Joseph L. Peterson,[1] D.Crim. and Penelope N. Markham,[2] Ph.D.

Crime Laboratory Proficiency Testing Results, 1978 –1991, I: Identification and Classification of Physical Evidence

REFERENCE: Peterson, J. L. and Markham, P. N., "**Crime Laboratory Proficiency Testing Results, 1978–1991, I: Identification and Classification of Physical Evidence,**" *Journal of Forensic Sciences,* JFSCA, Vol. 40, No. 6, November 1995, pp. 994–1008.

ABSTRACT: The proficiency testing of crime laboratories began in the mid-1970s and presently assumes an important role in quality assurance programs within most forensic laboratories. This article reviews the origins and early results of this testing program and also examines the progress of proficiency testing in allied scientific fields. Beginning in 1978, a fee-based crime laboratory proficiency testing program was launched and has grown to its present level involving almost 400 laboratories worldwide. This is the first of two articles that review the objectives, limitations and results of this testing from 1978 through 1991. Part I reviews the success of laboratories in the identification and classification of common evidence types: controlled substances, flammables, explosives, fibers, bloodstains, and hairs. Laboratories enjoy a high degree of success in identifying drugs and classifying (typing) bloodstains. They are moderately successful in identifying flammables, explosives, and fibers. Animal hair identification and human hair body location results are troublesome. The second paper will review the proficiency of crime laboratories in determining if two or more evidentiary samples shared a common origin.

KEYWORDS: forensic science, criminalistics, proficiency testing, crime laboratories

One of the cornerstones of the forensic sciences is the presumed validity and reliability of scientific test results and interpretations. Up until the mid-1970s, however, there were virtually no procedures in place to test this assumption empirically; rather, primary responsibility for ensuring the reliability of forensic results rested with each individual scientist and his or her laboratory organization. Laboratories were expected to hire and train competent personnel, preserve the integrity of the evidence, use proper scientific methods, and write reports and deliver appropriate testimony in courts of law. The judiciary placed its faith in such legal procedures as voire dire, cross examination and the appearance of opposing experts to evaluate the credentials of examiners and challenge their findings and opinions. This situation changed when a program of crime laboratory proficiency testing was developed with the assistance of funding from the federal government.

Received for publication 11 Oct. 1994; accepted for publication 17 April 1995.

This research was supported, in part, by a grant from the Office of Social Science Research, University of Illinois at Chicago.

[1]Professor, Department of Criminal Justice, University of Illinois, Chicago, IL.

[2]Research Associate, Department of Medicinal Chemistry and Pharmacognosy, University of Illinois, Chicago, IL.

Review of Laboratory Proficiency Testing

Crime Laboratory Testing

Proficiency testing for criminalistics laboratories began on a broad scale in 1974 with a grant from the National Institute of Law Enforcement and Criminal Justice (LEAA) to the Forensic Sciences Foundation (FSF). That grant enabled the FSF to manufacture and issue a series of twenty-one tests, covering a broad range of "evidence" types, to voluntarily participating forensic laboratories.

While the primary purpose of this research (hereafter referred to as the "LEAA study") was to *develop* mechanisms for administering such a testing program (that is, *how* to prepare and distribute samples, and *how* to analyze and report results), the study, nonetheless, produced results that called into question the long-held assumption that crime laboratory results were of uniform high quality. The project found that some crime laboratories were experiencing serious problems in the examination and interpretation of several types of test specimens [*1*]. In particular, laboratories exhibited high rates of unacceptable proficiency in identifying animal hair specimens and in determining if samples of paint, soil, glass, blood, and handwriting shared a common origin. Laboratories also revealed they were lagging in the adoption of serum protein and isoenzyme analysis techniques to assist in the discrimination of bloodstains. The project staff and advisory committee attributed these findings to several factors: misinterpretation of test results by examiners who were careless or lacked necessary training and/or experience, mislabeled or contaminated standards, inadequate data bases, and faulty testing procedures.

Following the initial release of these test results in 1977, and publication of the final report in 1978, the print and broadcast media ran a number of stories critical of crime laboratory performance [*2*]. Legal scholars took cognizance of these results and began to incorporate these findings in their treatments of scientific evidence, and the reliability of crime laboratory examinations emerged as an issue receiving serious legal scrutiny [*3,4*].

Equally important, the LEAA study resulted in a number of recommendations including the need for the commitment of greater resources to these laboratories, improved education and training opportunities, implementation of accreditation and certification programs, as well as the need for ongoing proficiency testing and quality assurance programs. Progress has been notable on many of these fronts over the past fifteen years [*5*], including the continuation and expansion of proficiency testing at the local, state, federal, and international levels.

Despite ongoing proficiency testing programs, there have been relatively few forensic (crime) laboratory proficiency testing results reported in the scientific literature. Publications have basi-

Copyright © 1995 by ASTM International

cally described programs introduced in particular laboratory systems [6,7], with Brunelle observing that laboratories of the Bureau of Alcohol, Tobacco and Firearms had achieved 100% accuracy in their blind proficiency trials. Several presentations were made by representatives of local, state and federal laboratories at the Annual Meeting of the American Academy of Forensic Sciences (AAFS) in February 1989 in a workshop on proficiency testing and quality assurance [8]. Presenters indicated proficiency tests were conducted regularly in their laboratory systems, but provided no performance data. In England, the Central Research and Support Establishment of the Home Office submits a variety of proficiency tests to its examiners [7], including biologists trained to conduct DNA tests [9]. Results of such testing were not published. At the 1993 AAFS meeting, an official with the FBI described their DNA proficiency testing program and, though the data remain unpublished, he noted there had been no reports of a false association.

A 1985 article by Lucas, Leete and Field [10] reviewed the participation and response rates in the post LEAA, Collaborative Testing Service (CTS) program, and offered general comments on areas where laboratories were performing well plus those needing improvement. Additional articles that relied on CTS data, usually focusing on a single area of evidence testing, have appeared. Sensabaugh [11] examined the CTS results of electrophoretic tests of blood and physiological fluids between the years 1975 and 1986 and found that of almost 8000 tests performed, fewer than 3% were in error. Using the same data, Grunbaum's review of physiological fluid proficiency test results from 1979 through 1983 found ABO error rates to be about 6%, and electrophoretic typing of bloodstains in various systems to range between .3% and 5.2% [12]. Grunbaum was also cited in the Office of Technology Assessment's *Genetic Witness: Forensic Uses of DNA Tests Report* as stating that "25 percent of laboratories returning results made errors" [13].

Risinger et al. [14] published a critical review of the results of handwriting comparisons administered through CTS for the period 1984 to 1987 and reported unsatisfactory results exceeding 40%. Risinger et al.'s stinging conclusion that there was no evidence to support the "existence of handwriting identification expertise" [14, pp. 750–751] was addressed in a study published in 1994 which set out to test the hypothesis that professional handwriting examiners were more proficient in performing writer identification than nonprofessionals [15]. This study, contrasting the performance of seven professional FBI examiners and ten graduate students with no training in handwriting comparison, found the professional examiners were significantly better in comparing and classifying handwriting samples than the college-educated nonexperts.

In another article using his own data, Miller found student accuracy in performing human hair comparisons varied depending upon the number of choices (knowns, unknowns) presented to students [16]. While such published data are generally indicative of examiner performance, it becomes clear that the manner in which "error" is defined and computed significantly affects the reported "proficiency" results.

Clinical, Toxicological, and Drug Screening Proficiency Results

An extensive history and literature exist for proficiency testing in a wide range of clinical analysis. The method and content of testing in these areas are in many respects based on procedures that have been in place in the clinical laboratory area beginning in the mid-1940s [17]. Early proficiency testing surveys in the clinical laboratory area documented widespread problems which, coupled with media coverage of unsatisfactory results and U.S. Congressional scrutiny, led to passage of the Clinical Laboratory Improvement Act (CLIA) of 1967 (Public Law 90-174).

CLIA '67 enacted many standards and regulations for clinical laboratories engaged in interstate specimen testing and laboratories receiving medicare funding, but excluded physician's office and forensic toxicology laboratories. The Centers for Disease Control, College of American Pathologists (CAP), and several state agencies furnished affected laboratories with proficiency test samples for the entire spectrum of clinical laboratory testing analysis and specimen types [18]. The results of these surveys have been communicated in the cited literature as annual summaries of surveys (American Journal of Clinical Pathology, 1975–1983). Since 1983, CAP surveys were published in the Archives of Pathology and Laboratory Medicine.

There were many proposals to toughen and expand these regulations over the years but it was not until 1988, based in large measure on data showing the superior performance of regulated laboratories versus those that were unregulated [19], that the law was strengthened and extended via Public Law 100-578 (CLIA '88) to virtually all clinical (but non-forensic) laboratories in the United States [20]. This legislation consolidated various regulatory requirements and introduced mandatory standards for technical and supervisory staff, licensing requirements, and uniform quality assurance procedures [21]. Proficiency testing was strengthened and became the centerpiece of efforts to determine a laboratory's competence; proficiency test samples were to be issued regularly, treated as routine specimens, and results of their analyses published. Uniform grading procedures were established and laboratories failing to meet minimum standards would risk losing their certification. Criteria were also established for assessing agencies engaged in the business of administering proficiency testing programs [22]. Proficiency surveys are also recognized as an important tool for the evaluation of new and emerging disciplines. Currently, this is evident in the field of molecular pathology in which surveys under the auspices of CAP are being initiated to assess areas including forensic DNA and paternity testing [23].

The more specialized field of drug screening first began on a large scale in the late 1960s as a result of the growth of methadone treatment programs and the development of immunoassay techniques to analyze specimens. With the increase in the use of marijuana and other drugs of abuse in the armed services, the workplace, and other sectors of society, drug screening became increasingly used both to identify and to deter potential abusers. Willette reviewed the programs of many federal agencies and private companies engaged in drug testing of job applicants and employees that have the purpose of creating a drug-free work environment [24]. Given the substantial percentage of subjects testing positive for drugs of abuse and the growing use of sanctions against them, many agencies of government (the military in particular) recognized they had to address the reliability of drug screening tests [25]. The threat of sanctions against subjects testing positive raised obvious forensic implications, and mandated that drug screening laboratories not only perform high quality analytical work on large numbers of specimens, but also uphold traditional evidentiary (for example, chain of custody) requirements [26].

The Armed Forces Institute of Pathology (AFIP) took responsibility for conducting proficiency testing of laboratories serving the military, which collectively handled more than 2 million specimens per year [27]. In the early 1970s the Centers for Disease Control (CDC), the American Association for Clinical Chemistry (AACC), and the National Institute on Drug abuse (NIDA) initiated profi-

ciency testing of laboratories engaged in federally funded urinalysis programs and published several studies documenting their proficiencies [26–32]. These studies demonstrated several things: (1) the value of simulated clinical specimens; (2) the relatively low incidence of false positive results and comparatively higher rates of false negatives; and (3) that results varied greatly as a function of testing mechanics such as drug concentration thresholds, and if the tests were "open" (the examiner knew it was a test) or "blind" (the sample was introduced as a "routine" case). Hansen et al. [32] reviewed CDC testing for the period 1972–1981 and found high rates of false negatives for certain drugs, and comparatively poor performance of labs when presented with blind specimens. These studies also pointed out the wide variations in drug quantitation results, the need for confirmatory testing of positive results, and the desirability of accreditation standards, including routine proficiency testing. The recent literature documents well the improvements made in drug screening efforts resulting, in large measure, from regulation and proficiency testing [33,34].

During this same period (1970s and 80s) there were also attempts to introduce proficiency testing in the general forensic toxicology area [35–40]. The initial study results of Dinovo and Gottschalk indicating "startling interlaboratory differences" [35, p. 843] were criticized for deficiencies in sample selection and statistical interpretation of results. Peat et al.'s study [39] revealed some problems in detecting opiates in blood and low concentrations of barbiturates in blood and urine, with results showing wide interlaboratory variations in quantitation. On the other hand, researchers concluded proficiency testing was feasible and desirable, and that greater attention needed to be paid to the types of samples issued and methods used to interpret the data generated [39]. Forensic toxicologists acknowledge the importance of urine drug testing in today's society and the need for guidelines to regulate personnel, facilities and test results [40].

With studies showing that controls and regulations enhanced the quality of testing, the federal government strengthened its urine-screening provisions and now mandates a quality assurance program covering "chain of custody, security and reporting of results, initial and confirmatory testing and validation of analytical procedures" [41]. These regulations include a requirement to subscribe to a blind proficiency testing program offered by organizations certified by the Department of Health and Human Services.

With increased drug screening of arrestees, greater attention is also being paid to the sensitivity and reliability of analytical techniques used in criminal justice monitoring programs. A publication of the National Institute of Justice [42] found thin layer chromatography to be an inferior screening technique (compared with three immunoassay techniques) and recommended confirmation of all positive tests by GC/MS.

Proficiency Testing and the Regulation of Crime Laboratories

Although certain crime laboratories and laboratory systems have required proficiency testing of its examiners for years, from a national perspective, the participation of examiners in such testing has been mostly an elective form of quality control. This has changed with the establishment of laboratory accreditation, introduced in 1981, and examiner certification, begun in the field of criminalistics in 1993. The American Society of Crime Laboratory Directors (ASCLD), and its accrediting body ASCLD/LAB, expect laboratories that seek accreditation to satisfy various criteria, including maintenance of a "quality system." [43, pp. 17–21]. This requires the laboratory to follow various practices, including: validation and written documentation of all technical procedures, verification of the quality of standard samples and reagents, periodic calibration of instruments, administrative review of all technical reports, and proficiency testing. A proficiency testing program must be in place in which the laboratory receives samples from an approved external provider and each examiner in the lab must successfully complete a proficiency test annually in each functional area in which they perform casework. For DNA examiners, ASCLD/LAB has endorsed the Technical Working Group on DNA Analysis Methods (TWGDAM) quality assurance standards [44], and requires every examiner to complete successfully a minimum of two proficiency tests annually (at least one of them external). Although a voluntary program, ASCLD/LAB has thus far accredited about 130 crime laboratories [45].

Proficiency testing is also an essential component of the criminalistics certification process implemented in 1993. After an applicant passes a general written examination and becomes a diplomate of the American Board of Criminalistics, he/she may elect to advance to Fellow status within one or more specialty areas (drug identification, forensic biology, paints and polymers, hairs and fibers, and fire debris analysis) [46]. In addition to passing a written exam in their specialty field, applicants must submit certifiable proficiency test results in that area. To maintain their specialty certification, they are required to submit acceptable proficiency tests annually. So, although proficiency testing is still a voluntary process for most examiners and results are rarely disseminated beyond the individual's laboratory or certifying body, there is little question that proficiency testing has become an important step in the attainment of professional standards.

Blind trials of forensic DNA testing in its early days (1987) were organized by the California Association of Crime Laboratory Directors and involved three commercial laboratories. Two firms each declared one false match involving fifty samples and a third laboratory had no false matches. In a second blind trial test, one laboratory reported a false match [47]. See, also, Thompson and Ford's interpretation of these same results [48].

Jonakait [49] published a highly critical review of crime laboratory performance in 1991, using as his basis the LEAA, toxicology, and DNA proficiency test studies noted earlier. He called for tighter regulation of laboratories using the Clinical Laboratory Improvement Act as a model, beginning with mandatory crime laboratory proficiency tests and publication of results. This article prompted an immediate and vehement response from the professional community [50], members of which complained he was relying on outdated information, failed to appreciate the limitations of proficiency test results, and had neglected to recognize that many laboratories had adopted proficiency testing and many of his other recommendations to insure consistently high performance. While it is unfortunate Jonakait chose to emphasize the proficiency data and other shortcomings of the field at the expense of the many positive advancements made in forensic science in recent years, he nonetheless did identify a number of important issues that merit serious consideration—the need to bolster education, training and research programs, consideration of alternative organizational environments for forensic science laboratories, and assessment of mechanisms for insuring that all forensic science laboratories adopt quality assurance systems.

Crime Laboratory Proficiency Testing, 1978–1991

Beginning in 1978, the year in which the LEAA sponsored proficiency testing report was issued, the FSF and Collaborative

Testing Services (with the support of the American Society of Crime Laboratory Directors (ASCLD)) launched a fee-based proficiency testing program. This program has grown steadily over the years with now more than 390 laboratories participating, including 65 laboratories in 21 foreign countries [51].

This article summarizes the results obtained by laboratories in these tests between the years 1978 and 1991. The performance results of given examinations were conveyed in reports to participants following each set of tests; this paper is the first to summarize the laboratory results on all tests of various evidence types during this period. We also examine whether laboratories have improved their performance in comparison with the results published under the original LEAA study in 1978.

As stated by the proficiency testing managers, the specific objectives of the CTS-FSF program have been:

- to provide laboratories the opportunity to examine physical evidence samples and to compare their performance with others receiving duplicate samples
- to provide laboratory directors with a mechanism for self-evaluation and improvement
- to encourage laboratories to handle samples according to "normal operating procedures"
- to provide manufacturers' sample specifications and present tabulations and analyses of laboratory responses
- to summarize methods used and, in selected tests, time expended by, and experience/qualifications of, examiners
- to identify problems and make recommendations as to how laboratories might enhance their examinations [52].

Limitations

The data generated through this program have limitations in terms of accurately reflecting the quality of casework in today's crime laboratories. First, given the nature of evidence typically submitted to forensic laboratories, construction of proficiency samples is a much more complicated task than that faced by manufacturers of specimens in other laboratory testing fields. Whether it was the preparation of a simulated controlled substance seizure involving multiple drug types, the creation of fire debris samples containing different petroleum products, scenarios involving various animal or human hair samples, or swatches of cloth stained with a variety of body fluids, each presented its own set of challenges to the manufacturer. In addition to the pitfalls encountered in production of samples, there was always the possibility of human error or inadvertent contamination during the packaging and mailing process.

Second, the goal of creating tests representative of actual casework was sometimes compromised by the necessity of producing a large number of identical samples. For example, imprint and impression tests usually did not include actual suspect tools or weapons with which laboratories could make their own knowns. Similarly, with document and fingerprint tests, laboratories were issued photographs of so-called standards rather than originals, a practice at variance with routine requirements in most laboratories.

Third, these were declared proficiency tests, and examiners knew they were being tested. Also, because the testing was voluntary, with about two thirds of U.S. laboratories subscribing to the program and one third responding with data, the results do not necessarily represent all laboratories engaged in this type of casework. There are various possible explanations for the high rate of nonresponses, ranging from some laboratories' reluctance to have even their anonymous replies recorded and disseminated, to laboratory systems that procured multiple samples for training and reference purposes, but only returned single responses.

Fourth, the level of difficulty of tests varied among and within evidence types and, therefore, did not always reflect typical cases routinely examined in forensic laboratories. Unlike these routine cases, the proficiency testing sponsors would continually try to devise new scenarios and specimens that would interest and challenge the participating laboratories.

A final limitation is that the testing service had no control over how the sample was treated in the laboratory upon receipt. While laboratories were generally asked to examine the samples as they would handle routine case material of that variety, the program could not dictate which tests or procedures were to be followed nor the qualifications of the particular examiner doing the work. Anecdotally, we know many labs treated the sample as a collaborative effort, some as a training exercise for new personnel, and still others as an experimental effort to evaluate new tests or instruments. We know, too, based on the number of tests and hours of effort reported by laboratories on several tests, that many laboratories invested more time examining samples than would be expected or required on actual casework. Similarly, we cannot be sure if laboratories were inclined to be overly conservative in reporting their responses or to take more risks since these were not actual cases and their responses were anonymous.

In sum, caution must be exercised in interpreting the meaning of these data and extrapolating these results to the crime laboratory profession at large. The data provide a view of how a subset of laboratories responded to particular case specimens over the time period encompassed by this testing. Still, even with these limitations, these results have value since they are one of the only sources of data on laboratory performance in a broad spectrum of evidence categories since the original LEAA study data were published in 1978.

The Data

The data examined in this article were gathered from individual test and supplementary reports issued under the CTS/FSF proficiency testing program. Both tabular data and summaries of results appearing in these reports were used. Comments of the Proficiency Advisory Committee (PAC) published in these reports were also helpful in interpreting the results and identifying important problems and issues.

Table 1 enumerates the various kinds of samples issued during the testing, as well as information on the number of tests issued, the net increase/decrease in the number of laboratories participating in the tests, and the participation rates. In all, a total of 175 separate tests were conducted and results tabulated. Drugs and body fluids had the most tests issued (29 and 28, respectively). For all types of samples, save for explosives, there were substantial increases in the number of subscribing laboratories; in eleven categories, there was more than a doubling of subscribers over the 14 year testing period. The percent of laboratories receiving samples that responded with data ranged generally between 45% and 65%, with rates of responses generally increasing over time.

The test results are organized into two major sections: those in this article concern the identification of substances; in a followup article results of comparisons of "known" and "unknown" samples to determine possible common origin are detailed. We initially

TABLE 1—*Proficiency test reports (1978–1991).*

Evidence Type	Number of Tests	Subscribing Laboratories[a]	Overall Participation Rate[b]	Purpose
Drugs	29	79–256 (+224%)	60%	Identification Quantitation
Flammables	11	80–196 (+145%)	56%	Identification
Explosives	4	80–71 (−11%)	25%	Identification
Fibers	13	67–199 (+197%)	43%	Identification Comparison
Blood/body fluid	28	69–220 (+219%)	44%	Identification Comparison
Hair	8	82–147 (+79%)	41%	Identification Comparison
Firearms	14	42–173 (+312%)	49%	Comparison
Toolmarks	12	72–163 (+126%)	52%	Comparison
Glass	11	63–148 (+134%)	51%	Comparison
Footwear	7	84–160 (+90%)	62%	Comparison
Paint	18	67–164 (+144%)	49%	Comparison
Questioned documents	8	41–83 (+102%)	68%	Comparison
Latent prints	9	38–141 (+271%)	67%	Comparison
Metals	3	40–103 (+153%)	44%	Comparison
Total	175			

[a]The number of laboratories subscribing to the first and final test in the series, with the percent change in parentheses.
[b]Percent of laboratories receiving samples that responded with results.

review the performance of laboratories in identifying and classifying the following types of evidentiary materials:

- Controlled substances—qualitative and quantitative determinations of drugs of abuse and diluents
- Hair—determination of species of origin for animal hair and body area of origin of human hair
- Flammables—presence or absence of accelerants and their identification
- Explosives—identification and classification of materials
- Fibers—identification
- Blood and body fluids—determination of species of origin of bloodstains, identification of blood and nonblood body fluids (saliva or semen), and blood grouping information.

It should be noted that for some categories—fibers, blood and body fluids, and hair—laboratories were asked to identify the samples as well as answer questions of common origin. The identity/classification results are included in this article, while those addressing common origin appear in the next paper.

Table 2 summarizes the fraction/percent of laboratory responses identifying substances that agreed with the manufacturers' specifications. The materials are listed in descending order of percent correct identifications for six evidence categories, beginning with bloodstains and drugs where, on average, responses were in agreement with manufacturers' specifications the highest percent of time (94%).

Noteworthy findings for each of the six evidence categories follow.

TABLE 2—*Identification/classification of common evidence types.*

Evidence Type	Number of Identification Tests	Fraction/Percent Correct Identifications
Blood/body fluids typing	13	17,050/18,150 = 94%
Drugs	29	2228/2364 = 94%
Flammables	10	1670/1826 = 91%
Fibers	9	1170/1318 = 89%
Explosives	4	185/228 = 81%
Hair	4	105/193 = 54% (animal species) 223/339 = 56% (human body area)

Blood/Body Fluids

There were 28 separate physiological fluid tests issued during the testing, the reports for two of which (1 and 9), however, did not include test results. There were ten exercises (12, 18, 85-2, 86-2, 87-2, 88-2, 89-2, 90-2, 90-13, 91-14) in which the principal objective was for laboratories to supply typing results on various bloodstain samples. There were four more where laboratories were presented with scenarios and blood grouping results were requested in addition to information on their possible sources (80-2, 82-9, 86-11, and 88-14). (88-14 involved a fetal bloodstain pooled from different sources, which caused confusion and is not included in this or subsequent discussions.) For the purposes of this tabulation we did not include typing results in exercises that called for the identification of nonblood body fluids (80-9, 87-13, 91-2) and others where scenarios involved blood and/or other body fluid mixtures and where the primary purpose was to resolve issues of common origin (nine additional tests). The reader may wish to consult Sensabaugh's tabulation and review of electrophoretic typing error rates which also included many of these latter tests [*11*]. His overall error rate (2.4%) is similar to ours. The results of the exercises involving questions of common origin will be discussed in the following article.

In this first section, we will discuss those samples where laboratories were asked to determine the blood grouping of various

TABLE 3—*Blood group determination.*
Totals (12,18,80-2,82-9,85-2,86-2,86-11,87-2,88-2,89-2,90-2, 90-13,91-14).

System	Results: Agree	Disagree[a]	Inconclusive	Total
ABO	2395 (96%)	43 (2%)	68 (3%)	2506
PGM	2138 (98%)	27 (1%)	21 (1%)	2186
PGMsub	1207 (98%)	12 (1%)	12 (1%)	1231
EAP	1966 (91%)	99 (5%)	92 (4%)	2157
EsD	1914 (94%)	32 (2%)	90 (4%)	2036
AK	1610 (97%)	3 (.2%)	48 (3%)	1661
ADA	1491 (95%)	14 (1%)	61 (4%)	1566
Hp	1027 (92%)	27 (2%)	67 (6%)	1121
GLO	1100 (85%)	29 (2%)	159 (12%)	1288
Gc	675 (86%)	8 (1%)	98 (13%)	781
PEP A	394 (90%)	3 (1%)	42 (10%)	439
Tf	343 (97%)	1 (.3%)	9 (3%)	353
CA II	431 (97%)	3 (1%)	11 (2%)	445
Hb	249 (95%)	0 (0%)	13 (5%)	262
6PGD	110 (93%)	0 (0%)	8 (7%)	118
Total	17,050 (94%)	301 (2%)	799 (4%)	18,150

[a]Percent of total responses that agreed and disagreed with the manufacturers' specifications, or were inconclusive for each of the respective systems.

bloodstains and, in two exercises, where laboratories were asked to determine the stains' species of origin. With respect to species testing, only three (0.4%) of the 670 responses in these two tests were at variance with the manufacturer's target values (in test #12, one lab reported goat blood as human, and one lab reported one of the human stains as nonhuman; in test #18, one lab reported the chicken bloodstain not to be blood).

In terms of identification and classification (typing) of bloodstains, laboratories were more successful in this series than in any other identification exercises. Of the more than 18,000 typing results reviewed (data were tabulated for the 15 most frequently reported systems) respondents were in agreement with the manufacturers' specifications an average of 94% of the time (Table 3).

Laboratory results disagreed with manufacturers' specifications in about 2% of responses and yielded inconclusive results in another 4%. Laboratories were most successful in their typing of PGM, PGMsub, AK, Tf, and CA II systems, where results were in agreement 97% of the time or higher. Laboratories mistyped stains in only 1% or less of these trials, with the balance of results inconclusive. Laboratories had the greatest difficulties with the GLO, Gc, Pep A, EAP and Hp systems. Gc, GLO, and Pep A had the highest rates of inconclusive replies, 13%, 12%, and 10%, respectively.

The EAP system had a significantly higher rate of error than any other single system and contributed practically a third of all incorrect responses in the approximately 18,000 results reviewed. The PAC attributed the higher number of inconclusive and incorrect responses with this system to "sample transposition" and the lack of "knowledge, training or experience" on the part of the examiner in distinguishing the relative intensities of the bands. Use of known controls of pertinent phenotypes was highly recommended. While there did not appear to be a correlation between years of experience and success in grouping the stains, it was apparent that a few laboratories were responsible for a disproportionately large fraction of mistakes, suggesting these were due to a few examiners who lacked the necessary knowledge and skills across several systems. ABO was the system typed most often and while results were in agreement about 96% of the time, they disagreed in 2% and were inconclusive 2% of the time.

Compared with the results published in the 1978 report, laboratory performance in ABO typing has remained about the same (about 96% correct identifications) while performance in other systems has improved substantially. More systems are being typed with greater accuracy and systems with which examiners experienced problems (for example, MN, Rh) have been discontinued.

It also appears there was general improvement in typing over the years, 1978 to 1991. From 1978 to 1982, error rates averaged about 2.9% over four typing tests administered. There were no typing results tabulated for the years 1983–1984. In the period 1985 to 1988, five bloodstain typing exercises resulted in incorrect replies in 1.6% of responses. From 1989 to 1991 the rate of error lowered slightly to 1.5%, and would have been less than 1.0% were it not for test 90-13 where the error rate jumped to 2.1% due to problems experienced by five laboratories that supplied almost half of the incorrect results.

Drugs

A total of 29 drug tests were administered to crime laboratories between the years 1978 and 1991 (see Table 12). Laboratories were typically issued a substance and requested to identify it using their normal laboratory procedures. They were also requested to describe the methods they used and to report all quantitative and qualitative data developed. Usually the test involved the identification of one or more drugs in the presence of other substances (incipients, diluents) with which the drug(s) had been mixed. In 12 of the final 16 tests, laboratories were instructed to identify *and* quantitate the sample if a controlled substance was found. In exercise 83-2, quantitation was necessary in order to answer the request to determine if two different samples might have shared a common origin. Some scenarios also indicated the substance had been seized from a clandestine laboratory and participants were asked if they could predict the route of synthesis/method of manufacture used in production of the drug (84-1, 86-9).

The number of laboratories participating in the testing more than tripled, increasing from 79 in 1978 to 256 in 1991. With response rates hovering around 50% to 60% during the period, the number of laboratories which returned results rose from 56 to 192. Rates of participation did not increase or decrease appreciably over time, although they did vary by type of controlled substance tested. Samples issued to laboratories fell into the four major categories of controlled substances: narcotics (seven tests), stimulants (eleven tests), depressants (two tests) and hallucinogens (five tests) and four tests containing exclusively noncontrolled drugs. Participation rates for the analysis of narcotics, stimulants and hallucinogens were higher on average (60% to 65%) than for depressants (51%).

The nature of samples issued to laboratories were of four types: 1) those containing a single controlled drug; 2) samples containing more than one type of controlled substance; 3) those having one or more controlled substances *mixed* with noncontrolled drugs; and 4) those containing only noncontrolled substances. Table 4 shows the overall success of laboratories in identifying the substances—both singly, and when mixed with other controlled and/or noncontrolled substances. Laboratories identified the *controlled drug* present in an average of 94% of responses.

In nine exercises, 100% of reporting laboratories correctly identified the controlled substance in question. An exercise, though, which gave laboratories some difficulty was test 89-12 where analysts were issued a recently scheduled controlled substance, N,N dimethylamphetamine HCl. A number of laboratories were unprepared (several lacked the necessary standard) to identify this substance and, not surprisingly, the identification rate was among the lowest (86%) for samples containing a single controlled drug.

A more common problem, though, was where laboratories did not identify *all* controlled substances present in a sample as in 89-4, where laboratories were asked to examine a sample containing both cocaine and methamphetamine. Whereas 95% of labs identified the cocaine and 92% the methamphetamine, only 85% identified them both. A single laboratory identified neither.

Table 5 provides a retabulation in which the success in identifying only controlled substances in the samples is contrasted with success in identifying both controlled *and* uncontrolled substances

TABLE 4—*Rates of successful drug identifications.*

Controlled Substance To Be Identified	Samples	Identification Rate	Rates (Range)[a]
Opiates	7	537/564 = 95%	83%–100%
Stimulants	11	1163/1201 = 97%	86%–100%
Depressants	2	87/91 = 96%	90%–100%
Hallucinogens	5	419/486 = 86%	47%–100%
Total	25	2206/2342 = 94%	

[a]"Range" includes the low and high successful identification rates for the tests administered within given substance categories.

TABLE 5—*Successful drug identification rates.*

Category[a]	Samples	Samples Including Noncontrolled Drugs[b]	Samples Excluding Noncontrolled Drugs[c]
Opiates	7	$\frac{802}{854}$(94%)	$\frac{537}{564}$(95%)
Stimulants	11	$\frac{1363}{1533}$(89%)	$\frac{1163}{1201}$(97%)
Depressants	2	$\frac{87}{91}$(96%)	$\frac{87}{91}$(96%)
Hallucinogens	5	$\frac{531}{698}$(76%)	$\frac{419}{486}$(86%)
Other/ noncontrolled	4	$\frac{399}{581}$(69%)	...
Total	29	$\frac{3182}{3757}$(85%)	$\frac{2206}{2342}$(94%)

[a]Category of primary drug cited, but in mixtures occasionally included other drug types (for example, see 86-1).
[b]Rate of identification for *all* drugs (controlled and noncontrolled) present in the various samples.
[c]Rate of identification for only controlled substances, either singly or in mixtures.

TABLE 6—*Rate of identification of noncontrolled substances.*

Substances	Samples Rate	Identification	Range
[a]Exclusive noncontrolled samples	4	$\frac{399}{581}$(69%)	(22%–100%)
[b]Noncontrolled substances that were mixed with controlled drugs	7	$\frac{577}{834}$(69%)	(36%–94%)

[a]80-5 (diphenhydrine), 81-1 (ephedrine and theophylline), 81-7 (PCC), 90-4 (testosterone).
[b]83-7 (catnip), 85-9 (tetracaine), 86-1 (procaine), 87-1 (ephedrine), 90-12 (nicotinamide), 91-5 (lysergol), 91-13 (procaine).

(singly and in mixtures). Including uncontrolled substances lowers the identification rate to 85%. There are a number of examples.

In 86-1, laboratories were issued a mixture of heroin, cocaine and procaine (noncontrolled) and were asked to identify and quantitate any controlled substances present. While 95% identified heroin, 83% the cocaine, and 86% the procaine, only 75% identified *both* the heroin and cocaine, and 67% of laboratories identified all three. In test 87-1, laboratories were issued a sample containing cocaine and ephedrine and were asked to identify and quantitate any drugs found. Whereas 100% of the responding laboratories identified the cocaine, only 39% also identified the ephedrine.

In the great majority of cases where laboratories experienced problems, therefore, it was a *failure* to identify either a controlled or noncontrolled substance present (false negative) rather than a misidentification (false positive). Both situations may reflect laboratory policy not requiring laboratories to reach beyond the first controlled substance identified or to identify noncontrolled substances, diluents or adulterants. False positives, however, can be a very serious problem and might result in the improper charging and conviction of an otherwise innocent party. There were only a handful of such mistakes, as in 81-1 where a laboratory incorrectly identified phenobarbital, and 87-1 where a laboratory identified methaqualone when none was present. In 83-7, one laboratory incorrectly identified catnip as marijuana (or somehow confused the samples). In 87-1, a sample containing cocaine and ephedrine, one laboratory also reported pseudoephedrine and another, phenmetrazine. In 87-8, a sample containing N-Ethyl M.D.A. Hydrochloride, five laboratories incorrectly identified the active ingredient and failed to identify the MDA derivative. In terms of total responses, such false positives constituted less than .5% of all replies.

The last category of exercises were those containing only noncontrolled drugs (see Table 6). In these four tests, laboratories were successful in about 69% of their attempts. The lowest successful rate of identification was for test 87-1 in which laboratories were issued a sample containing PCC—an analog of PCP. Only 22% of responding laboratories correctly identified this noncontrolled substance. One explanation for this low rate is that PCC was not on the list of controlled substances when this test was issued. Nevertheless, the PAC was highly critical of the procedures employed by many of the laboratories in this exercise, noting their "misuse of instrumentation."

With respect to quantitation, the percent of laboratories reporting results and their accuracy varied widely as a function of the type and amount of controlled substance present. For narcotics, the percent of laboratories supplying quantitative information ranged from a low of 18% on test 82-1 (morphine) to a high of 100% on test 23 (heroin). In 82-1, while only 6 laboratories quantitated the morphine, those that did, obtained excellent results. In general, the mean quantitative values reported by laboratories correlated well with the manufacturer's values, but the ranges were generally great. For eleven of the tests, laboratories were specifically requested to quantitate the drugs present. The percent of laboratories performing quantitative analyses, and the mean, range, and standard deviation for these tests are presented in Table 7.

When these results as a whole are compared with results published in the 1978 report, the controlled substance identification rates have improved by about eight percentage points—from an average rate of 86% successful identifications to an average identification rate of 94%. It is also possible to compare performance where laboratories were issued similar unknowns. Comparing sample #6, a mixture of heroin, cocaine and procaine, with sample 86-1, a similar mixture, the rate of identification of heroin was about the same, while the identification of the other drugs was significantly better—13 and 4 percentage point improvements for cocaine and procaine, respectively. Sample #15 in the LEAA study contained methamphetamine and ephedrine while tests 86-9 and 89-4 in the CTS program contained methamphetamine and methamphetamine and cocaine, respectively. Comparing the success of laboratories in identifying just the methamphetamine we see, again, laboratories exhibited higher rates of identification by a margin of 10 to 15 percentage points.

Flammables

Between the years 1980 and 1991 a total of eleven flammable test samples were issued to crime laboratories, ten of which called for the identification of accelerants (Table 8) and one which posed a question of common origin. The number of laboratories subscribing to flammables more than doubled in this time period (from 80 to 196) and the participation rate averaged 56% (with no trends apparent). A total of 127 laboratories reported results in the final test (90-8), compared with 39 in the first (80-3).

TABLE 7—*Drug quantitation results.*

Report	Drug	Target Concentration	Number (%) Quantitation[b]	$\bar{X}$	Range	S.D.
83-2	Heroin (A)	3.5%	22(36%)	3.1%	1.5–6.6%	1.19
	Heroin (B)	5.0%	23(38%)	4.0%	2–7%	1.15
84-8	Cocaine	20%	60(77%)	18.8%	11–30%	4.11
85-9	Cocaine-HCl	15.4%	78(91%)	14.2%	6.3–30%	3.53
86-1	Heroin	10%	75(77%)	10.3%	2.1–24.9%	3.74
	Cocaine	5%	65(66%)	4.2%	2–10.2%	1.51
86-9	Methamphetamine HCl	13%	57(65%)	12.1%	3.2–26.7%	3.78
87-1	Cocaine base	37.9%	63(87%)	37.3%	26–54%	5.75
87-8	MDA	14.4%	34(34%)	13.2%	3.0–28.6%	4.47
88-1	Mix-cocaine (base)	10%	17(15%)	15.4%	.4–33.1%	10.24
	Cocaine (HCl)	20%	43(37%)	24.3%	14–38%	4.85
	Cocaine (all)[a]	30.0%	55(47%)	25.3%	14.7–50%	6.46
88-8	(TCP) Thienylcyclohexylpiperidine	9.3%	35(30%)	9.7%	6–25%	4.07
89-4	Cocaine	12%	86(70%)	12.0%	5.15–29%	3.70
	Methamphetamine HCl	5%	57(46%)	5.0%	1–22%	3.01
89-12	N-n dimethylamphetamine HCl	15.5%	18(13%)	14.7%	10–21%	2.86

[a]Whereas 17 labs reported quantitative results specifically for the "base" and 43 for HCl, 55 reported quantitative results for "cocaine" without distinguishing between base or HCl.
[b]This column gives the number and percent of responding laboratories that quantitated samples. On occasion, extreme values, possibly clerical errors, were excluded from calculations.

Samples issued included diesel, charcoal starter, gasoline, lacquer thinner, lantern fluid and various weights of petroleum distillates (light to heavy) on different substrates including carpet, soil, cotton and wood. The PAC designed different tests, adjusting both the quantity and condition of the accelerant (for example, evaporated), as well as the condition of the substrate (for example, charred/uncharred carpeting). All samples were placed in clean paint cans or sealable bags before distribution. In all cases the laboratories were asked to search for the presence of flammable substances/accelerants and, if detected, to identify the particular accelerant(s) present. In general, the vast majority of laboratories detected flammable substances when they were present. Twenty-one of the twenty-six individual test items issued contained a flammable substance; laboratories correctly reported the presence of a flammable in 92% of their 1437 responses. In the five samples without a flammable, laboratories correctly reported none was present in 89% of their 388 replies. Stated conversely, laboratories failed to detect a flammable when present in 8% of their replies, and incorrectly reported the presence of an accelerant in 11% of trials when none was present.

With regard to the detection of flammable substances the performance of laboratories differed according to the quantity and nature of the accelerant in the sample. Laboratories seemed to have the greatest success in detecting the presence of an accelerant where gasoline had been added to the sample (see results from tests 83-11, 86-10, 88-9, and 89-8 where correct responses ranged between 95-100%). Occasionally, a few laboratories failed to detect gasoline even though a considerable quantity (100 μL) of the flammable had been added as in 83-11; the PAC attributed this to poor technique on the part of the analyst. Highly volatile, light petroleum distillates (LPDs) gave laboratories the greatest trouble, as in test 84-10, where only 56% of laboratories reported a flammable present when only 5 μL of a light petroleum distillate was added to a sample of charred carpet; in test 88-9, only 54% reported the presence of a flammable even when 50 μL of a LPD was added to a sample of cotton. Increasing the quantity of volatile in the test sample did seem to make the task easier, however, as was demonstrated in 84-10 where 84% of the labs identified the accelerant when 100 μL of the LPD was added to the carpet.

In four tests, 83-11, 84-10, 87-9, and 90-8 laboratories were also given the opportunity to analyze (carpet) samples containing no accelerants (either as controls or evidence). In 83-11, 2/50 laboratories reported detecting an accelerant on unburned carpeting even though none was present. The PAC attributed these mistakes to the "qualifications and experience of analysts." In 84-10, 13/63 (21%) laboratories reported finding an accelerant in charred carpeting containing no accelerants. The PAC noted some laboratories confused the pyrolysis products of the charred carpeting with accelerant products and, once again, stressed the need for more training of these analysts in recognizing pyrolysis patterns in burned carpeting. In 87-9, 8/74, and 7/74 of reporting laboratories reported nonexistent accelerants in two burned carpeting samples. The PAC suggested laboratories making these errors review their chromatographic patterns to determine how they could have confused accelerants from combustion products. In 90-8, 7/127 (6%) of laboratories mistakenly reported the presence of an accelerant in an uncharred carpet sample.

In terms of laboratories' abilities to identify the particular accelerant in the samples, laboratories were successful, on average, about 65% of the time. In test 83-11, 88, and 72% of laboratories correctly identified gasoline in samples 1 and 2, respectively. Failures were attributed to lack of analyst expertise and not the techniques they were using. Accordingly, the PAC recommended laboratories evaluate the qualifications and experience of their analysts. In 86-10, laboratories, again, achieved high (around 90%) success in identifying gasoline, but were unable to identify kerosene contained in one of the samples (only 15% did so). The LPD in test 84-10 caused laboratories great difficulty with only about 30% correct identifications; only 17% of laboratories correctly identified the LPDs in both samples with it present. This sample also had a very high percentage (21%) of misidentification of accelerants.

Test 87-9 proved interesting in that around 10% of respondents reported an accelerant in two samples, even though none was present; most thought it fell in the class 3 (medium petroleum distillate) range. Test 88-9 was instructive in that only 8/73 laboratories reporting data correctly identified all four accelerant types, class 1, 2, and 3 accelerant types plus a fourth which was a mixture

TABLE 8—*Flammable identification.*

Report	Participation Rate	Nature of Sample	Accelerant Identification[a] (Yes/No)	(Correct Class)[a]	Comments
80-3	39/80 (49%)	1) Charcoal starter on carpet sample	38/39 (97%)	23/39 (59%)	Correctly id'ed charcoal starter
		2) Charcoal starter			
		3) Leaded gasoline			
		4) Lacquer thinner			
81-4	24/67 (35%)	1) Diesel fuel on balsa wood	19/24 (79%)	10/24 (42%)	Includes: Conforms plus correct, but incomplete
		2) Charcoal lighter on balsa wood	20/24 (83%)	17/24 (71%)	
		3) Gasoline on balsa wood	21/24 (88%)	17/24 (71%)	
82-4	29/56 (52%)	Mixture charcoal lighter and diesel/ fuel added to soil	29/29 (100%)	23/29 (79%)	
83-11	50/75 (66%)	1) 100 μL, 58% evap. gasoline[1]	48/50 (96%)	44/50 (88%)	
		2) 5 μL, 58% evap. gasoline[1]	39/50 (78%)	36/50 (72%)	
		3) unburned control[2]	48/50 (96%)	2/50 (4%)	Mis id'ed accel.
		[1]added to burned carpet			
		[2]unburned carpet			
84-10	63/107 (59%)	1) 5.0 μL LPD/charred carpet	35/63 (56%)	11/63 (17%)	Id'ed LPD/ClassII
		2) Charred carpet-control	50/63 (79%)	13/63 (21%)	Mis id'ed accel.
		3) 100.0 μL LPD/charred carpet	53/63 (84%)	31/63 (49%)	Id'ed LPD/ClassII
			also:	11/63	Corr. id'ed LPDs 1 and 3
				32/63	Mis id'ed 1 and/or 3
86-10	65/123 (53%)	1) 25 μL unweathered gasoline[1]	64/65 (98%)	60/65 (92%)	
		2) 10 μL weathered gasoline[1]	63/64 (97%)	58/65 (89%)	
		3) 20 μL 50:50 mix gas/fuel oil[1]	64/65 (98%)	61/65 (94%)	Id'ed gasoline
				10/65 (15%)	Id'ed fuel oil
		4) Control-uncharred carpet	(no results)		
		[1]added to uncharred carpet			
87-9	74/135 (55%)	1) Charred carpet-no accel.	62/74 (83%) none	8/74 (11%)	Mis id'ed accel.
		2) Charred carpet-no accel.	64/74 (86%) none	7/74 (9%)	Mis id'ed accel.
		3) Control carpet			
88-9	74/161 (45%)	1) 50 μL lacquer thinner (I)	39/72 (54%)	36/72 (50%)	
		2) 50 μL 95% evap. gasoline (II)	72/72 (100%)	50/72 (70%)	
		3) 50 μL 1:1mx gs/diesel (II,V)	72/72 (100%)	27/72 (38%)	
		4) 50 μL MPD, III	68/72 (96%)	64/72 (90%)	
		5) Control cotton			
89-8	112/176 (64%)	1) Carpet-15 μL unweath.gas	112/112 (100%)	110/112 (98%)	
		2) Carpet-10 μL weath.gas	112/112 (100%)	107/112 (96%)	
		3) Carpet-20 μL mx.gs/fuel oil	112/112 (100%)	33/112 (29%)	
90-8	127/196 (65%)	1) 25 μL lantern fluid/charred carpet	122/127 (96%)	121/127 (95%)	(Per Report 90-8, LPD, MPD, gasoline and Coleman fuel all considered correct)
		2) 25 μL lantern fluid/uncharred carpet	124/127 (98%)	122/127 (96%)	
		3) Uncharred carpet	120/127 (94%)	7/127 (6%)	Mis id'ed accel
Total	656/1176 (56%)	1670/1826 (91%)	Correctly reported accelerant present/not present.		
		1326/1437 (92%)	Correctly reported accelerant when present.		
		334/388 (89%)	Correctly reported no accelerant when none present.		
		975/1504 (65%)	Correctly identified class of flammable.		

[a]This column lists the percent of labs successfully identifying the class of flammable present; for samples 83-11(3), 84-10(2), 87-9(1 and 2), and 90-8(3) the figure is the percent of laboratories that reported a flammable when none was present.

of class 2 and 5. In this mixture of gasoline and diesel fuel, only 38% of laboratories identified both. As with earlier tests, the LPDs gave the laboratories the greatest challenge, given the infrequency with which they are found in the laboratory and their few identifying characteristics. In about half the chromatograms submitted in cases where labs failed to identify LPDs, the PAC found peaks present to justify an identification. Laboratories also had difficulty in interpreting the evaporated gasoline sample with several misinterpreting chromatograms as being heavy petroleum distillates. The PAC found that many chromatograms submitted did not show the presence of gasoline and the PAC surmised the gasoline had been lost during the "concentration step" of the extraction procedure. It was clear from these results that many laboratories needed to reexamine their procedures and upgrade examiner qualifications.

Test 89-8 included another sample containing a mixture of gasoline and fuel oil. Whereas >96% of respondents correctly detected and identified the gasoline present in samples 1 and 2, and 97% identified the gasoline in the third, only 29% also found the fuel oil in sample 3. Identifying the components of such mixtures is a challenge, pushing the state of the art of forensic examinations. Still, the PAC recommended the employment of improved accelerant recovery and identification procedures in order to distinguish grades of fuel oils in such mixtures.

Laboratories generally performed well on test 90-8, correctly reporting the presence/absence of Coleman lantern fluid on charred and uncharred carpeting in 96% of replies. In the single sample of carpet issued without a flammable liquid 6% incorrectly reported finding an accelerant. The accuracy of results did not correlate with years of experience/time devoted to this type of analysis, or the particular recovery/identification method utilized.

Test 91-9 presented a different scenario in the series of flammable analysis in which laboratories were asked to determine if a substance (napalm) found at a crime scene could have originated from the same source as a similar substance found in a suspect's residence, and if a liquid (gasoline) found in the suspect's garage could have been used to manufacture both batches of napalm. While both samples of napalm originated from a common source, the gasoline found at the scene was not used in the manufacture of either (napalm) substance. While about 90% of respondents correctly answered the first question as to the common origin of the napalm, only about 49% correctly reported the gasoline mixture could not have been used to manufacture either napalm sample.

The single flammables test in the earlier LEAA study did not lend itself to comparison with these later results.

Fibers

Laboratories were issued a total of thirteen fiber tests during the period covered by this review (culminating in 91-7), nine of which yielded identification results. Four of the tests (7, 86-8, 88-10, and 89-6) asked laboratories only to identify fibers. The results of comparative fiber examinations will be discussed in the next article (Part II). The number of laboratories subscribing to the tests rose from 67 in the early years (1978) to 199 in 1991, which represents almost a tripling of participants. The percentage of subscribing laboratories that actually responded with data remained steady, averaging 46% for the years covered.

In the exercises where laboratories were asked to identify fibers they were also asked to report the methods they employed and the information developed from each of the methods. Laboratories performed relatively well in the identification exercises, on average correctly identifying about 89% of fibers issued to them (Table 9). Respondents clearly performed better identifying synthetic fibers (a successful identification rate of 94%) than they did natural fibers (a type of fiber evidence only occasionally submitted for analysis) where the success rate was 77%. In test 88-10, less than half of the laboratories correctly identified all five natural fibers. It should be noted that many of the respondents in this latter test acknowledged their lack of experience with these fiber types and under actual casework conditions would have referred the case to an outside expert ("botanist skilled in plant fiber identification").

Explosives

Laboratories were presented with a total of four explosives proficiency tests from 1980 through 1985. They were asked to identify suspected explosive residues and to specify the methods used to determine their responses. The number of laboratories subscribing to these tests actually decreased over the period of testing, declining from 80 to 71 (Table 10). The average participation rate (25%) was substantially lower than on other tests; however, the rate steadily increased from an initial low of 14% to a concluding rate of 37%. Laboratories posted an average overall rate of successful identifications of about 81% (Table 10). While laboratories generally performed well on these tests, few laboratories were able to identify *all* explosive components, particularly when two or more were mixed together.

In the first test (80-8) laboratories were asked to identify each of five samples and to report methods employed. There were three explosives (PETN, RDX, TNT), one oxidizer ($NaClO_4$), and one sample containing starch. Ten laboratories reported results and supplied a total of fifty responses. Overall, laboratories were correct in 41/50 (82%) of these responses. Success in identifying the samples ranged from a low of 60% on the first sample (PETN—Nitroxyl Ester) to 90% on three others. The most common methods employed were spot tests, IR spectrophotometry, thin-layer chromatography and optical crystallography.

In the second test (81-8), laboratories were again requested to report the identity of five suspected explosive residues. Less than half (44%) of the 17 reporting laboratories correctly and completely identified all five mixtures (a total of 85 potential identifications). Laboratories experienced their greatest success in identifying the $KMNO_4$ and NH_4NO_3—aluminum powder with 100% and 71% correct identification rates, respectively. Where the criteria for correct responses is slightly relaxed and partial identifications are included, a much higher percentage (89%) of replies would be included as partially or totally correct.

In test 82-7 laboratories were issued a white powder and were asked if it was an explosive mixture and to identify its components. About two-thirds of the 15 laboratories responding with data indicated it was an explosive mixture and one-third reported it was not. This question (Is it an explosive mixture?) proved to be somewhat ambiguous and dependent upon what the various laboratories deemed the minimum necessary to constitute an explosive sample. Although about two-thirds of the laboratories successfully identified one or more of the four major components in the sample, only a third identified them all.

In test 85-10, three samples of debris from an explosion were submitted to laboratories and they were asked if an explosive substance was present and, if so, what type. Two of the samples contained a mixture of black and smokeless powder (Hercules Red Dot) and soil, and the third contained only soil and a small amount of lawn fertilizer (ammonium nitrate). Nineteen (19) (73%) laboratories identified both (black and smokeless) powders in the first sample, but only 13 (58%) reported both for the second sample. The same number (13) correctly identified both powders in both samples. More than a quarter of responding laboratories (7/26) did not report finding black powder in either of the samples. The PAC observed that the difficulty laboratories experienced in identifying the black powder in the second sample might be attributed to the presence of (similar appearing) soil particles. In contrast the smokeless powder was easily distinguished by the red dots on some of the individual wafers. Overall, laboratories successfully identified explosives present at a rate of about 70%. Ninety-six percent (25/26) of labs correctly reported the third sample had no explosive present.

Hair

Between 1980 and 1991 a total of eight hair proficiency tests were issued to crime laboratories. The number of laboratories subscribing to these tests grew from 82 to 147 (an increase of about 80%) during the testing, while the number of laboratories responding with data increased from 25 to 67. The overall response rate of laboratories was 41% with, as with other areas of testing, the reply rate in the latter half of testing (46%) substantially higher than during the initial stages (28%). Laboratories were given different exercises, ranging from identification of species of origin of different animal hairs, the area of the (human) body where hairs originated (Table 11), to more common scenarios where laboratories were asked to determine if two or more human hairs could have shared a common origin (see the following article).

Laboratories had difficulty determining the proper species of origin of animal hair, an exercise many crime laboratories would

TABLE 9—*Fiber identification.*

Report	Participation Rate	Fiber Type	Identification Rate	Comments
22	22/71 (31%)	A) Polypropylene B) Polypropylene C) Polypropylene	16/20 (80%) Polypropylene 16/20 (80%) Polypropylene 15/18 (83%) Polypropylene $\bar{X}$ = 47/58 (81%)	
80-6	30/82 (37%)	A) Polyester & Cotton B) Polyester & Cotton C) Polyester & Cotton D) Polyester & Cotton	18/21 (86%) Polyester & Cotton 25/25 (100%) Polyester & Cotton 23/24 (96%) Polyester & Cotton 25/25 (100%) Polyester & Cotton $\bar{X}$ = 91/95 (96%)	
83-6	37/68 (54%)	A) Nylon 6.6 B) Nylon 6.6	35/37 (95%) Nylon 35/37 (95%) Nylon $\bar{X}$ = 70/74 (95%)	16/37 (43%) id'ed Nylon 6.6/420 15/37 (41%) id'ed Nylon 6.6
84-4	52/93 (56%)	A) Nylon B) Nylon C) Nylon	39/40 (98%) Nylon 39/39 (100%) Nylon 39/39 (100%) Nylon $\bar{X}$ = 117/118 (99%)	
85-4	48/105 (46%)	S-1 Polyester S-2 Polypropylene K-1 Polyester K-2 Polyester	42/42 (100%) Polyester 25/27 (93%) Olefin/Polypropylene 41/41 (100%) Polyester 41/42 (98%) Polyester $\bar{X}$ = 149/152 (98%)	13/27 (48%) id'ed polypropylene
86-8	61/124 (49%)	1) Cotton 2) Nylon 6 3) Polyester	60/61(98%) Cotton 59/61(97%) Nylon 58/61(95%) polyester $\bar{X}$ = 177/183 (97%)	31/61 (51%) id'ed Nylon 6
87-10	54/130 (42%)	S1 Polyester S2 Polyester S3 Nylon K1 Polyester	51/53(96%) 51/53(96%) 37/49(76%) 51/53(96%) $\bar{X}$ = 190/208(91%)	
88-10	62/155 (40%)	S1 Jute S2 Sisal S3 Abaca S4 Silk S5 Ramie	46/62(74%) 50/62(81%) 38/62(61%) 54/62(87%) 37/62(60%) $\bar{X}$ = 225/310(73%)	
89-6	60/153 (39%)	1) Acrylic 2) Modacrylic	56/60 (93%) 48/60 (80%) $\bar{X}$ = 104/120(87%)	
Total	426/981 (43%)		1170/1318 = 89%	

not ordinarily undertake. In test 80-12, 44% (on average) of respondents correctly identified deer, opossum, and black bear hair samples. In test 81-10, only 30% correctly identified the species of origin of moose, grey squirrel and grey fox hair. Were correct responses to include correct "family of origin" conclusions (for example, the deer family as opposed to specific mention of "moose") the correct identifications rose above 50%. Many laboratories objected to both exercises because samples did not include the hair root, which would have facilitated the identification process.

In a subsequent test, (85-6), of five hairs of unknown origin to be compared with samples taken from the victim and two suspects, one originated from a German Shepherd dog. Because species of origin was not specifically requested in this exercise, it is not possible to offer a percent figure; however, all laboratories noted it was of animal origin and most said it originated from a dog. Many laboratories, again, reported they felt handicapped by the absence of roots on the dog hair specimens. There was a fourth exercise (87-6), a very difficult test, in which labs attempted to determine the location on the human body where various hairs had originated. Overall, labs were correct in 56% of their designations, having much greater success (86% accuracy) identifying human hair as being of head and pubic origin, but much lower success (30%) in noting the origin of beard, arm and chest hair. Very

TABLE 10—*Explosives.*

Report	Participation Rate	Explosive Type	Identification Rate	Comments
80-8	11/80 (14%)	A) PETN	6/10	
		B) RDX	9/10	
		C) TNT	8/10	
		D) Starch	9/10	
		E) $NaClO_4$	9/10	
		Subtotal	41/50 (82%)	(Total of 6/50 misidentifications)
81-8	17/67 (25%)	A) $KMnO_4$	(17/17)	
		B) Baratol (80%TNT 20%$BaNO_3$)	(1/17) 14/17[a]	
		C) Hercules Bullseye gunpowder	(1/17) 13/17[a]	
		D) $KClO_3$-KIO_4 (60:40 mixture)	(6/17) 16/17[a]	
		E) NH_4NO_3-Aluminum Powder	(12/17) 17/17[a]	
		Subtotals	37/85 (44%) 76/85 (89%)[a]	( ) Complete identification [a]Partial identification
82-7	15/56 (27%)	A) KNO_3 (25%)		
		B) $KClO_4$ (25%)		
		C) $Sr(NO_3)_2$ (25%)		
	Mixture of	D) $Ba(NO_3)_2$ (25%)		
		E) Corn starch (minor)		
		F) Diatomaceous earth (minor)		
		Subtotal	9/15 (60%)	Explosive mixture
85-10	26/71 (37%)	1) Wood/debris, black powder, smokeless powder	19/26 (73%)	no explosive present
		2) Soil, black powder, smokeless powder	15/26 (58%)	
		3) Soil, lawn fertilizer (ammonium nitrate)	25/26 (96%)	
		Subtotal	59/78 (76%)	
Total	69/274 (25%)		185/228 (81%)	

TABLE 11—*Hair identification.*

Report	Participation Rate	Hair Type	Identification Rate	Comments
80-12	25/82 (30%)	A) White tailed deer	19/25(76%)	70/75(93%) Correctly labelled hairs "non-human"
		B) Opossum	7/25(28%)	
		C) Black Bear	7/25(28%)	
		Subtotal	33/75(44%)	
				Correct family
81-10	22/79 (28%)	A) Moose	8/22(36%)	20/22(91%) Deer Family
		B) Grey Squirrel	10/22(45%)	10/22(50%) Squirrel Family
		C) Grey Fox	2/22(9%)	4/22(18%) Fox/Canidae
		Subtotal	20/66(30%)	34/66(52%)
85-6	52/109 (48%)	1) Animal (dog) hair	52/52(100%)	
		5) Dyed human head hair	52/52(100%)	
		Subtotal	104/104(100%)	
			Correct Species	Correct Body Area
87-6	57/106 (54%)	1) Human head hair	57/57(100%)	50/57(87%)
		2) Human pubic hair	57/57(100%)	48/57(84%)
		3) Human beard hair	56/57(98%)	19/57(33%)
		4) Human arm hair	56/57(98%)	23/57(40%)
		5) Human beard hair	37/57(65%)	14/57[a](25%)
		6) Human chest hair	57/57(100%)	14/57[b](25%)
		7) Human (wig) head hair	55/57(96%)	55/57(96%)
		Subtotal	375/399(94%)	223/399(56%)
Total	156/376 (41%)		532/644(83%)	

[a]Beard/facial.
[b]Chest/body.

TABLE 12—*Drug analysis.*

Drug	Report	Name	Concentration	Participation Rate	Identification Rate	Comments
Opiates	2	Heroin	4%	56/79(71%)	Not in Report	
	23	Heroin	.86%	28/76(37%)	Informed it was heroin	
	80-1	Pentazocine	30 mg/mL	43/88(55%)	41/43(95%)	
	82-1	Morphine	3%	33/88(38%)	31/33(94%)	
	83-2[a]	Heroin	a. 3.5% b. 5.0%	50/61(82%)	50/50(100%) 49/50(98%)	
	86-1[b]	Heroin Cocaine Procaine	10% 5% 74%	98/147(67%)	93/98(95%) 81/98(83%) 84/98(86%)	
	91-13[b]	Heroin Procaine	5.5% 10.0%	192/256(75%)	192/192(100%) 181/192(94%)	
Subtotal				500/795(63%)	537/564(95%)	Excludes two noncontrolled procaine components
Stimulants	13	Cocaine	8%	57/83(69%)	57/57(100%)	
	20	Phendimetrazine	28%	48/83(58%)	47/48(98%)	
	84-8	Cocaine	20%	78/119(65%)	78/78(100%)	
	85-1	Phendimetrazine	...	101/115(88%)	99/101(98%)	
	85-9[a]	Tetracaine HCl Cocaine HCl	a. 15.6% b. 15.4%	86/137(63%)	85/86(94%) 86/86(100%)	
	86-9	Methamphetamine HCl	13%	88/148(59%)	88/88(100%)	
	87-1[b]	Cocaine base Ephedrine	37.9% 5.3%	101/158(64%)	101/101(100%) 39/101(39%)	
	88-1[b]	Cocaine base Cocaine HCl	10% 20%	116/192(60%)	116/116(100%)	
	89-4[b]	Cocaine base Methamphetamine HCl	12% 5%	123/224(55%)	117/123(95%) 113/123(92%)	
	89-12	N,N-dimethylamphetamine HCl	15.5%	135/224(60%)	116/135(86%)	
	90-12[b]	Cocaine HCl Nicotinamide	39.9% 9%	145/235(62%)	145/145(100%) 76/145(52%)	
Subtotal				1078/1718 (63%)	1163/1201(97%)	(Excludes tetracaine (85-9), ephedrine (87-1), and nicotinamide (90-12))
Depressants	82-6	Glutethimide	15%	39/84(46%)	35/39(90%)	
	84-1	Methaqualone	...	52/95(55%)	52/52(100%)	
Subtotal				91/179(51%)	87/91(96%)	
Halluci-nogens	5	Phencyclidine (PCP)	8%	63/79(80%)	Not in report	
	83-7[a]	Cannabis Catnip Extract		55/87(61%)	a. 55/55(100%) b. 55/55(100%) c. 27/55 (47%)	
	87-8	N-Ethyl M.D.A. HCl	14.4%	101/171(59%)	81/101(81%)	
	88-8	TCP HCl morpholine analog	9%	118/216(55%)	108/118(92%)	
	91-5[a]	LSD Lysergol		157/238(66%)	148/157(94%) 57/157(36%)	
Subtotal				494/791(62%)	419/486(86%)	(Excludes lysergol and catnip)
Other	80-5	Diphenhydramine HCl	20%–29%	58/90(64%)	58/58(100%)	
	81-1[b]	Ephedrine Theophylline	20% 20%	45/81(55%)	36/45(80%) 39/45(87%)	
	81-7	PCC (PCP analog)	2%	34/81(42%)	7/34(22%)	
	90-4[b]	Testosterone Propionate Testosterone Cypionate Testosterone Enanthate	50 mg/mL 50 mg/mL 50 mg/mL	133/220(60%)	113/133(85%) 72/133(54%) 74/133(56%)	
Subtotal				270/472(57%)	399/581(69%)	

[a]Separate samples.
[b]Mixture.

few laboratories would even attempt this latter exercise in an actual investigation.

In sum, laboratories were no more successful in identifying the correct species of origin of animal hair (a relatively uncommon request made to crime laboratories) than they were in the earlier LEAA study. The average successful identification rate remains at about the 50% level.

Summary

This first article on proficiency testing of crime laboratories covering the period 1978–1991 reveals a wide range of results. Laboratories meet with the greatest success in the examination of bloodstains and controlled substances. There is a very low percent of misidentifications or false positives in both categories. Laboratories are moderately successful identifying/classifying flammables and fibers. Of concern for flammables is the sizeable percent of false positive results; in the fibers category, natural fibers pose the major challenge. Laboratories are only moderately successful identifying various explosives. Animal and human (body area) hair identifications are clearly the most troublesome of all categories tested, and for which labs successfully identify the animal species or area of the human body from which the hair originated only about 50% of the time.

Acknowledgments

We wish to acknowledge the assistance of David Stoney in developing the original design for classification of the test results. We also acknowledge the cooperation and assistance of Charles Leete and Collaborative Testing Services for providing copies of individual test reports.

References

[1] Peterson, J. L., Fabricant, E. L., Field, K. S., and Thornton, J. I., *Crime Laboratory Proficiency Testing Research Program,* U.S. Government Printing Office, Washington, DC, 1978.
[2] Ostrow, R., "Many Crime Labs Fail Tests: Evidence in Trials May Be Hurt," *Los Angeles Times,* 9 April 1977, p. 1.
[3] For example, see: E. J. Imwinkelried, Ed., *Scientific and Expert Evidence* (2nd edition) Practising Law Institute, New York City, 1981, p. 32; Moenssens, A. A., Inbau, F. E., and Starrs, J. E., *Scientific Evidence in Criminal Cases,* 3rd edition, The Foundation Press, Inc., Mineola, N.Y., 1986, pp. 5–6. P. C. Giannelli and E. J. Imwinkelried, *Scientific Evidence,* 2nd edition, The Michie Company, Charlottesville, VA, 1993, p. 187.
[4] Giannelli, P. C. "The Admissibility of Laboratory Reports in Criminal Trials: The Reliability of Scientific Proof," *Ohio State Law Journal,* Vol. 49, No. 1, 1988, p. 689.
[5] Bashinski, J., "Laboratory Standards: Accreditation, Training, and Certification of Staff in the Forensic Context," *DNA Technology and Forensic Science,* J. Ballantyne, G. Sensabaugh, and J. Witkowski, Eds., Cold Spring Harbor Press, New York, 1989, pp. 159–171.
[6] Brunelle, R. L., Garner, D. D., and Wineman, P. L., "A Quality Assurance Program for the Laboratory Examination of Arson and Explosives Cases," *Journal of Forensic Sciences,* Vol. 27, No. 4, Oct. 1982, pp. 774–782.
[7] Pereira, M., "Quality Assurance in Forensic Science," *Forensic Science International,* Vol. 28, 1985, pp. 1–6.
[8] *Proficiency Testing and Quality Assurance Workshop,* Annual Meeting of the American Academy of Forensic Sciences, Las Vegas, 1989.
[9] Werrett, D. J., Lygo, J. E., and Sutton, J. G., "The Introduction of DNA Analysis into Home Office Forensic Science Laboratories in England and Wales," *DNA Technology and Forensic Science,* J. Ballantyne, G. Sensabaugh, and J. Witkowski, Eds., Cold Spring Harbor Laboratory Press, New York, 1989, pp. 233–241.
[10] Lucas, D. M., Leete, C. G., and Field, K. S., "An American Proficiency Testing Program," *Forensic Science International,* Vol. 27, 1985, pp. 71–79.
[11] Sensabaugh, G. F., "Genetic Typing of Biological Evidence: Comments for the Cooper Amicus Brief," *California Association of Criminalists Newsletter,* July 1987, pp. 11–17.
[12] Grunbaum, B., "Physiological Stain Evidence—Assuring Quality Analysis," *Cal Defender,* Spring 1985, p. 20, cited by Jonakait, R., "Forensic Science: The Need for Regulation," *Harvard Journal of Law and Technology,* Vol. 4, Spring 1991, p. 119.
[13] Grunbaum, B., personal communication, cited in U.S. Congress, Office of Technology Assessment, *Genetic Witness: Forensic Uses of DNA Tests,* OTA-BA-438, U.S. Government Printing Office, Washington, DC, 1990, p. 79.
[14] Risinger, D. M., Denbeaux, M. P., and Saks, M. J., "Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification 'Expertise,'" *University of Pennsylvania Law Review,* Vol. 137, No. 3, Jan. 1989, pp. 731–792.
[15] Kam, M., Wetstein, J., and Conn, R., "Proficiency of Professional Document Examiners in Writer Identification," *Journal of Forensic Sciences,* Vol. 39, No. 1, Jan. 1994, pp. 5–14.
[16] Miller, L. S., "Procedural Bias in Forensic Science Examinations of Human Hair," *Law and Human Behavior,* Vol. 11, 1987, pp. 157–163.
[17] Sunderman, F. W., "The History of Proficiency Testing/Quality Control," *Clinical Chemistry,* Vol. 38, No. 7, 1992, pp. 1205–1209.
[18] Laessig, R. H., Ehrmeyer, S. S., Lanphear, B. J., Burmeister, B. J., and Hassemer, D. J., "Limitations of Proficiency Testing Under CLIA '67," *Clinical Chemistry,* Vol. 38, No. 7, 1992, pp. 1237–1244.
[19] Schaeffer, M., Ed., *Federal Legislation and the Clinical Laboratory,* G. K. Hall Medical Publishers, Boston, MA, 1981.
[20] Peddecord, K. M. and Hammond, H. C., "Clinical Laboratory Regulation Under the Clinical Laboratory Improvement Amendments of 1988: Can It Be Done?" *Clinical Chemistry,* Vol. 36, No. 12, 1990, pp. 2027–2035.
[21] Rej, R. and Jenny, R. W., "How Good Are Clinical Laboratories? An Assessment of Current Performance," *Clinical Chemistry,* Vol. 38, No. 7, 1992, pp. 1210–1217.
[22] Hinkel, C., "Regulation of Proficiency Testing Under the March 14 Rule," *Clinical Chemistry,* Vol. 38, No. 7, 1992, pp. 1251–1253.
[23] Nakamura, R. M., "Quality Assurance and Proficiency Testing in Molecular Oncology," *Cancer* (suppl.) Vol. 69, No. 6, 1992, pp. 1633–1635.
[24] Willette, R. E., "Drug Testing Programs," in *Urine Testing for Drugs of Abuse,* R. L. Hawks and C. N. Chiang, Eds., NIDA Research Monograph 73, Department of Health and Human Services, Rockville, MD, 1986, pp. 5–12.
[25] Kalita, S. J., *Proficiency Testing of Laboratories Performing Analyses of Drugs of Abuse,* Master of Science Thesis in Criminalistics, University of Illinois at Chicago, Chicago, IL, 1987.
[26] Davis, K. H., Hawks, R. L., and Blanke, R. V., "Assessment of Laboratory Quality in Urine Drug Testing: A Proficiency Testing Pilot Study," *Journal of the American Medical Association,* Vol. 260, No. 12, Sept. 1988, pp. 1749–1754.
[27] Willette, R. E., "Proficiency Testing and Quality Control Programs," in *Urine Testing for Drugs of Abuse,* R. L. Hawks and C. N. Chiang, Eds., NIDA Research Monograph 73, Department of Health and Human Services, Rockville, MD, 1985, pp. 20–23.
[28] Guerrant, G. O. and Hall, C. T., "Drug Abuse Proficiency Testing," *Clinical Toxicology,* Vol. 10, No. 2, 1977, pp. 209–219.
[29] McCormick, W., Ingelfinger, J., Isakson, G., and Goldman, P., "Errors in Measuring Drug Concentrations," *New England Journal of Medicine,* Vol. 299, 16 Nov. 1978, pp. 1118–1121.
[30] Boone, D. J., Hansen, J. H., Hearn, T. L., et al., "Laboratory Evaluation and Assistance Efforts: Mailed, On-Site, and Blind Proficiency Testing Surveys Conducted by the Centers for Disease Control," *American Journal of Public Health,* Vol. 72, 1982, pp. 1364–1368.
[31] Morgan, J. P., "Problems of Mass Urine Screenings for Misused Drugs," (letter) *Journal of Psychoactive Drugs,* Vol. 16, No. 4, 1984, pp. 305–317.
[32] Hansen, H. J., Caudill, S. P., and Boone, J., "Crisis in Drug Testing," *Journal of the American Medical Association,* Vol. 253, No. 16, 26 April 1985, pp. 2382–2387.
[33] Frings, C. S., White, R. M., and Battaglia, D. J., "Status of Drugs of Abuse Testing in Urine: An AACC Study," *Clinical Chemistry,* Vol. 33, No. 9, 1987, pp. 1683–1686.

[34] Kelly, K. L., "The Accuracy and Reliability of Tests for Drugs of Abuse in Urine Samples," *Pharmacotherapy,* Vol. 8, No. 5, 1988, pp. 263–275.
[35] Dinovo, E. C. and Gottschalk, L. L., "Results of a Nine-Laboratory Survey of Forensic Toxicology Proficiency," *Clinical Chemistry,* Vol. 22, 1976, pp. 843—846.
[36] Dinovo, E. C. and Gottschalk, L. L., "More on Proficiency Testing in Forensic Toxicology," *Clinical Chemistry,* Vol. 22, No. 12, 1976, p. 2056.
[37] Kelly, R. C. and Sunshine, J., "Proficiency Testing in Forensic Toxicology: Criteria for Experimental Design," *Clinical Chemistry,* Vol. 22, No. 8, 1976, pp. 1413–1414.
[38] McCloskey, K. L. and Finkle, B. S., "Proficiency Testing in Forensic Toxicology," letter *Journal of Forensic Sciences,* Vol. 22, No. 4, Oct. 1977, pp. 675–678.
[39] Peat, M. A., Finnigan, J. S., and Finkle, B. S., "Proficiency Testing in Forensic Toxicology: A Feasibility Study," *Journal of Forensic Sciences,* Vol. 28, No. 1, Jan. 1983, pp. 139–158.
[40] Caplan, Y., "Drug Testing in Urine," Workshop on Drug Use Testing: Forensic Science and Legal Issues, *Journal of Forensic Sciences,* Vol. 34, No. 6, Nov. 1989, pp. 1417–1421.
[41] Department of Health and Human Services, "Mandatory Guidelines for Federal Workplace Drug Testing Programs," *Federal Register,* Vol. 53, No. 69, 11 April 1988, pp. 11970–11989.
[42] Visher, C. and McFadden, K., *A Comparison of Urinalysis Technologies for Drug Testing in Criminal Justice,* June 1991, National Institute of Justice, Washington, DC.
[43] *Laboratory Accreditation Board Manual.* American Society of Crime Laboratory Directors, Laboratory Accreditation Board, Jan. 1994.
[44] Technical Working Group on DNA Analysis Methods (TWGDAM), "Guidelines for a Quality Assurance Program for DNA Analysis," *Crime Laboratory Digest,* Vol. 22, No. 2, April 1995, pp. 33–37.
[45] Valadez, M., Jr., personal communication, 13 July 1995.
[46] *ABC Policy Manual,* Proficiency Testing Program, American Board of Criminalistics, 1995, pp. L1–L25.
[47] U.S. Congress, Office of Technology Assessment, *Genetic Witness: Forensic Uses of DNA Tests,* OTA-BA-438, U.S. Government Printing Office, Washington, DC, 1990, pp. 80–81.
[48] Thompson, W. C. and Ford, S., "The Meaning of a Match: Sources of Ambiguity in the Interpretation of DNA Prints," in *Forensic DNA Technology,* M. A. Farley and J. J. Harrington, Eds., Lewis Publishers, Inc., Chelsea, MI, 1991, pp. 142–145.
[49] Jonakait, R., "Forensic Science: The Need for Regulation," *Harvard Journal of Law and Technology,* Vol. 4, Spring 1991, p. 178.
[50] Letters Forum: Forensic Labs, *Harvard Journal of Law and Technology,* Vol. 5, Fall 1991, pp. 1–17.
[51] Leete, C., personal communication, 12 July 1995.
[52] Language taken from selected laboratory test reports, 1978-1991.

Address requests for reprints or additional information to
Joseph L. Peterson, D.Crim.
Department of Criminal Justice
University of Illinois-Chicago
1007 W. Harrison St.
Chicago, IL 60607