "there are no standardized protocols. According to the DOJ Bureau of Forensic Services, 'a pre-set list of minimum performance does not apply to forensic science where the individual skill, judgment, and experience of the scientist are the final determiners of what specific tests are required in each instance.'" Grunbaum, *supra* note 73, at 31 (quoting STATE OF CAL. DEPT. OF JUSTICE, BUREAU OF FORENSIC SERVICES, TECHNICAL GUIDELINE 85–1 SEROLOGY 1).

See also Lander, *supra* note 172, at 505 (discussing a case in which questions were raised about the interpretation of DNA results because adequate controls had not been used.)

182. *See, e.g.*, Koenig, *supra* note 56, at 115 (no standard procedures used to make spectrographic comparisons); State v. Sager, 600 S.W.2d 541, 563 (Mo. App. 1980) (no standard procedure for arriving at conclusions by forensic odontologists). *Cf.* Lander, *supra* note 172, at 153 (citing urgent need for agreement of general guidelines for the practice of DNA typing).

ment, and the evidence.[183] Forensic scientists, then, are often using procedures that have never been truly scientifically tested.[184]

Forensic scientists defend the absence of tested procedures by asserting that such protocols cannot be used. They maintain that standardized practices can only be used with standard samples and crime labs do not analyze such samples.[185]

---

183. One commentator summarized testimony about one forensic technique as follows:

> Professor Sensabaugh agreed with Dr. Grunbaum that there was no uniform procedure of electrophoresis being used in crime laboratories. Each analyst used procedures with which he or she was most comfortable. Sensabaugh explained that analysts routinely modify the electrophoretic techniques as fits their convenience and equipment.

Bretz, *supra* note 38, at 518.

184. *See* Comment, *supra* note 156, at 674:

> The crime laboratories' diversity of procedure reflects ... disunity. For example, individual laboratories, and even individual technicians, frequently set their own idiosyncratic standards concerning testing protocols for the same basic serological test. Variation of protocols ("Protocol Drift") may cause inconsistent test results. Especially troublesome, the interpretation of test results may represent only one analyst's opinion.

Cf. Grunbaum, *supra* note 76, at 13:

> [Because established protocols are not consistently used,] it is hard to determine whether the most reliable methods of examination are used. There is no way of knowing what innovations [the analyst] has made to a standard reliable method, and it is difficult to establish whether or not he is using the method within its limits of reliability.

185. *See* Peterson, *Observations on Access to Expertise*, 101 F.R.D. 642, 643 (1983). A report by the Forensic Science Foundation, the research arm of the American Academy of Forensic Sciences, makes this contention: "[A] basic requirement of a standard method is a standard sample. In reality, such samples rarely exist in the forensic sciences. Consequently, forensic scientists are reluctant to use the standard or official approach, preferring instead, to use the methods they themselves have validated." Forensic Sciences Foundation, *Foreword* to CRIMINALISTICS METHODS OF ANALYSIS FEASIBILITY STUDY (1980), *quoted in* Grunbaum, *Potential and Limitations of Chemical Analysis of Physiological Evidence*, in ANALYTICAL METHODS IN FORENSIC CHEMISTRY 417, 426 (M. Ho ed. 1990). A pair of forensic drug analysts have concluded "that the use of standard methods of analysis is not applicable to the field of forensic drug chemistry." Frank & Gunn, *supra* note 165, at 163. They reasoned that

> because there is a great variance in the type of material being submitted for [the analyst's] examination, it is necessary to use a wide range of methodology in analyzing the material. Where one method of analysis may be suitable in several instances, the same method may give inconclusive results in other situations....
>
> [S]tandardization of chemical procedures is of value only under controlled circumstances, that is, when dealing with products of known composition or when examining a product to determine whether it meets certain established criteria.

This assertion is simply bad science. Although the forensic sample may be unknown, proper analysis will not be enhanced by introducing additional unknowns. Accurate analytical determinations require the elimination of variables, not the introduction of more. Adherence to protocols eliminates unnecessary unknowns. A distinguished scientist has asked, "If there is a proper way to do it, why not do it that way every time?"[186] If there is no proper way, how can anyone be confident of the results?

Even though samples may be non-standard, standardized methods must be devised and followed if consistently accurate results are to be obtained. The protocol, like a recipe that gives additional high altitude baking directions, should consider variability of the sample.[187] DNA profiling illustrates that variation can be included. The samples analyzed by DNA technology vary—dried blood, fresh blood, hair roots, semen—and often have unknown histories. Even so, forensic scientists are capable of devising protocols for DNA profiling.[188] If protocols can be established for DNA identifications, they can be devised for any area

---

Id. at 166–67. *See also* LEAA, *supra* note 1, at 26 ("These different approaches do not lend themselves to the type of testing that is carried out by most other types of 'testing' laboratories wherein a set protocol for the examination of a given sample of anything must be followed.").

186. BANBURY REPORT, *supra* note 159, at 186 (comment of Dr. Richard Roberts, Cold Spring Harbor Laboratory).

187. *Cf.* Peterson, *supra* note 185, at 643 ("[T]he uniqueness of a case does not justify use of methods that have not been evaluated by other scientists in other laboratories."); Bashinski, *supra* note 27, at 161 ("The evidence samples to be analyzed are uncontrolled and unpredictable, often minute in quantity, and contaminated or degraded by the environment. The criminalist must be capable of assessing the potential effects of these factors on the analysis and of adjusting the procedures accordingly.").

Good scientific practice also requires the keeping of records indicating how the protocol was carried out. Unfortunately, inadequate recordkeeping appears to be endemic in crime labs. *See* Comment, *supra* note 156, at 674 n.49 ("[A]dequate test records are neither consistently generated nor preserved. Again, the adequacy of record-keeping varies from facility to facility, and examiner to examiner."). *See also* Note, *supra* note 50, at 492; Winek, *supra* note 56, at 181; Kearney, *supra* note 171, at 46.

188. *See, e.g.*, Note, *DNA Identification Tests and the Courts*, 63 WASH. L. REV. 903, 927 (1988) ("[T]he Central Research Establishment, a government forensic lab in the United Kingdom, has produced a 51-page lab protocol outlining the procedure for receiving a sample for DNA analysis and describing in detail the isolation, sorting, and probing of the DNA.") *See also* Kearney, *supra* note 171 (twenty page outline for assuring quality in DNA analysis). *Cf.* Sensabaugh, *Forensic Biology—Is Recombinant DNA Technology in its Future?*, 31 J. FORENSIC SCI. 393, 394 (1986).

The fact that some labs have drafted protocols does not mean that all labs have. *See, e.g.*, Note, *supra*, at 928 ("[N]ot all [DNA] labs have clearly codified their protocols or made them available for peer review."). The fact that a lab has a protocol does not mean it is followed. For a discussion of a DNA lab deviating from its protocol, see generally Lander, *supra* note 172.

of forensic science.[189] That forensic scientists in fact do not institute and routinely follow tested analytical procedures is an indication that forensic labs cannot produce consistently accurate results.[190]

C. Bias and Subjectivity in Forensic Laboratories

Failure to follow established protocols is not the only inferior analytical technique. Another is the failure of forensic scientists to shield themselves from possible bias.

Evidentiary material often is presented to the forensic scientist in a needlessly suggestive manner. The analyst is given crime scene evidence (hair, fingerprints, blood) and one other sample, labelled as the defendant's (or victim's). This is frequently accompanied by a synopsis of the investigation indicating the reasons that the investigators believe the suspect is guilty.[191] Such a presentation, of course, suggests to the

---

189. "Some procedures (for example, firearms identification) lend themselves to standards more easily than do other procedures (that is, handwriting comparisons), but each procedure is subject to standards, which can be composed by a group of qualified experts from the fields involved." Thomson, *supra* note 175, at 510.

190. "The key for a forensic confirmation is to establish objective criteria for testing and reporting and then apply the objective criteria to the laboratory data, reporting as positive *only* those specimens that fully meet the criteria." Council on Scientific Affairs, *Scientific Issues in Drug Testing*, 257 J. AM. MED. A. 3110, 3111 (1987). *See supra* Parts I and II.

191. *See, e.g., State v. Sager, 600 S.W.2d 541, 564 (Mo. Ct. App. 1980)*, (where the court indicates that only the picture of the bite mark on the victim and the dental casts of appellant were delivered to the forensic odontologist for identification; Miller), *Procedural Bias in Forensic Science Examinations of Human Hair*, 11 LAW AND HUM. BEHAV. 157, 158 (1987).

> In the examination and identification of human hair, the usual procedure involves obtaining hair samples from a suspect for comparison purposes (known versus questioned). Police investigators usually submit the questioned and known suspect hair samples along with a synopsis of facts surrounding the investigation. The main purpose of the synopsis is to provide information to the examiner that may assist in the analysis. The synopsis usually contains the facts and circumstances leading to the arrest of a suspect. In some cases, the synopsis may even include eyewitness accounts, other forms of physical evidence collected in the case, and admissions or confessions made by the suspect.

Id. Use of lineups overcomes many prior abuses, and a parallel can be drawn with other investigative techniques. When the crime laboratory for the prosecution is presented with handwriting samples, hairs or other biological specimens, voice exemplars, and similar evidence, the laboratory personnel know that they are from persons accused of crime. There are no blind experiments in which some false samples deliberately are introduced to test the validity of results. This is standard procedure in many aspects of scientific methodology, yet we do not take sufficient advantage of it in our forensic applications. *See id.*; Starrs, *Comments on Discovery and Its Application to the Lindbergh Kidnapping Case*, 101 F.R.D. 625, 626–27 (1983).

analyst what the "right" answer should be.[192] This suggestiveness coupled with the understandable prosecutorial orientation of many forensic scientists[193] will naturally, even if unconsciously, skew subjective judgments.[194]

The effects of suggestiveness and bias are well-known generally in the scientific world[195] and have been empirically validated in a forensic

---

192. Miller observes:

> In criminal investigations, the police generally have little or no doubt regarding the suspect's guilt. Their preoccupation lies with obtaining sufficient proof for a conviction.... That attitude may be communicated to the forensic examiner through personal interaction or through the written synopsis accompanying the evidence to be examined. It is conceivable that the forensic examiner may unconsciously believe that the suspect must be guilty or the police would not have made the arrest. Such unconscious beliefs may potentially create prejudice, bias, and stereotypes on the examiner's part regarding conclusions about the evidence.

Miller, *supra* note 191, at 158. *See also* Grunbaum, *supra* note 76, at 15; Aitken & Robertson, *supra* note 136, at 687.

193. Thomson, *supra* note 175, at 509–10:

> [T]he preponderance of well-qualified forensic laboratories are located within the resources of the State.... They work hand in hand with the police from the beginning of the case.... Is the witness who has his job and salary controlled by the State completely free from pressure, conscious or unconscious, to be entirely impartial?

See also Kates & Guttenplan, *Ethical Considerations in Forensic Science Services*, 28 J. FORENSIC SCI. 972, 972 (1983) (describing hand-in-hand relationship between the forensic scientist and the prosecutor). A survey of forensic science laboratories found that 57% would only examine evidence submitted by law enforcement officials. Peterson, Mihajlovic & Bedrosian, *supra* note 97, at 13. *Cf.* Lucas, Leete & Field, *supra* note 3, at 72 ("Of the more than 300 'crime labs' in the USA, about 80% are within law enforcement agencies"); M. SAKS & R. VAN DUIZEND, *supra* note 25, at 53, conclude, "Given what is known about reference group phenomena, the need that people have for social support of attitudes and conduct; and the process of socialization in occupational settings, it strains credulity to believe that these experts do not identify with prosecutors"; *Cf.* Note, *supra* note 50, at 500 ("The incentives for the private laboratories to get results that they can sell to a legal community eager to convict create examiner bias.").

194. *See* Kates & Guttenplan, *supra* note 193, at 975 ("The expert witness realizes that the testimony must be presented so that it is advantageous to the prosecutor, if the prosecutor is to use that witness again."). *See also* Note, *supra* note 50, at 485 ("[T]he examiner who looks for a match is more likely to find one particularly because it would be consistent with the prosecutor's theory of the case."). Sometimes the prosecutorial pressure put on forensic scientists to give "correct" opinions can be quite direct. *See, e.g.*, McCarty v. State, 765 P.2d 1215, 1219 (Okla. Cr. App. 1988) (The Oklahoma County District Attorney's Office may have placed undue pressure upon the forensic analyst to give an expert opinion beyond scientific capabilities.).

195. "Scientists are subject to self-deception or forms of experimental bias that they cannot reasonably be expected to control. For example, science that relies heavily on human observation skills is especially prone to the self-deception phenomena.... Experimental bias in studies using human subjects is so common that the use of a 'double blind' has become standard practice not only in human research but also in animal experimentation.

science context. Half of a group of students trained in hair comparisons were given samples in the normal casework manner—two samples and a statement that one sample was from the crime scene and the other from the suspect. The remaining students were given hair from the crime scene plus hair samples from five possible suspects. Thus, the experimenter presented the evidence in a suggestive method to half, and a "lineup of hair" to the rest.

Startling differences were observed. While sixteen percent of all the opinions rendered were wrong,[196] the error rate varied significantly according to the method of presentation. When the hair was presented in the usual suggestive crime lab manner, 30.8% of the conclusions were wrong. When a lineup was presented, only 3.8% of the opinions were incorrect. The investigator concluded:

> A preconceived conclusion that a questioned hair sample and a known hair sample originated from the same individual may influence the examiner's opinion when the samples are similar. The data in the present study indicate that the method by which hair evidence is submitted for analysis may encourage unintentional bias on the part of the examiner.[197]

A double blind prevents both the experimenter and the subjects of the experiment from knowing which subject are the controls." Anderson, *The Federal Government's Role in Regulating Misconduct in Scientific and Technological Research*, *supra* note 101, at 128–29.

196. The "crime scene" hair in fact did not match any of the other submitted samples. In other words, the correct answer to each analysis was "no match." *See* Miller, *supra* note 191, at 160.

197. *Id.* at 161. Seven inconclusive results were omitted from the incorrect/correct tabulations. All the inconclusives came from the lineup presentation. "Although an 'inconclusive' opinion may not be accurate, it is more correct than rendering an erroneous conclusion based on a preconceived opinion. The 'line-up' method appears to have increased opinion accuracy since none of the inconclusive opinions were expressed using the traditional method of hair analysis. . . ." *Id.*

The results for hair are generally applicable to the forensic sciences. For example:

> Usually [the analyst] knows what genetic markers to look for to strengthen the evidence against a suspect. For instance, he may analyze a victim's blood, a suspect's blood, and [a] bloodstain found on the shirt of the suspect. Consciously or unconsciously, he will be attempting to match the bloodstain with the blood of the victim. Because of his mental set, he may "identify" a genetic marker that he would otherwise declare to be inconclusive.

Grunbaum, *supra* note 76, at 15.

A similar bias may also affect validation studies done in research labs. *Cf.* Note, *supra* note 102, at 44 ("Whether scientists admit it or not, their views as to what constitutes a reliable scientific study may depend upon the source of their paycheck. . . ."); Thompson & Ford, *DNA Typing: Acceptance and Weight of the New Genetic Identification Tests*, 75 VA. L. REV. 45, 88–89 (1989).

Since subjectivity is common in forensic science opinions, routine procedures that induce bias can have a great effect on quality.[198] The fact that laboratories do not rigidly employ systems insuring unbiased examiners, even when that would be simple to do,[199] again raises concerns about the accuracy of crime lab work.

> [I]t is important to realize that the priorities of forensic experts . . . may be quite different than the priorities of scientists working to validate DNA typing techniques. [T]he police criminalist . . . may adopt a lower threshold for declaring a match in order to avoid exculpating a guilty person. For a scientist employed by a company attempting to market a DNA typing test, however, the key priority is to demonstrate that there is a low probability of a match between samples from different individuals. These scientists may therefore adopt a higher threshold, in order to minimize the number of coincidental matches.

Id.

198. Professor Giannelli has noted:

> [M]any commonly employed procedures are based on subjective judgments. . . . Even apparently routine and objective procedures involve an element of objectivity. For example, a firearms identification examiner may conclude that two bullets had been fired from the same weapon. Although a positive identification is based on objective data—the striations on the bullet surfaces—the examiner's conclusion rests on a subjective evaluation. There are no objective criteria used for this determination.

Giannelli, *supra* note 80, at 691. *See also* Imwinkelried, *supra* note 2, at 638 ("Some scientific fields have relatively objective standards for evaluating test results. . . . However, other scientific fields have few, if any, objective standards.").

> Certain types of evidence involve accepted scientific principles and the measurement of scientific properties which can be described in absolute terms by the expert. Others involve comparisons of individual characteristics on a case-by-case basis. Here the expert opinion is based partially on scientific principles and supplemented by the expert's personal opinion. An element of subjectivity will enter into the formulation of opinion about the second form of evidence.

Note, *supra* note 138, at 328–29.

199. "The possibility of bias is minimized for any testing procedure if all analysts are unacquainted with circumstances of the alleged crime and unaware of any previous results." Selvin & Grunbaum, *Genetic Marker Determination in Evidence Bloodstains: The Effect of Classification Errors on Probability of Non-discrimination and Probability of Concordance*, 27 J. FORENSIC SCI. SOC'Y 57, 61 (1986). *See* Note, *supra* note 188, at 927 ("Results should be interpreted by a technician who was not involved in the test and is unaware of the sources of the sample."). *See also* Note, *supra* note 50, at 486 ("[E]xaminers should be told neither the origin of the samples nor the prosecution's theory of the case.").

D. Lack of Replication

According to the demands of good science, the ability to replicate a scientific finding must exist before the finding can be accepted. Forensic science, however, employs careless procedures by not insisting on replication whenever possible.[200]

Sometimes tests are not repeated because repetition is impossible. Forensic samples are often so minute that they may be consumed in a single test or so degradable that they cannot be preserved.[201] Even with these samples a form of replication can be obtained by having the results independently interpreted by separate examiners.[202] This is not regularly done, nor is it routine to replicate analyses when replication is possible.[203]

---

200. "Scientific experiments which produce a finding of interest are usually repeated, sometimes many times, to be sure they are accurate. Often several scientists will work independently on a given problem.... The situation in a forensic laboratory is quite different. Tests often are not repeated." Thompson & Ford, *supra* note 197, at 56–57.

201. Giannelli, *supra* note 9, at 1242.

202. *See* Pereira *supra* note 74, at 5, discussing quality assurance in British crime labs where a "particularly important requirement [of the program is] the confirmation of all significant findings by an appropriate member of the staff." *See* Note, *supra* note 50, at 486 ("[A]ny declaration of a match should be confirmed independently by at least two examiners.").

203. "[M]ost case work is never reexamined by a second expert and, thus, many things may go undetected. Of the cases we reexamined over recent years, some 15 to 20% could be considered as having questionable procedures or deficiencies of some sort." De Zeeuw, *Procedures and Responsibilities in Forensic Toxicology*, 27 J. FORENSIC SCI. 749, 752 (1982). *See also* Lavett, *Electrophoresis: A Continuation of the Discussion*, 29 J. FORENSIC SCI. 704, 704 (1984) ("In scientific laboratories, crucial tests are always repeated. Although the results of criminal laboratories have an effect on the accused's life and/or liberty, I have never seen a crime lab report that gives more than one test result on a given sample.").

Thompson and Ford contend that only results unfavorable to the prosecutor are likely to be reconsidered. "Errors 'stand out' and invite additional scrutiny only if they are inconsistent with the prosecution's theory of the case (of which the laboratory analyst is invariably aware). The most dangerous errors, those which falsely incriminate someone on whom suspicion has already focused, are likely to go unchallenged...." Thompson & Ford, *supra* note 197, at 57.

> [A]n electrophoretic run ... indicated an EsD type of 1 for the bloodstain taken from the porch of the victim's home. This result was followed by a question mark and the stain was retested on June 6, 1978. At the later testing, the same stain showed an EsD type of 2–1. Without retesting a third time, [the forensic scientist] only reported the latter result. Defendant's blood, tested on July 7, 1978, was an EsD type 2–1.

Bretz, *supra* note 38, at 519.

Cf. Baird, *Quality Control and Quality Assurance*, in BANBURY REPORT 32: DNA TECHNOLOGY AND FORENSIC SCIENCE 175 (J. Ballantyne, G. Sensabaugh & J. Witkowski eds. 1989) ("If there appears to be a sufficient amount of DNA available for analysis, ... the evidence should be analyzed in duplicate.").

The failure to reproduce results is not good science and fails to minimize errors.[204]

E. Special Need for Crime Laboratory Quality Control

The concern about the lack of routine is magnified by the forensic sciences' special needs for demanding analytical practices. The error rate in a crime lab's analyses is inherently higher than the rate for other types of testing.

Forensic facilities deal with more "unknowns" than other laboratories. For example, a hospital lab typing blood for a transfusion types the ABO factor in two unknowns — the patient's blood and the donated blood. The parallel procedure in crime work, however, normally requires analysis of at least three unknowns — the suspect's blood, the victim's blood, and a bloodstain, perhaps found at the crime scene or on the suspect's clothes. If the error rate for typing ABO is X, then for a transfusion that rate will be 2X[205] while in the crime lab it will be at 3X if three samples are involved (and higher if more samples are involved).[206]

---

204. *See* Grindon & Eska, *Error Rate, Precision, and Accuracy in Immunohematology*, 17 TRANSFUSION 425, 428 (1977): "If an error in ABO grouping is made once in 1,000 determinations, and a net error rate of that magnitude is felt to be unacceptable, it would be important to have a second person perform the test independently and then have the results compared. If the error rate remains constant under these conditions, the net error rate becomes one per million." *See also* Saltzburg, *Frye and Alternatives*, 99 F.R.D. 208, 217 (1983) ("There is sufficient evidence that crime labs are not as careful and dependable as they might be to warrant rules requiring second tests of physical evidence that may have an important impact on trials. These tests should be blind — i.e., one tester should not know what another has found.").

205. Actually, the rate is slightly less than 2X. If one imagines a simple system with two possibilities, A or B, then a "correct" result is obtained if an error is committed in typing both samples. Such a result will occur with a probability of X^2. Thus, the true error rate is $2X-X^2$. For small error rates, this approximates 2X. For more intricate systems, such as the ABO multiple allele bloodsystem, the calculations are more complex, but the principle is the same. The resulting error rate is reasonably close to 2X.

206. This pattern is not limited to serology, but applies generally to all crime lab work. Thus, if for some reason, a chemist wanted to know the paints on a car, he might examine a paint chip from that vehicle. The forensic scientist might also examine the chip from the car, but only to compare it with chips left at the scene of a hit and run accident — two unknowns versus one, and twice the inherent error rate. *Cf.* Lander, *supra* note 172, at 501:

> DNA diagnostics requires simply identifying whether each parent has passed to a child the ... pattern inherited from his or her mother or father. Because the four discrete patterns are known in advance, these investigations have built-in consistency checks which guard against many errors and artefacts.
>
> DNA fingerprinting, by contrast, is more like analytical biochemistry: one must determine whether two completely unknown samples are identical.

These errors in the crime lab are also less likely to be noticed and corrected than in other labs. Forensic tests may have fewer inherent controls than other kinds of testing.[207] Also, errors in forensic labs will not be highlighted as are errors in research labs. Good researchers often recognize incorrectly performed procedures because the results will conflict with existing knowledge.[208] For example, a scientist studying the effects of a drug on hormone levels would notice if the drug appeared to elevate the hormone level higher than had been seen before. The experiment certainly would have to be reproduced a number of times before a result considered unusual would be accepted. If the first result had been an error, it would be discovered. In contrast, if the forensic scientist typing the ABO factor in a bloodstain incorrectly finds AB instead of O, that result often will not stand out as inconsistent with existing knowledge. As a result, the test will normally not be repeated.

Since errors in the crime lab are inherently more frequent and less likely to be noticed than in other kinds of facilities, forensic procedures should be carried out even more meticulously than other kinds of tests. Forensic labs should require more stringent quality control programs than labs where errors will be fewer and more apparent.[209]

IV. THE COURTS AND FORENSIC SCIENCE QUALITY

The courts seem to provide an external mechanism to enforce higher quality in forensic science. However, they can not adequately accomplish this task. Courts only deal with the evidence in the cases before them. This prevents them from confronting the full range of issues concerning forensic laboratories.[210] Courts only see a fraction of a forensic scientist's analyses.[211] They almost never examine the inaccurate work

207. *See supra* text accompanying notes 167–170.

208. *See* Thompson & Ford, *supra* note 197, at 56 ("errors tend to stand out [in a research lab] because they are inconsistent with scientific knowledge and theory").

209. "[R]esearch scientists can tolerate relatively high rates of error and unreliability in their procedures [, but] . . . error rates that are acceptable in research laboratories are intolerable in forensic laboratories." *Id.* at 56–57.

210. *Cf.* Comment, *supra* note 156, at 678. ("Because judges are limited to assessing the admissibility of DNA printing as evidence, they are constrained, at least until the issues reach them in a justiciable form, from examining the broader concerns of crime lab performance.").

211. "The prosecution often uses scientific evidence to build a case but does not call upon an expert to testify. In a recent national survey of forensic science laboratories, respondents estimated that they testified in only about ten percent of the cases in which they examined the evidence." *Symposium, supra* note 5, at 630 (comments of Prof. Joseph Peterson).

that produces inconclusive findings or leads to a false exclusion.[212] Indeed, despite widespread inaccuracies in forensic lab reports, courts are unlikely to discover such problems, and are therefore unable to remedy the widespread consequences of poor forensic work.

Courts are also unlikely to be effective reformers of laboratory quality because of their limited tools. Courts are essentially restricted to excluding evidence. If the courts regularly refused to admit analyses when laboratories had not followed programs assuring high quality work, improvements could follow.[213] Judges, however, are unlikely to adopt such a procedure. Furthermore, such court action would not be an efficient way to produce the needed systemic reforms. For example, until the highest court in a jurisdiction had ruled on an issue, laboratories would have to guess what quality control efforts each court might demand of them. Labs might be subject to as many regulators as there are courts in a state. Such multiple regulation would lead to inefficiencies.

Even if the courts were willing to devise and supervise scientific quality control programs, the courts would be dependent upon the adversary system. That system often does not operate effectively in assessing science. An adversary process that performs properly requires lawyers who can effectively expose weaknesses in scientific evidence and a judiciary not overawed by science.[214] These qualities are often absent because lawyers and judges usually lack high levels of scientific training and are

---

212. Forensic testing often eliminates people as suspects. For example, David Werrett, of the British Home Office, in referring to DNA profiling, stated, "When we began, about 15% of 'very good suspects,' according to the police, were eliminated. The current rate is 25%." Westin, *A Privacy Analysis of the Use of DNA Techniques as Evidence in Courtroom Proceedings,* in BANBURY REPORT 32: DNA TECHNOLOGY AND FORENSIC SCIENCE 25 (J. Ballantyne, G. Sensabaugh & J. Witkowski eds. 1989).

213. *Cf.* M. SAKS & R. VAN DUIZEND, *supra* note 25, at 74. ("How can the courts assure the integrity of the knowledge offered to them? One means is to ensure that individuals who testify as experts have the requisite training and experience and that forensic laboratories maintain rigorous quality control.").

214. *Cf.* Levy, *supra* note 159, at 383: An effective adversary system for scientific evidence requires that

> litigants against whose interest scientific testimony has been introduced have available an adequately trained and effective expert to rebut its implications, a lawyer who can effectively expose the weaknesses of the testimony through rigorous cross-examination, a judge who is neutral and objective and willing to exercise his evidence admissibility discretion wisely and unlikely to be seduced by the "aura of infallibility" of the testimony.

thus unable to challenge or evaluate science.[215] Both prosecutors and defense attorneys seem unwilling to challenge the accuracy of test results. Even though shoddy forensic science may help free the guilty more often than inculpate the innocent, prosecutors are unlikely to verify whether the exculpatory finding is warranted.[216] Similarly, defense attorneys often seem to trust the reliability of forensic test results.[217] Thus, despite the high error rates revealed by the LEAA testing program,

---

215. A. MOENNSENS, F. INBAU & J. STARRS, *supra* note 5, at 7:

> [L]awyers as a group evidence an appalling degree of scientific illiteracy, which ill equips them to educate and guide the bench in its decisions on admissibility of evidence proffered through expert witnesses. This scientific illiteracy is shared by a large segment of the trial and appellate bench; many judges simply do not understand evidence based on scientific principles.

See also McCord, *Syndromes, Profiles and Other Mental Exotica: A New Approach to the Admissibility of Nontraditional Psychological Evidence in Criminal Cases*, 66 OR. L. REV. 19, 25 (1987) ("[M]ost lawyers and judges do not have extensive training or experience in dealing with scientific matters. Accordingly, many lawyers and judges are uncomfortable dealing with scientific evidence."); Sperlich, *Social Science Evidence and the Courts: Reaching Beyond the Adversary Process*, 63 JUDICATURE 280, 288 (1988): ("Judges and their clerks are not normally able to distinguish relevant from irrelevant materials and good science from bad."); *but see* Ellinger, *DNA Diagnostic Technology: Probing the Problem of Causation in Toxic Torts*, 3 HARV. J.L. & TECH. 31, 54 (1990) ("The DNA fingerprinting cases support the thesis that courts are capable of familiarizing themselves with the 'science' of DNA diagnostic technology and of reaching informed decisions regarding the admissibility of test results.").

216.

> [T]he legal community that is pushing the evidence through the courts appears unconcerned about the possibility of spurious alteration of the [DNA] profile which would exculpate the guilty, a far more likely result of this test than a false match incriminating the innocent. If false incrimination does not give prosecutors pause, false exculpation should. However, until squarely faced with a case where all circumstantial evidence except the DNA profile point to a defendant's guilt, prosecutors will continue to opt for the odds that the test will be performed correctly.

Note, *supra* note 50, at 50 n.83.

217. *See Symposium, supra* note 5, at 634 ("Often the plaintiff or the prosecution alone calls an expert who testifies with little or no cross-examination. The opponent more or less accepts the expert's evidence as being true if it does not concern a major issue in contention.") (comments of Prof. Michael Graham). Even when novel scientific evidence is presented, it often goes unchallenged by opposing experts. Giannelli, *supra* note 9, at 1243 (A "surprising number of novel techniques have gained admissibility without the presentation of defense expert testimony."). *Cf.* Peterson, Ryan, Houlden & Mihajlovic, *supra* note 12, at 1749:

> [D]efense attorneys [use] a variety of tactics . . . to challenge forensic science evidence, ranging from efforts to have the evidence ruled inadmissible . . . to attacks on the expert's qualifications or intense cross-examination of the expert's conclusions. Usually, however, defense counsel will attempt to "explain away" the physical evidence by supplying a reasonable and lawful explanation for its presence. If the above tactics cannot be used, defense counsel will usually stipulate to the evidence and attempt to draw as little attention to it as possible. Contrary to a commonly expressed attitude that defense attorneys distrust the analyses and testimony of

"[t]here is little reason to think that many of such errors were detected by counsel for the defense."[218] Furthermore, meaningfully challenging the evidence may not be possible because of discovery limitations,[219] the inability to retest the evidence,[220] and the lack of adequate expert assistance.[221] Courts, therefore, are unlikely to bring about better forensic

---

> "prosecution" experts, defense counsel we interviewed are basically satisfied with the competence and nonpartisanship of forensic scientist[s] with whom they have contact.

218. M. SAKS & R. VAN DUIZEND, *supra* note 25, at 10.

219. "[P]retrial discovery is critical when scientific evidence is admitted at trial." Giannelli, *Observations on Discovery of Scientific Evidence*, 101 F.R.D. 622, 622 (1983). For a discussion of the limitations on discovery in criminal cases as compared to civil cases, see Saltzburg, *Pretrial Discovery of Scientific Evidence*, 101 F.R.D. 613, 613 (1983).

Rule 16 of the Federal Rules of Criminal Procedure does permit defense access to the results and reports of scientific tests and experiments, but this is of limited value because Rule 16 excludes important items such as:

> the identity of witnesses and discovery of the prior statements of any witnesses other than the results or reports of a scientific expert. . . . Moreover, the provision making scientific reports discoverable is of limited benefit since rule 16 does not require that a scientific expert make a report and does not specify the data that must be included in such reports. Hence, in those cases in which a report is prepared, the defense lawyer facing a scientific expert will know the name of the expert and the intention of the government to call him, but may have little other information to help the lawyer prepare for cross-examination.

Eads, *Adjudication by Ambush: Federal Prosecutors' Use of Nonscientific Experts in a System of Limited Criminal Discovery*, 67 N.C. L. REV. 577, 583 (1989).

220. Although "reliability of a technique means that the test results can be reproduced . . . ," Note, *supra* note 50, at 524, replication of crime laboratory analyses are rare. *See supra* text accompanying notes 200–04. Sometimes retesting is impossible simply because the sample was not saved. The Supreme Court, however, failed to find a constitutional violation where the prosecution did not preserve forensic science samples. California v. Trombetta, 467 U.S. 479 (1984) and Arizona v. Youngblood 488 U.S. 51 (1988). For a summary of state law on the right of defendants to retest forensic evidence, see Note, *supra* note 50, at 523–25.

221. Although Ake v. Oklahoma, 470 U.S. 68 (1985) held that an indigent defendant had a due process right to a defense psychiatrist when the defendant's mental condition was at issue: "the courts do not consider all denials of defense requests for expert services to be constitutional violations." Note, *supra* note 50, at 521. *See, e.g.*, Moore v. Kemp, 809 F. 2d 702 (11th Cir.)(en banc), *cert. denied*, 481 U.S. 1054 (1987) (denial of the indigent defendant's request for expert assistance did not violate the Constitution despite the state's reliance on a microanalyst and serologist.) *See* Harris, *Ake Revisited: Expert Psychiatric Witnesses Remain Beyond Reach for the Indigent*, 68 N.C. L. REV. 763, 769–74 (1990) (discussion of *Moore*.)

Even when jurisdictions allow indigents to employ experts, the pay for the experts may be so low as to make it almost impossible to get expert assistance.

> Texas, for instance, provides that indigents may receive five hundred dollars for "expenses incurred for purposes of investigation and expert testimony." If the expert is required to travel any considerable distance, it is reasonable to assume that after expenses there would be little left with which to pay a fee for testifying, much less any money to cover the costs of independent experimentation.

science practices because those practices are seldom fully analyzed and challenged in our adversary system.

Even if the courts had the technical ability to engage in such rigorous analysis, they are unlikely to have the time. Because there are no enforceable systems of quality assurance in place, the courts would have to do in-depth, individualized examinations of the practices, routines, and results of the laboratories in each case involving forensic science evidence. The time required for such a process would be prohibitive.

Furthermore, the current legal standard for admissibility of scientific evidence limits the courts' inquiry into the quality of the forensic work performed. In determining admissibility, courts examine the reliability and acceptance of a technique, but generally do not assess how well a particular analysis has been performed or what quality controls were used.[222] The dominant position, echoed by the California Supreme Court, states that the admissibility standard "tests the fundamental validity of a new scientific methodology, not the degree of professionalism with which it is applied.... Careless testing affects the weight of the evidence, not its admissibility."[223]

Note, *supra* note 111, at 374 (quoting TEX. CRIM. PROC. CODE ANN. § 26.05(d) (Vernon Supp. 1985)). *See also* Note, *supra* note 50, at 522. ("Although about half the states and the federal government have statutes which provide funds for defense experts, those funds are prohibitively low, and the statutes suffer from the problematic interpretations of what is considered 'necessary' to a defense.").

Even when adequate funds are available, "a fundamental problem is locating a qualified expert to analyze the testimony to be advanced by the opponent's expert." Imwinkelried, *Observations on Access to Expertise*, 101 F.R.D. 639, 640 (1983). *See also* Neufeld & Colman, *supra* note 157, at 52 (describing the difficulties defense attorneys have had finding qualified experts in DNA profiling cases). Sometimes defendants have trouble obtaining experts because "the experts are either employees or former employees of law enforcement agencies. This inhibits in particular the ability of the defense to present independent evidence of ballistics and fingerprints." M. SAKS & R. VAN DUIZEND, *supra* note 25, at 26.

222. *See* Recent Development, *DNA Fingerprinting: The* Castro *Case*, 3 HARV. J.L. & TECH. 223, 230–233. (Courts should look at the particular procedure performed, not just the acceptance of that type of procedure.).

223. People v. Farmer, 47 Cal.3d 888, 913, 765 P.2d 940, 956, 254 Cal. Rptr 508, 524 (1989).

The refusal to consider the quality of the work in admitting crime lab evidence may harm accurate factfinding. *See, e.g., Symposium on Rules for the Admissibility of Scientific Evidence*, 115 F.R.D. 79, 135 (1987).

> It is not difficult to determine that some technical methods, such as chromatography, are generally accepted and can be used with confidence that they are reliable. But it is an altogether different matter to determine whether a particular datum derived from a specific sample of urine or blood is trustworthy. We should concern ourselves with such matters as how the sample was collected, whether it was correctly identified, how it has been transmitted and stored, who performed the analysis, whether the reagents were fresh and the instruments calibrated, and related factors that bear on the reliability of the results.

Id. (comments of Vincent Dole).

If the court system is to have a role in improving the quality of forensic science, that job belongs not to the judiciary when it decides what evidence is admissible, but to the jurors when they decide what weight is to be given to scientific evidence.[224] Unfortunately, there is little reason to believe that jurors can make adequate assessments about the quality control programs of labs when they weigh scientific evidence.[225] Moreover, while juries may determine that the crime lab work in a particular

---

Some courts do employ an admissibility standard that looks to how well the work was performed. P. GIANNELLI & E. IMWINKELRIED, SCIENTIFIC EVIDENCE 39–43 (1986 & Supp. 1988). (courts have divided on whether showing of proper application of a valid technique affects admissibility or just weight of evidence). For example, the Minnesota Supreme Court has held that the "admissibility of specific test results in a particular case hinges on the laboratory's compliance with appropriate standards and controls . . ." People v. Schwartz, 447 N.W. 2d 422, 428 (Minn. 1989). The dominant position, however, maintains that compliance with standards and controls only goes to the weight of the evidence. *See, e.g.*, Note, *supra* note 188, at 938 ("challenges to proper administration of [a DNA test] go to the weight given to the evidence, not to the evidence's admissibility.") *See also* Black, *Evolving Legal Standards for the Admissibility of Scientific Evidence*, 239 SCI. 1508, 1508 (1988) ("Judges . . . have generally refused . . . to hold experts to the standards and criteria of scientific practice.").

224. Even if the courts used a different standard for judging the admissibility of novel scientific evidence, the quality of forensic science work would still be largely left to juries. Once appellate courts have found novel scientific evidence to be admissible, trial courts no longer hold admissibility hearings. Instead, the evidence is heard by the jury, and the opponent of the evidence is limited to attacking the weight of that evidence. *See, e.g.*, People v. Kelly, 17 Cal.3d 24, 32, 549 P.2d 1240, 130 Cal. Rptr. 1240 (1976) ("Once a trial court has admitted evidence based upon a new scientific technique, and that decision is affirmed on appeal . . . , the precedent so established may control subsequent trials, at least until new evidence is presented reflecting a change in the attitude of the scientific community."); People v. Brown 40 Cal.3d 512, 530, 726 P.2d 516, 230 Cal. Rptr. 834 (1985) ("appellate endorsement of a technique ends the need for case-by-case adjudication.").

225. *See* Note, *The Frye Doctrine and Relevancy Approach Controversy: An Empirical Evaluation*, 74 GEO. L.J. 1769, 1776–87 (1986), for a discussion of recent research on jurors' perceptions of novel scientific evidence. None of the studies examined whether jurors were able to evaluate the effect of quality control programs on the reliability of scientific evidence. Nevertheless, it seems to be asking a lot of jurors to make such evaluations. *See, e.g.*, Thompson & Ford, *supra* note 197, at 58:

> If all that a court requires is acceptance of a general approach [to DNA typing], then anyone with the requisite scientific training who sets up a laboratory to do [DNA] analysis would, presumably, be able to testify about his or her results. It would be up to the trier of fact . . . to assess whether the techniques of [DNA] analysis had been implemented properly by the lab. This is a heavy burden to place on the trier of fact . . . [since] the procedures involved in DNA typing are quite complex. There are many points in DNA typing where nuances of laboratory procedure can have a dramatic effect on the reliability of the test. Evaluating the adequacy of laboratory protocols is likely to be a difficult, time-consuming task for attorneys and jurors.

Finally, if juries were providing incentives for crime labs to improve their quality, then we could expect prosecutors, concerned about losing future convictions, to demand better quality controls from crime labs. However, nothing indicates that prosecutors are making significant attempts to improve the performance of forensic scientists.

case was substandard and should not be relied upon, such determinations are unlikely to induce improved performance in the forensic sciences generally. The jurors' assessments are not officially communicated. Juries in criminal cases render general verdicts. A not guilty verdict, even if it were influenced by sloppy crime lab practices, does not explicitly state the jurors' views about the scientific evidence. Similarly, the jury might convict on other grounds despite finding the scientific evidence itself unreliable. Since the jury's determination of scientific evidence reliability is never definitively known, it is not likely to spur better quality.

Thus, the procedural and substantive limitations of the court system cannot adequately ensure the reliability of forensic science. No mechanism currently in place assures quality.

V. EFFECT OF REGULATION ON LABORATORY QUALITY

Only some medical laboratories were regulated before 1988.[226] A 1987 House committee investigating the quality of clinical laboratories concluded "that as much as 25% of all testing is done in laboratories which are either unregulated by the federal government or subject to diminished or ambiguous regulatory standards."[227]

The uneven application of regulation had a beneficial side-effect. An

---

226. With the Clinical Laboratory Improvement Act of 1967 ("CLIA"), 42 U.S.C. § 263a (1982), the federal government regulated all laboratories sending specimens in interstate commerce. The Social Security Act also subjected labs in the medicare program to federal regulation. 42 U.S.C. §§ 301–1397 (1982). CLIA exempted from federal regulation very small laboratories, most laboratories located in physicians' offices, and laboratories with local operations provided that they did not participate in medicare.

For an overview of how that regulation was conducted, see LaMotte & Robinson, *supra* note 20.

227. H.R. REP. NO. 899, 100th Cong., 2d Sess., 13–14, *reprinted in* 1988 U.S. CODE CONG. & ADMIN. NEWS 3828, 3833–35 [hereinafter HOUSE REPORT]. Not all labs that fell within the purview of CLIA were regulated by the federal government:

> Laboratories which are in hospitals accredited by the Joint Commission on the Accreditation of Hospitals (now the Joint Commission on the Accreditation of Health Care Organizations), or laboratories which are accredited by the College of American Pathologists . . . are exempted from the licensure requirement of CLIA. Private accreditation bodies which have standards equal to or more stringent than those of [CLIA] may be approved to accredit laboratories for the purpose of this exemption from CLIA.

Id. at 11. In addition, many states had regulations affecting clinical laboratories, but the House committee concluded, "While some of these state regulatory systems are comprehensive and rigorous, others are rather cursory and fail to assure appropriate levels of quality." *Id.*

experimental condition was created whereby the effect of regulation on the quality of laboratory performance could be meaningfully examined. Studies consistently discovered that unregulated, unlicensed laboratories performed lower quality work than licensed laboratories.[228] A survey of the relevant literature concluded,

> Each study . . . confirms that unregulated laboratories perform considerably less accurately than regulated laboratories. . . . Unregulated laboratories do not engage in sufficient quality control and they have more random error and less precision than is found in conventional clinical laboratories.
>
> The literature concludes that there is a significant decrease in the accuracy of laboratory tests conducted in unlicensed laboratories as compared to licensed laboratories and suggests that: "the legislation and enforcement of laboratory regulations improves the accuracy of clinical laboratory testing."[229]

In 1986 the Oregon Public Health Laboratory surveyed forty unlicensed and twenty licensed physician office laboratories. Even though the unlicensed facilities had all volunteered, and presumably, therefore, had confidence in their work, the quality control was considerably higher in the licensed laboratories than in the unlicensed ones. "By all measures employed in the study, licensed physician office laboratories performed better in proficiency testing programs and on internal index of performance [than] did unlicensed POL's. Clear evidence exists to show that laboratories inspected for state licensure produce less variability in laboratory testing."[230]

---

228. The House Report summarized the findings:

> [N]umerous studies . . . showed that the absence of regulation in physician office laboratories has resulted in diminished quality. Data collection in California . . . suggest that unregulated laboratories perform lower quality work than regulated labs. Studies in Idaho showed that, according to proficiency testing data, physician office laboratories improved when quality assurance standards were improved. Finally, a study of licensed and unlicensed physician office laboratories in Oregon showed that licensed laboratories performed better than unlicensed labs.

Id. at 14.

229. *Hearings, supra* note 20, at 338–40 (statement of American Clinical Laboratory Association) (quoting Grayson, *Effects of Regulatory Controls on the Accuracy of Clinical Laboratory Tests,* 1 J. MED. TECH. 633 (1984)).

230. *Id.* at 229 (statement of Virginia L. Worrest, Medical Consultant). *Cf. id.* at 242 (statement of Herbert W. Dickerman, New York State Department of Health) ("Numerous reports in the literature have demonstrated that voluntary proficiency testing simply does not perfect laboratory performance as well as a mandatory program does. When laboratories must perform to obtain an operating permit, greater incentive exists to achieve state-of-the art standards.").

Similarly, the Nebraska Society for Medical Technology found that fifty-five percent of the blood chemistry samples tested in unregulated labs fell outside of the acceptable range while the percentage was only thirteen percent at the accredited facilities.[231] Another study discovered that laboratories employing only certified medical technologists had a mean accuracy score of ninety-five percent on proficiency tests. Labs with all noncertified technologists recorded seventy-five percent accuracy.[232]

These studies demonstrate that regulation improves clinical laboratory performance better than noncompulsory programs. Perhaps, a laboratory director best summarizes the situation: "Most of us would prefer that regulation be on a voluntary basis. But frankly, voluntary systems don't work worth a damn in the long run."[233]

Congress responded to this evidence[234] by enacting the Clinical Laboratory Improvement Amendments of 1988, which mandate regulation of all clinical labs.[235]

---

231. *See id.* at 194 (statement of the American Society for Medical Technology).

232. *See* Lunz, Castleberry, James & Stahl, *The Impact of the Quality of Laboratory Staff on the Accuracy of Laboratory Results*, 258 J. AM. MED. A. 361 (1987). This study also found that the greater the percentage of certified technologists employed, the higher the lab's accuracy.

Patricia Ashton of the American Society for Cytotechnology testified summarizing a study showing that proficiency test scores had improved in regulated labs over a fourteen-year program in New York. "The study concludes that periodic inspection and testing of laboratories, withholding of permits from substandard facilities, and especially a comprehensive program of continuing education, have contributed to this improvement...." *Hearings*, *supra* note 20, at 147 (quoting Collins & Patacsil, *Proficiency Testing in Cytology in New York*, 30 ACTA CYTOLOGICA 633 (1986)).

233. Bogdanich, *Medical Labs, Trusted as Largely Error-Free, Are Far From Infallible*, Wall. St. J., Feb. 2, 1987, at 1, col. 6, (statement of Walter Johnson, the director of the cytogenetics lab at Case Western Reserve University).

234. It did take considerable time and effort for the new legislation to be enacted:

> Since Congress passed its laboratory legislation 20 years ago, it has tried five time to strengthen the law. Each effort, opposed by organized medicine, failed.... Growing evidence of testing abuses, along with fear of malpractice litigation, has begun to stir concern in some quarter of the medical establishment. Since late 1985, the Journal of the American Medical Association has been running a series of articles on office testing. Physicians, the authors have written, "are unfamiliar with ... the subtleties of testing, the idiosyncratic personalities of 'foolproof' electronic equipment, or even the basic concepts of quality control."

Id. at 14.

235. Clinical Laboratory Improvement Amendments of 1988, Pub. L. No. 100–578, 102 Stat. 2903 (1988). The House concluded that not only were regulated labs superior to the nonregulated facilities, it also found that the existing regulatory structure was "a seriously flawed system for assuring compliance and an inadequate system of enforcement of Federal standards...." HOUSE REPORT, *supra* note 227, at 12. The 1988 Amendments not only extended regulation to more laboratories, they also closed loopholes in the compliance and enforcement mechanisms. Mandating stricter compliance with existing regulations to pro-

Nothing in the data indicates that the positive effect regulation has on laboratory quality is limited to the clinical field. In fact, since the unregulated forensic laboratories, which bear many similarities to clinical labs, perform poorly, a regulatory scheme imposed on crime labs can improve the quality of forensic science.

Such regulation has been advocated in the area of DNA typing or profiling. The State of New York recently convened the Forensic DNA Analysis Panel (the "Panel") to study the implementation of the new technology. This group, consisting of forensic science lab directors, prosecutors, defense attorneys, police officers, judges, and legal and scientific academics, noted:

> There are no widely accepted criteria for quality control or proficiency testing for forensic laboratories at a state or national level. Concern is mounting in the scientific community that the forensic laboratories performing DNA typing are not following all of the necessary and appropriate practices. If proper quality control procedures are not used, the reliability of the data produced is questionable.[236]

The Panel recommended that "a state accreditation process be developed to monitor public and private laboratories providing forensic DNA analysis services in New York state." The Panel further recommended that only DNA typing evidence performed by accredited laboratories be admissible.[237]

Legal and scientific scholars also contend that mandatory quality assurance standards are necessary for DNA typing. For example, Professor Alan Westin concludes, "Both government and private laboratories doing DNA analysis for forensic use should meet designated quality standards and be subject to periodic inspections."[238]

---

duce better scientific testing is consistent with a recommendation for pharmaceutical products. New drugs must undergo clinical trials to demonstrate the new products' safety and efficacy. An examination of scientific misconduct in these trials brought forth the recommendation that the Food and Drug Administration "should continue or even expand its monitoring program [of these trials], while ensuring its rigor." Shapiro & Charrow, *supra* note 115, at 736.

236. J. POKLEMBA, *supra* note 50, at 15.

237. *See id.* at 46, 49.

238. Westin, *supra* note 212; *cf.* Lander, *supra* note 172, at 505 ("[T]here is an urgent need for the scientific community to agree on clear guidelines for the procedures and standards needed to ensure reliable DNA fingerprinting. Legislators should also consider whether licensing and proficiency testing should be required in forensics."). *See also* Comment, *supra* note 156, at 701 ("To safeguard the performance and evidentiary value of DNA printing, legislatures must regulate crime laboratories' performance.").

In spite of pleas for mandatory quality assurance standards, no jurisdiction requires DNA laboratories to meet performance standards. Members of the forensic science community strongly resist DNA laboratory regulation. Two forensic scientists contend:

> In principle, introduction of DNA analysis results is no different from the introduction of latent fingerprints, gunshot residue analysis, bullet and cartridge case comparisons, or comparisons of Fourier transform infrared (FTIR) spectra of fibers, which are all different forms of pattern recognition and comparisons. Provided reliable and reproducible techniques are used, comparison of DNA autoradiographs is no different than comparisons of gas chromatograms or striations on bullets.
>
> There appears to be no need or reasons for any special regulation, licensing, or oversight of forensic scientists or laboratories because of the introduction of DNA typing. No one is in a better position to understand the scientific and legal requirements of evidence testing than forensic scientists themselves.[239]

This position makes sense. As these scientists suggest, other forensic procedures are demanding. Why regulate DNA typing when the rest of forensic science is not subjected to regulation? The New York panel concluded, it is true that "[q]uestions about the quality of the work being done by the private [DNA] laboratories have not been satisfactorily answered, and the laboratories' adherence to accepted scientific procedures has not been demonstrated."[240] That does not separate DNA typing from other forensic procedures. Since crime labs are unregulated, no forensic laboratory is routinely required to demonstrate its adherence to accepted scientific procedures. Furthermore, the importance of DNA testing cannot be used as justification for its being the only forensic procedure which is regulated. While DNA typing may have great value to some criminal cases, other forensic techniques are used more often and affect more cases.[241]

---

239. Lee & Gaensslen, *The Need for Standardization of DNA Analysis Methods*, in BANBURY REPORT 32: DNA TECHNOLOGY AND FORENSIC SCIENCE 212 (J. Ballantyne, G. Sensabaugh & J. Witkowski eds. 1989). Jack Ballantyne, a forensic scientist with the Suffolk County, New York, Medical Examiner, has also stated, "I do not believe that the quality assurance procedures should be any more stringent for DNA than for any other serological evidence." BANBURY REPORT, *supra* note 159, at 244.

240. J. POKLEMBA, *supra* note 50, at 26.

241. *See* Peterson, *Impact of Biological Evidence on the Adjudication of Criminal Cases: Potential for DNA Technology*, in BANBURY REPORT 32: DNA TECHNOLOGY AND FORENSIC SCIENCE 55 (J. Ballantyne, G. Sensabaugh & J. Witkowski eds. 1989). *See also* Ballantyne, *DNA Technology in a County Forensic Laboratory Setting*, in BAN-

These similarities raise the question of why all forensic laboratories are not required to satisfy quality assurance regulations.

The federal government's response to this question has led to regulation in only one area — urine drug testing for federal employees and for workers in the transportation industry.[242] To assure that urine drug tests of employees are as accurate as possible, laboratories must be certified and must meet continuing performance standards to maintain that certification.

The federal government regulates and monitors clinical laboratories and drug testing facilities analyzing workers' urine. Why, then, are crime laboratories not subjected to similar mandatory quality assurance programs? Arguments against regulation are not persuasive. First, some may argue that forensic labs in comparison to the other facilities are not important enough to require a regulatory structure. Admittedly, health and jobs are important, but freedom and imprisonment are equally so. It is also true as some argue that clinical laboratories affect significantly more people than do crime laboratories. Although most people probably have a clinical procedure done every one or two years, very few have been involved with forensic scientists. But to conclude, therefore, that the quality of forensic laboratories is unimportant is like contending that the abilities of judges hardly matters because only a fraction of the populace is involved in litigation, or that the quality of police is trivial because most of us are not arrested.

Forensic science is significant. A recent survey found reports from

---

BURY REPORT 32: DNA TECHNOLOGY AND FORENSIC SCIENCE 212 (J. Ballantyne, G. Sensabaugh & J. Witkowski eds. 1989):

> DNA typing methods are not expected to dramatically increase the number of cases where blood is found to be associative with respect to the defendant and the crime scene because . . . a powerful set of conventional genetic markers are available to perform the task of excluding the majority of the population. However, the converse is true for seminal stain evidence, in that a dramatic increase in the number of rape cases is expected where the suspect is associated with the victim.

In fact, "the bulk of [crime lab] casework concerns the identification of drugs and alcohol." Peterson, Ryan, Houlden & Mihajlovic, *supra* note 12, at 1733.

242. *See* Department of Health and Human Services Mandatory Guidelines for Federal Workplace Drug Testing Programs, 23 Fed. Reg. 11,970 (1988); Procedures for Transportation Workplace Drug Testing Programs, 53 Fed. Reg. 47,002 (1988). The urine analyses are similar to tests done in crime labs. *See also* Caplan, *Drug Testing in Urine*, 34 J. FORENSIC SCI. 1417, 1418 (1989) ("Urine drug testing should be considered a special application of analytical forensic toxicology.").

crime laboratories present in about a third of all criminal cases.[243] Forensic science enhances the accuracy of fact-finding.[244] It certainly affects verdicts. Jurors believe that they comprehend scientific evidence as well as or better than other evidence.[245] Furthermore, "[a]bout one quarter [of jurors who] were presented with scientific evidence believed that had such evidence been absent, they would have changed their verdicts—from guilty to not guilty."[246] Forensic science, thus, determines whether people are jailed as well as whether the guilty are mistakenly freed.

Accurate forensic science is essential to justice, but abysmal quality remains widespread. The quality of forensic science must be improved. Since regulation can produce better performance, forensic laboratories should be regulated.

VI. MODELS FOR THE REGULATION OF FORENSIC SCIENCE

A. *Clinical Laboratory Improvement Act as a Model*

Just as regulation has improved the performance of clinical laboratories, regulation can force improvements in forensic science laboratories. The scheme embodied in CLIA, which has proved successful for the regulation of clinical labs, provides a model for improving forensic labs.[247]

---

243. *See* Peterson, Ryan, Houlden & Mihajlovic, *supra* note 12, at 1733. That study found that murder and drug prosecutions almost always have lab reports. The use of scientific evidence in criminal cases has grown rapidly over the last two or three decades. Estimates in the early 1960s indicated that fewer than one percent of criminal cases received a forensic science examination. *See id.* at 1731. It has been predicted that "the use of scientific evidence will . . . likely increase." Giannelli, *supra* note 9, at 1200. A recent multi-jurisdiction survey of the use of forensic science found, however, that the usage rates were fairly steady from 1975 to 1981. Peterson, Ryan, Houlden, & Mihajlovic, *supra* note 12, at 1734.

244. *Symposium, supra* note 5 (comment of Edward Imwinkelried).

245. *See* Peterson, Ryan, Houlden & Mahajlovic, *supra* note 12, at 1748. *See also* Note, *supra* note 225 (discussion of empirical studies examining the effect of scientific evidence on jurors).

246. Peterson, Ryan, Houlden & Mihajlovic, *supra* note 12, at 1748. This survey also discovered that the presentation of forensic evidence correlates with increased prison sentences. For example, "the presence of a laboratory report adds about 23 months to attempted murder sentences, 27 months to robbery terms, and 4 months to theft sentences." *Id.* at 1746.

247. "The regulation of forensic laboratories has an excellent model, the Clinical Laboratories Improvement Act. . . . The law was enacted to ensure that such service laboratories, which are not subject to the same peer scrutiny as research laboratories, would nonetheless provide reliable products and service." Neufeld & Colman, *supra* note 157, at 53.

Under that law, all clinical laboratories must be certified.[248] Certification requires compliance with standards "designed to assure that such laboratories will consistently perform tests in a valid and reliable manner."[249] There are four major components to the program: "maintenance of a quality assurance and quality control program by the laboratory; maintenance of appropriate records, equipment and facilities; personnel standards; and proficiency testing."[250] Although all the standards are meant to further accuracy in clinical lab results, "proficiency testing [is] the central element in determining a laboratory's competency, since it purports to measure actual test outcomes rather than merely gauging the potential for accurate outcomes."[251] A lab must undergo quarterly testing for each type of analysis that it performs.[252] Most important, the testing must be done in a blind fashion[253] with the results made available to the public.[254]

In order to ensure that a laboratory routinely uses proper quality control measures, the laboratories can be inspected with or without prior

248. No laboratory can do any clinical analyses unless it receives a certificate issued by the Secretary of Health and Human Services at least every two years. 42 U.S.C. § 263a(b) & (c).

249. HOUSE REPORT, *supra* note 227, at 3848.

250. *Id.* The law requires the Secretary of Health and Human Services to issue standards for these four components and any other desirable ones. 42 U.S.C. § 263a(f)(1)(A)-(E).

251. HOUSE REPORT, *supra* note 227, at 3849.

252. "The standards shall require that a laboratory issued a certificate ... be tested for each examination and procedure ... for which it has received a certificate.... The testing shall be conducted on a quarterly basis, except where the Secretary determines for technical and scientific reasons that a particular examination or procedure may be tested less frequently (but not less often than twice per year.)" 42 U.S.C. § 263a(f)(3)(A).

253. To be certified, a laboratory must agree "to treat proficiency testing samples in the same manner as it treats materials derived from the human body referred to it for laboratory examinations or other procedures in the ordinary course of business." *Id.* § 263a(d)(1)(E). The House Report concluded, "A significant deficiency in the current proficiency testing regime is its inability to assure that proficiency samples are treated like patient samples.... The only way to guarantee that samples are treated by the same personnel, at the same speed, using the same equipment as the patient specimens is [through] blind or on-site proficiency testing." HOUSE REPORT, *supra* note 227, at 3837. The drafters were concerned that some labs "might run repeated tests on the sample, use more highly qualified personnel than are routinely used for testing, or send the sample out to another laboratory. Such practices obviously undermine the purpose of proficiency testing and [§ 263a(d)(1)(E)] seeks to prevent them through this agreement." *Id.* at 3845.

254. "The Secretary shall establish a system to make the results of the proficiency testing programs ... available, on a reasonable basis, upon request of any person." 42 U.S.C. § 263a(f)(3)(F).

notice.[255] The drafters of the law expected frequent audits, many of them unannounced.[256]

Finally, the law provides a range of sanctions. The Department of Health and Human Services can suspend, revoke, or limit a clinical laboratory's certification.[257] It can also take intermediate steps, such as fines, to induce compliance.[258]

Congress concluded that this rigorous, mandatory quality control program was necessary after studying the performance of clinical laboratories and the positive effects regulation can have on the quality of those labs.[259] If such a regime is necessary for clinical labs, a similar regulatory scheme — involving inspections, personnel standards, quality control, and external proficiency testing[260] — seems in order to improve the endemic poor quality of forensic laboratories. However, regulation of crime laboratories, raises several difficulties not present with the regulation of clinical facilities.

First, who is to devise, implement, and enforce such crime laboratory quality control programs? The present standards for federal regulation of clinical facilities build upon decades of national regulation and an

---

255. "The Secretary may, on an announced or unannounced basis, enter and inspect, during regular business hours of operation, laboratories which have been issued a certificate under this section." *Id.* § 263a(g)(1). To be certified, a laboratory must agree to permit the inspections and make its records available to the Secretary of Health and Human Services. *Id.* § 263a(d)(1)(C) & (D).

256. "It is the Committee's intent that such inspection normally be conducted every two years, but the Secretary could establish criteria or guidelines for more frequent or less frequent inspections, depending on the identified needs and risks involved." HOUSE REPORT, *supra* note 227, at 3846. "The Committee expects the Secretary to have many of these inspections conducted on an unannounced basis, in order to observe the laboratory during typical operations, rather than under ideal circumstances arranged in preparation for a prearranged inspection." *Id.* at 3853.

257. 42 U.S.C. § 263a(i)(A)-(G).

258. *Id.* § 263a(h). "This provision would authorize the Secretary to impose one or more of a range of sanctions on a laboratory which is no longer in substantial compliance with [CLIA]. The intermediate sanctions include directed plans of correction, civil monetary penalties for each day of non-compliance, and payment of the costs of onsite monitoring." HOUSE REPORT, *supra* note 227, at 3853.

259. The federal government imposed even more strict regulations on laboratories doing employee urine drug analyses. *See supra* text accompanying note 242. Laboratories doing urine drug analyses for government employees or for employees in the transportation industry must be certified. 53 Fed. Reg. 47,013 (1988). Before certification, a laboratory must undergo three cycles of blind testing without producing any false positives. *Id.* at 47,015. False positives on proficiency testing after certification can result in certification suspension or revocation. The regulations also require the laboratories to have a procedures manual and a quality assurance program for all phases of testing. *Id.* at 47,011.

260. *See Hearings*, *supra* note 20, at 345–46 (statement of American Clinical Laboratory Association) ("These components are interdependent and no one element is more important than another.").

even longer history of mandatory state programs.[261] Consequently, federal regulators have considerable information for determining appropriate personnel standards, in-house quality control programs, record-keeping requirements, and appropriate proficiency testing programs. Potential forensic laboratory regulators, on the other hand, lack such data. Little thought has been given to forensic personnel standards and quality control programs. Consensus does not even exist over such basic matters as protocols for routine analyses. Regulation of forensic labs truly must start from scratch. The imposition of a rigorous, comprehensive, mandatory scheme will require extensive study and effort.

A CLIA-type regulatory scheme requires well-staffed institutions, especially for monitoring in-house quality control programs and for on-site inspections. To meet these demands, the Secretary of Health and Human Services does not have to depend solely on his own department; the Secretary can also approve private nonprofit organizations as accreditation bodies if those organizations use standards at least as stringent as those required under CLIA.[262] Such organizations exist and have been effectively enforcing their own accreditation standards for decades.[263] Similarly, the Secretary can exempt from federal regulation laboratories regulated by states with stringent standards.[264] Many states have such regulation. This situation contrasts with that of forensic labs. Since no institutions currently regulate crime laboratories, the source of potential forensic regulators is not clear. New organizations would have to be created or significant new duties would have to be given to existing organizations.

Besides a workforce, experience and knowledge, money will be necessary for forensic laboratories to be well regulated. For clinical regulations, congress found a politically expedient, if somewhat disingenuous, method of funding. Laboratories are charged fees for certification.[265] This allows the regulation of clinical laboratories

---

261. Clinical labs have been regulated by the federal government since 1967. For the history of that regulation, see LaMotte & Robinson, *supra* note 20, at 39; HOUSE REPORT, *supra* note 227, at 3831–32. New York City and a number of states had regulated clinical facilities before this. *See* Hamblin, *Growth and Development in California*, in FEDERAL LEGISLATION AND THE CLINICAL LABORATORY 14 (M. Schaeffer ed. 1981) (brief overview of state and local regulation).

262. 42 U.S.C. § 263a(e)(2)(A)(ii).

263. For example, the College of American Pathologists has a rigorous, voluntary accreditation program for clinical labs. LaMotte & Robinson, *supra* note 20, at 40.

264. 42 U.S.C. § 263a(p)(2).

265. "The Secretary shall require payment of fees for the issuance and renewal of certificates." 42 U.S.C. § 263a(m)(1).

without the use of tax money, although the public ultimately pays for the regulatory scheme.[266] A legislature, however, will not find such a convenient method of funding forensic regulations. Overwhelmingly, crime labs are public agencies that cannot pass additional costs on to insurance companies or consumers. Public money must be directly allocated if forensic facilities are to be regulated.

Finally, CLIA-style sanctions also pose problems for the regulation of forensic science. Fines imposed on crime laboratories would almost always have to be paid out of the public till. It is unlikely that a regulator would so sanction a governmental crime lab. Even if this occurred, the penalty would probably not have much deterrent or rehabilitative effect on an organization that does not make money. Regulators are also unlikely to suspend or revoke an accreditation necessary for a forensic laboratory to operate or present its results in court. Proficiency testing has taught that "accreditation on the basis of adherence to certain standards of accuracy ... would apparently mean that a good number of laboratories would be shut down."[267] Often, a major city may be dependent on a single crime lab, and its forced closing would have a huge effect on criminal justice. It would take great political will to enact and enforce such regulation even if it improved the accuracy of forensic science.[268]

While the pervasive poor quality of forensic science cries out for regulation like that imposed on clinical facilities, the difficulties in devising and implementing such regulations, as well as the fierce resistance by the forensic science community to any kind of enforced quality control, may make needed regulation politically infeasible. Nevertheless, less rigorous regulations may present fewer difficulties and still provide beneficial results.

B. Mandatory Proficiency Testing

A certification program that only required labs to participate in proficiency testing could be a significant step towards achieving better crime lab quality. Such a program would not impose personnel

266. "Given rising federal expenditures for health care and the scarcity of federal money to carry out enforcement activities, the Committee has determined that the costs of such regulation should be borne by the laboratories.... The Committee understands that consumers bear the ultimate costs of regulation and notes that the proper performance of tests not only protects patients from undergoing unnecessary and dangerous treatment, but ultimately saves money." HOUSE REPORT, *supra* note 227, at 3858.

267. M. SAKS & R. VAN DUIZEND, *supra* note 25, at 97.

268. *Id.*

standards, require in-house quality control measures, prescribe methods of analyses, or mandate a certain level in proficiency testing accuracy. Such limited regulation would certainly be a big departure from an ideal program,[269] but the restricted scheme would not encounter many of the problems of CLIA-type regulations.

Limited regulation avoids the difficult problem of creating institutions to devise and implement a thorough, mandatory quality control program. The institution of good proficiency examinations alone will be much easier than the imposition of an entire quality control program. Of course, the creation of a comprehensive proficiency testing program will take effort, but it can draw upon the knowledge gained from past and continuing forensic testing programs. Similarly, the proficiency component will require fewer people[270] and less money[271] than an entire quality control program.

Sanctions in a limited program would also pose fewer problems than in a full quality control program. A laboratory would not be forbidden from doing analyses or from presenting evidence in court because it did not have appropriate methods manuals, because it was not staffed or run

---

269. Valuable unannounced audits would also be given up. *Cf.* LaMotte & Robinson, *supra* note 20, at 41 ("The internal quality control systems of CLIA laboratories are evaluated by annual, unannounced, on-site inspections. By not announcing their visits, examiners get a truer picture of a laboratory's routine quality control system."). *See also Hearings, supra* note 20, at 115 (comments of Dr. Loyd R. Wagner, Vice-President, College of American Pathologists):

> Proficiency testing is the testing of an event in time, and it simply reflects what that laboratory did on that particular situation. It is one of the measures of laboratory quality, but it cannot be utilized alone to assess whether or not a laboratory is performing quality work. There are other things like quality control, space, other activities that go along with that, that better assess quality.

Such an abbreviated regulatory system would also not require some easily instituted procedures, now seldom in effect, that would guarantee improved accuracy such as replication of the test by another analyst or at least an independent reading of the same test by another analyst. *See supra* note 204 for the effect of such replication on error rates.

270. *Cf.* Pereira, *supra* note 74, at 6: ("The operation of a reliable [forensic] quality assurance system has heavy manpower resource implications.... One might ask 'Can we afford it?'—I think the more appropriate question is 'Can we afford to be without it?'").

271. Although I have not tried to calculate how much money it would take for a comprehensive proficiency testing program, the testing done by the Collaborative Testing Services gives some idea of costs. In 1989, Collaborative Testing Services charged $52 for each proficiency test. BROCHURE, *supra* note 28. There are almost 350 forensic laboratories in the country. Lee & Gaensslen, *Forensic Science Laboratory/Forensic Science Program Cooperation and Relationships: The View From the Forensic Science Laboratory*, 33 J. FORENSIC SCI. 1071, 1071 (1988) ("Today, there are approximately 40 federal and 295 nonfederal forensic science laboratories in the United States."). All these labs do not do analyses in all of the possible forensic areas, but if each lab is tested quarterly in twenty different categories, then the total cost would be about $1,500,000.

by people with a certain education, or because the accuracy of its results did not reach a predetermined level. Thus, sanctions would not be imposed because an outside body had determined how a laboratory should operate or had determined that the lab's proficiency was too low, but only on the objective ground that a facility failed to participate in proficiency testing.

Proficiency testing can significantly raise the quality of forensic science, but the testing must be mandatory for all crime laboratories. As we have seen, voluntary proficiency testing is not sufficiently effective.[272] Particularly, noncompulsory testing for forensic facilities is undersubscribed.[273] The starting point for improving forensic laboratory performance is fuller knowledge about the problems. Complete proficiency testing of all laboratories is essential to gain that knowledge.[274]

The proficiency testing must also be blind.[275] The important issue for criminal justice is not how accurately laboratories perform when they are aware of being tested, but how well they do on actual cases.[276] We can draw few useful inferences about real casework by administering trials

272. "Numerous reports in the literature have demonstrated that voluntary proficiency testing simply does not perfect laboratory performance as well as a mandatory program does." *Hearings, supra* note 20, at 242 (statement of Herbert W. Dickerman, New York State Department of Health).

273. *Cf. id.* at 368 (statement of Mark Birenbaum, American Association of Bioanalysts) ("Voluntary proficiency testing programs for [physician office laboratories] sponsored by the medical profession have been notoriously under-subscribed for years."). Of course, those labs that do participate in proficiency testing are atypical; they are usually the better ones. *See supra* text accompanying notes 19–20.

274. As the British have noted, a forensic proficiency program should also test all the lab functions, not just its scientific analyses: "The whole system is tested; not all problems which have come to light related to scientific technique; some aspects of the work are particularly prone to clerical error." Pereira, *supra* note 74, at 4. *Cf.* Caskey & Hammond, *supra* note 177, at 127, 129 (contending that human error is the most common problem in the use of DNA based technology and that the experience in medical genetic diagnoses shows that chain of custody difficulties are the most frequent). *See also Hearings, supra* note 20, at 21.

275. "Passing a battery of blind tests should be required before a crime laboratory is allowed a license to make determinations that affect an individual's freedom." Note, *supra* note 50, at 494. *See* Thompson & Ford, *supra* note 197, at 69 ("General concerns about the difficult nature of the [DNA profiling] procedure and its susceptibility to error can most likely be addressed only through blind proficiency testing of the analysts who actually perform the tests in forensic laboratories."); Note, *supra* note 188, at 928 ("Periodic evaluations of laboratories, including 'blind tests,' should be performed by independent, disinterested experts to ensure that the results of the DNA tests are accurate and that lab protocols are accurate and that the lab protocols are properly implemented.").

276. *Cf.* Lunz, Castleberry & James, *The Impact of the Quality of Laboratory Staff on the Accuracy of Laboratory Results,* 258 J. AM. MED. A 361, 361 (1987) ("The ultimate test of (laboratory) competence is the ability to perform the task accurately in the appropriate environment.").

that the analysts know to be tests. Studies indicate that performance will be better on known examinations than on either blind tests or real casework.[277] To learn about the accuracy and reliability of lab work, forensic facilities must be subjected to blind testing that simulates real cases as much as possible.[278]

In addition, the results of the proficiency testing must be made public. The true state of forensic science will not be known until the testing results from all laboratories are disclosed. This information will not only reveal the extent of the problems, but also how, for example, education and training of analysts, size of budgets, and kinds of equipment correlate with performance. Such data is essential for determining the necessary steps for improvement.

The dissemination of such information should also serve as a direct spur to improvements. The Clinical Laboratories Improvement Amendments, as we have seen, make testing results public. The drafters of the Amendments recognized that

---

277. *See supra* note 22. *See also* BANBURY REPORT, *supra* note 159, at 247 (comments of British forensic scientist David Werett).

278. A forensic scientist, in explaining the quality assurance program in Great Britain, has also explained the importance of blind proficiency testing:

> Our trials fall under two distinct headings; these are declared trials where participants receive material to test and are fully aware that they are under scrutiny and undeclared trials which enter laboratories disguised as genuine cases.... Undeclared trials are generally regarded as more profitable. With these, not even Directors of the laboratories are told about them until after the event. They are much more difficult to set up but have the advantage of testing the whole system from the documentation on receipt at the laboratory to the quality of the notes and the scientific work, the way in which the statement is laid out and the information conveyed and also the time taken to complete the case.... Because this type of trial is so highly appreciated, [the Home Office Central Research Establishment] is devoting more of its effort in this direction.... This system of trials enable the Controller, the Directors and senior members of the laboratories to evaluate the general performance of the Service and to identify any particular area of weakness.

Pereira, *supra* note 74, at 3–4.

The view expressed in the House Report is similar:

> [P]roficiency testing ... is considered one of [the] best measures of laboratory performance. It is arguably the most important measure, since it reviews actual test results rather than merely gauging the potential for good results....
>
> A significant deficiency in the current proficiency testing regime is its inability to assure that proficiency samples are treated like patient specimens.... The only way to guarantee that samples are treated by the same personnel, at the same speed, using the same equipment as patient specimens is [through] blind or on-site proficiency testing.

HOUSE REPORT, *supra* note 227, at 3836–37.

> [t]he release of such information would help physicians select laboratories on the basis of quality. Second, it may serve as a deterrent to laboratories violating standards. Third, it would give consumers and their representatives an opportunity to assess the performance of individual laboratories on an objective basis. Finally, release of such information would stimulate constructive debate about, and result in development of a consensus on, proper measures of quality.[279]

Public information about forensic proficiency testing can have much the same effects. As programs in other disciplines have established,[280] publication of proficiency testing results will give crime labs the data they need to improve themselves. When the proficiency results are available, a laboratory can see where other facilities are performing better, and seek information to produce more accurate analyses. For example, the lab may find out that while it is using one method of analysis, other facilities are using a different technique with consistently better accuracy.[281]

Widespread dissemination of the test results will allow all scientists to

---

279. HOUSE REPORT, *supra* note 227, at 3859. As discussed *supra* note 254 and accompanying text, CLIA requires that the results of proficiency testing be available to the public. Furthermore, the Secretary must make additional information public that would be useful for evaluating lab performances, including whether a lab had been subject to sanctions, denied participation in federal programs, or been convicted of fraud or billing abuses. 42 U.S.C. § 263a(n) (1988).

280. *See, e.g., Hearings, supra* note 20, at 147 (statement of Patricia Ashton, American Society for Cytotechnology) ("The [clinical] proficiency testing program in Ontario, Canada showed that test results pinpointed specific areas where continuing eduction was needed. Program officials developed and distributed educational programs for these special problems."). *See also* Mason, *Effect on Clinical Chemistry*, in FEDERAL LEGISLATION AND THE CLINICAL LABORATORY 110 (M. Schaeffer ed. 1981) ("One of the noteworthy advances in clinical chemistry that would have been virtually impossible in the absence of widespread proficiency testing is the development of data bases useful for documenting the level of overall performance for individual tests.").

281. *See* Gilbert, *Progress and Analytic Goals in Clinical Chemistry*, 63 AM. J. CLINICAL PATH. 960 (1975). Gilbert examined five years of data from the clinical laboratory proficiency testing done by the College of American Pathologists, noted that different methods of analyzing the same substances had different accuracy rates, and concluded:

> the major potential route for further improvement lies in encouraging widespread method change, with discontinuance of methods with poor precision and increased utilization of methods with good precision. The most widespread problem appears to be the need to eliminate poor methods, and it is hoped that data from the Survey program can play a major role in making participating laboratories aware of the need for change in providing reliable data that can be used as a guide in the selection of better methods.

Id. at 972–73.

examine the data to see where further studies would be most fruitful. This would improve use of resources, which is particularly important in a field where so little research is done. A thorough examination of findings may reveal that errors consistently occur in certain analyses, identifying important areas for investigation where present procedures are not particularly precise.[282] For example, in a 1986 proficiency test for GLO system typing, the false positive rate unexpectedly reached nearly fifty percent.[283] Since the analytical mistakes on the typing of other systems did not approach this level, the data suggest a particular problem in typing the GLO system that should prompt more research. Wide dissemination of proficiency testing data will enable as many people as possible to scrutinize the information, find possible problems, and seek solutions.

Furthermore, proficiency testing results must be made public to help courts make informed rulings on the admissibility of forensic evidence. Since the reliability of a technique is a component of admissibility, courts should examine proficiency testing results when evaluating scientific evidence. In addition to assessing scientists' and technicians' assertions about a scientific procedure, courts should have all the pertinent data on those issues. The outcomes of mandatory proficiency testing should be a prime source of that information.[284] If those results, for

282. *Id.* at 972. Gilbert found that the proficiency testing revealed that analyses of calcium and magnesium produced less accurate and reliable results than analyses of other substances. He concluded:

> The reason[ ] the results fall short of the desired goals is not clear, and the analysis warrants further study.... The data from the Survey program suggest that the [accuracy] goal is beyond the present level of technology.... It is difficult to believe that the failure to achieve this goal represents incompetent performance on the part of the practicing laboratories.

Id.

283. "Since the blood samples were drawn from only two sources, the groupings in the other systems confirm that there was no sample mix-up. Degradation in the GLO system with time is known to occur. The samples should have been at most only 7–10 days old when received by the labs.... Although the manufacturer's experience is that degradation does not begin until about 5 weeks for dried stains, even when it does occur a GLO 1 should not be confused with a 2–1." CTS 86–11, at 1–2.

284. The courts should also be aware of the difficulty or impossibility of designing meaningful proficiency testing for some areas of "scientific" expertise. Judges should realize that in such fields the opinion is, at best, based on deductions, not on empirical science. Bite mark evidence provides an example. *See supra* notes 130–35 and accompanying text for a discussion of bite mark identification research. How can a meaningful proficiency program be designed to test those who claim the ability to identify bite marks? Interestingly, a defense expert in State v. Sager, 600 S.W.2d 541 (Mo. Ct. App. 1980), did try to verify his testimony empirically. He concluded that defendant's cuspid would have left a mark that was not present in the pictures of the victim and performed an experiment: "In the morgue, we found a female cadaver of rather small breast size, and ... we pressed the

example, showed persistent, widespread difficulties in typing the GLO factor in bloodstains, courts should weigh this information in determining whether that typing is admissible.

Of course, such benefits as understanding the extent of crime laboratory problems, learning the best areas for research, and informing the courts how well a particular test has been done could all be achieved by releasing composite proficiency testing data, and keeping individual lab performances confidential, as the LEAA study did. That secretive system, however, would prevent other important advances. As the drafters of CLIA recognized, public disclosure of how labs perform can provide a powerful incentive to improve quality. If the identities of the laboratories are kept confidential, poorly performing facilities could not directly communicate with those doing better; thus, they would be deprived of the most efficient way of learning how to do better. Perhaps more important, forensic laboratory funding agencies ought to know the quality of individual facilities. With that information, government can expand facilities that perform well at the expense of those that do not. More analyses will be done accurately as a result.

The criminal justice system also needs to know about the quality of individual laboratory performances, both to spur more accurate and reliable performance and to dispense justice. Just as physicians and consumers with appropriate information can produce a greater number of correct analyses by choosing the better clinical facilities, consumers of forensic analyses can similarly affect quality when they know how indi-

---

articulated models [of defendant's teeth] into the breast of the cadaver material as firmly as we could.... Mark Sager's teeth left cuspid marks on that cadaver." *Id.* at 567. Such a demonstration, however, only highlights the difficulty of devising a meaningful proficiency program in this area. *Cf.* Whittaker, *Some Laboratory Studies on the Accuracy of Bite Mark Comparison*, 25 INT'L DENTAL J. 166 (1975). To test the accuracy of bite mark identifications, bites were placed in prepared pig skin placed over a rubber cylinder approximating an arm. This study found that when pictures were taken immediately after the bite, examiners could only make correct identifications about 70% of the time. When pictures were taken one hour after the bite, accuracy fell to 35%, and plummeted to 9–16% with pictures taken twenty-four hours after the bite. *Id.*

Cf. Risinger, Denbeaux & Saks, *supra* note 20, at 774 n.186 (discussing whether person claiming to be able to identify wines should be allowed to testify). They note that many people like to claim such expertise, and while that may be socially acceptable, "nobody ought to be subject to determination of legal rights without greater validation than the assertion of the practitioner." *Id.* This truth is not invalidated because a wine-tasting society or a journal attests to the skill's of the society's members or because the wine tasting invokes technical-sounding terms in testimony. "Either the proposed witness must be applying techniques that are replicable, and therefore available for validation to all who would take the trouble (which is not the case), or some blind testing validation would appear to be an absolute necessity." *Id.*

vidual laboratories and analysts perform.[285]

Juries should also know how particular laboratories and analysts perform on proficiency testing. When jurors determine what weight to give scientific evidence, they should have as much pertinent information as possible. Jurors' conclusions about the weight to be given to a fingerprint expert's opinion may logically change when they learn that the expert made mistakes ten percent of the time on proficiency testing. Without disclosure of those proficiency testing results, a proper assessment of weight cannot be made.[286] Better verdicts will be obtained if the results of proficiency tests are known.[287] Furthermore, attorneys,

---

285. Prosecutors aware that a laboratory they use performs poorly should exert pressure on that laboratory to improve its performance, and if they have the option, make a consumer choice to have more reliable facilities do their analyses. Without identification of how individual laboratories perform, prosecutors cannot take these quality-improvement actions.

Since defense attorneys are not as frequent consumers of forensic science services as prosecutors, defense attorneys have fewer opportunities to spur quality through their choices. Nonetheless, without proficiency testing information, they cannot exert what influence they have.

> The coexistence of a high degree of proficiency in many laboratories with the endemic presence of "unacceptable proficiency" in others causes both prosecutor and defense attorney to be faced with the dilemma of which laboratory is which. The defense attorney especially has no way of knowing whether he is dealing with a credible or a defective laboratory result.

Bradford, *supra* note 11, at 905.

286. *Cf.* Grunbaum, *supra* note 67, at 502–03 ("At present, no mechanism exists by which one can distinguish between the analysts offering reliable testimony and those whose work is scientifically unacceptable.").

287. It has been observed:

> In order to determine the significance of test results, fact-finders should be presented with information regarding the base rates for common phenomena analyzed by forensic scientists; the accuracy limits of the tests and analytic techniques used (such limits have been determined for virtually all biomedical laboratory tests) and the accuracy of the tests *in practice*, as determined by quality control studies run on the laboratories.... Judges and lawyers will then have to become sufficiently conversant with the statistical principles involved in order to employ these background data in assessing the testimony presented.

M. SAKS & R. VAN DUIZEND, *supra* note 25, at 76.

But see Symposium on Science and the Rules of Evidence, supra note 167, at 228 (comments of Miron Straf):

> But to say that the expert passed the test by being correct 95% of the time really does not help the fact-finder too much, because the issue in question concerns a particular case to which that probability does not necessarily apply. It would be additional evidence to be considered, but the distinction between the prior tests and the specific question at trial must be made clear.

A concern for accurate verdicts should also compel forensic laboratories to keep complete records of all their activities. "Standards for record keeping are even more crucial if knowing that proficiency testing results will be presented to the jury, will be concerned about the quality of the laboratories they use. Disclosure of testing results to jurors will provide a powerful incentive for prosecutors and defense attorneys to choose the best laboratories and to exert pressures for improved performances.

The proposal that crime laboratories be required to take part in frequent proficiency testing with results made public is quite modest. It does not require the drafting and enforcement of detailed regulations, or the creation of a large new bureaucracy. Nor would it prevent innovation.[288] Money, of course, is always a concern. But, since the required funds are not inordinate, at least this much regulation should be imposed on forensic laboratories.[289]

---

the results are admitted into court, since the records are one of the few ways for the opposing party to review the accuracy of the test." Note, *supra* note 50, at 493. *See also* Baird, *supra* note 203, at 175; Winek, *supra* note 56, at 181.

288. *Cf.* Gilbert, *supra* note 281, at 973. Gilbert noted that clinical chemistry calcium analyses lack precision and concluded:

> The data . . . suggest that the goal is beyond the present level of technology. . . . It is difficult to believe . . . that this could be corrected by laboratory regulation. On the contrary, it would appear far more probable that regulation would inhibit the free development of alternative methods by encouraging or even enforcing a single preconceived solution to a complex problem.

Id. Shapiro, *Biotechnology and the Design of Regulation*, 17 ECOLOGY L.Q. 1, 1 (1990) ("Regulation [of modern technology] can prevent, or at least mitigate, some of [the] dangers [to people and the environment], but it can also inhibit innovation and product development.").

289. This more limited proposal should have support within the forensic science community. *See, e.g.*, LEAA, *supra* note 1, at 257. ("There is a need for continuous proficiency testing programs at either the national, state or local levels to provide a means to monitor the progress of efforts to upgrade and maintain high quality criminalistics services."); Kelly & Sunshine, *supra* note 49, at 1414. ("No one can deny the value of an ongoing proficiency testing program . . . . Forensic toxicology will never achieve true professional stature until all concerned realize this goal."). Brunelle & Cantu, *Training Requirements and Ethical Responsibilities Performing Ink Dating Examinations*, 32 J. FORENSIC SCI. 1502, 1504 (1987) ("Forensic scientists should not attempt to examine actual . . . cases until they have been tested. Proficiency testing demonstrates the competency of the examiner's work."); Thomson, *supra* note 175, at 513 (advocating a particular method of testing). A survey of forensic science students found that they "are in strong agreement that even qualified forensic scientists will make occasional mistakes and should be required to demonstrate their competencies on proficiency tests. They believe scientist[s] who recognize they are incompetent and fail to take steps to improve their quality are unethical. . . ." Peterson, *Teaching Ethics in a Forensic Science Curriculum*, 33 J. FORENSIC SCI. 1081, 1083 (1988).

While mandatory, blind, public proficiency testing may garner some support from forensic scientists, it may also meet with strong resistance.

> The FBI . . . opposes independent proficiency testing, arguing that no outsider is qualified to evaluate the bureau's performance. In addition, at a recent FBI-sponsored symposium on DNA typing that attracted 300 forensic scientists from

CONCLUSION

All available information indicates that forensic science laboratories perform poorly. Logic, justice, and concern for the wise expenditure of money require improvement in forensic science performance. Current regulation of clinical labs indicates that a regulatory system can improve crime laboratories. Lack of manpower, money, experience, and an appropriate institutional superstructure make comprehensive regulation of crime labs infeasible. However, forensic facilities should at least be required to undergo mandatory, blind proficiency testing, and the results of this testing should be made public. The testing would be an important first step in correcting inherent problems in the forensic science system.

---

around the country, FBI personnel were alone in opposing proposals requiring laboratories to explain in writing the basis for their conclusions and to have their reports signed by the scientists and technicians who conducted the test. Neufeld & Colman, *supra* note 157, at 53.

At a symposium on scientific evidence, Professor Joseph Peterson, one of those in charge of the LEAA proficiency study, noted:

> The laboratories in our study insisted upon confidentiality in reporting their results for fear that their credibility might be impaired if they submitted erroneous results. Some laboratories refuse to participate in quality control programs because of fear that even minor deficiencies might erode their credibility in the courtroom.

Symposium, supra note 5, at 648. Professor Andre Moenssens responded, "They are afraid of losing the unearned credibility they attained simply by labeling themselves a forensic science laboratory." *Id.*