How to Improve Forensic Science*

(Forthcoming, 2005, ***European Journal of Law and Economics***)

Roger Koppl
Professor of Economics and Finance
Fairleigh Dickinson University
Madison, NJ 07940
USA

World Wide Web: http://inside.fdu.edu/pt/koppl.html
Internet: koppl@fdu.edu
Phone: (973) 443-8846
Fax: (973) 443-8377

* For useful comments I thank Richard Adelstein, Brian Arthur, Nigel Ashford, Chris Barker, John Becker, Don Bellante, Bruce Benson, Paul Boyer, William Butos, Bruce Caldwell, Deby Cassill, Mark Casson, Young Back Choi, David Colander, Paul Dower, Lester Embree, Alan Fask, David Friedman, Gloria Gadsden, Giampaolo Garzarelli, Adam Gifford, David Harper, Randall Holcombe, Sandy Ikeda, Lorenzo Infantino, Lawrence Kobilinsky, Chidem Kurdas, Brian Loasby, Ray Majewski, Thomas Marmefelt, Thomas McQuade, Jonathon Mote, Robert Mulligan, Meike Niedbal, Maria Paganelli, Francesco Parisi, Nelson Pole, Jonas Prager, David Prychitko, George Richardson, Mario Rizzo, Paul Rubin, Joe Salerno, John Schiemann, Richard Swedberg, Martti Vihanto, Richard Wagner, Glen Whitman, and two anonymous referees. I also benefited from conversations with Ron Calissi, Mark Campbell, Daniel Hausman, Robert Kurzban, Kevin McCabe, Maria Minniti, David Porter, David Rosen, Chris Sanchirico, Kenneth Vehrkens, and John Vanderkolk. Thanks, finally, to Vanessa Potkin of the Innocence Project for providing some source material. None of these people is responsible for any errors or insufficiencies of the paper.

Abstract

Some institutional structures for inquiry produce better approximations to truth than others. The current institutional structure of police forensics gives each lab a monopoly in the analysis of the police evidence it receives. Forensic workers have inadequate incentives to produce reliable analyses of police evidence. Competition would create such incentives. I outline a system of "competitive self regulation" for police forensics. Each jurisdiction would have several competing forensic labs. Evidence would be divided and sent to one, two, or three separate labs. Chance would determine which labs and how many would receive evidence to analyze. Competitive self regulation improves forensics by creating incentives for error detection and reducing incentives to produce biased analyses.

I. Introduction

The proper function of forensic science is to extract the truth. "As it is practiced today," note Saks *et al*. (2001), "forensic science does not extract the truth reliably. Forensic science expert evidence that is erroneous (that is, honest mistakes) and fraudulent (deliberate misrepresentation) has been found to be one of the major causes, and perhaps the leading cause, of erroneous convictions of innocent persons" (Saks *et al*. 2001, p. 28).

In the wake of DNA exonerations, an extensive literature has developed on the limited reliability of forensic testimony in the courts. (See, for example, Jonakait, 1991; Moenssens, 1993; Giannelli, 1997; Office of the Inspector General, 1997; Kaufman, 1998; Kelly and Wearne, 1998; Saks, 1998; Saks *et al*., 2001; Illinois, 2002; and Risinger *et al*., 2002.) Previous authors in this literature have recommended salutary reforms such as independence of forensic labs from law-enforcement agencies (Illinois, 2002; Gianelli 1997), improved documentation of forensic work (Kaufman, 1998), double-blind proficiency tests (Risinger *et al*., 2002), and the use of evidence line-ups (Miller, 1987; Risinger *et al*., 2002). These reforms, however, will have limited effect without further reform in the institutional structure of forensic work.

The institutional structure of forensic work is an important source of forensic error, insufficiency and, sometimes, malfeasance. The forensic worker has an effective monopoly on the analysis of the evidence presented to him and is therefore in a position to infer from the evidence what he pleases. Past calls for reform seem to have neglected

both the role of industrial organization in discouraging high-quality forensics and the importance of competition in the supply of forensic services.

I propose breaking up the forensic worker's monopoly by instituting "competitive self regulation." Competitive self regulation would put forensic labs into a competition similar to the competition characterizing pure science. Each forensic lab becomes a check on every other forensic lab. This system of checks and balances would reduce the errors committed by forensic scientists. It would even work to reduce the conscious and unconscious abuses committed by some forensic workers.

Science is a social process in which the truth emerges from a rule-governed competitive process (Merton, 1957; Polanyi, 1962; Koppl and Butos, 2003; and McQuade and Butos, 2003). It is a competitive process in which knowledge is public, the idiosyncrasies of individual researchers are checked by the results of other workers, and results are subject to criticism, review, and reproduction. As it is practiced today, forensic science departs needlessly from this model. Forensic analysis often depends too much on the personal qualities of each individual forensic scientist. Idiosyncrasies of individual forensic scientists may determine the final result, and there is limited criticism, review, and reproduction. A competitive process of self regulation is constantly at work eliminating errors in pure science. No such process is at work in forensic science. Pure science is self regulating, forensic science is not.

The differences between pure science and forensic science concern institutional structure. Forensic science is sometimes unreliable because the larger environment of knowledge seeking is not appropriately structured. Most forensic scientists are skillful and diligent practitioners of their trade. They find themselves in an environment,

however, that does not encourage the sort of institutional self-criticism characterizing pure science. They are in an environment that can induce unconscious bias and even give the unscrupulous an incentive to lie. If competitive self regulation has value, it is because it provides a better institutional structure for truth seeking and knowledge production.

Under competitive self regulation, forensic science would finally become "forensic" in the truest sense. The word "forensic" comes from the Greek word for the forum, where citizens would come to dispute over public questions. The other current meaning of "forensics" is "The art or study of formal debate; argumentation" (*The American Heritage Dictionary of the English Language*, 4th edition). In the current system, it is often difficult to challenge the analysis of a police forensic worker, even for the defense. In this sense, the forensic worker has power. The adversarial system of our criminal courts organizes a dispute between the prosecution and the defense. But the current institutional structure of forensic work put the results of forensic workers largely beyond dispute.

I begin with a brief outline of the argument in which I highlight those structural features of the current situation impeding better performance of forensic science. The following section is devoted to bias, especially unconscious bias, and how it might be eliminated. Then I outline my proposals, which are broad principles for improving forensics. I do not have detailed proposals. Drafting detailed proposals is a significant labor well beyond the scope of this paper. My principles for improving forensics might be called principles of "forensic science administration." The field of forensic science administration is largely unexplored. The systematic and scientific study of forensic science administration is, however, more urgently required than the similar study of

public administration, general business administration, or healthcare administration. Thus, in the conclusion I call for further work in the area. Appendix I reviews evidence that forensic practice is well below the level we might reasonably expect. It shows, in other words, that I am dealing with a real and important problem.

II. Overview

There are eight features of the organization of forensic science that needlessly reduce the quality of work performed by forensic scientists.

1) **Monopoly.** In most jurisdictions today, including those in the US, each laboratory has a monopoly on the evidence it analyzes. No other lab is likely to examine the same evidence.

2) **Dependence.** Forensic labs are often organized under the police and are thus dependent on the police for their budgets.

3) **Poor quality control.** Quality control systems tend to be weak. In the US, there are no required programs of accreditation and the principal accrediting agency, the American Society of Crime Lab Directors, is a professional group, not an independent organization.

4) **Information sharing.** Forensic scientists are privy to information that may be crucial to a criminal proceeding, but extraneous to the questions put to the forensic scientist.

5) **No division of labor between forensic analysis and interpretation.** The same scientist who, say, performs a test to establish blood type judges whether his test results exclude the police suspect.

6) **Lack of forensic counsel.** Indigent defendants rarely receive aid and counsel from forensic scientists hired to help them. In common law countries this creates an asymmetry whereby the prosecution has forensic counsel and, indeed, sometimes great batteries of forensic specialists, whereas the defense has no forensic counsel and, often, an attorney unable to adequately understand and challenge forensic testimony.

7) **Lack of competition among forensic counselors.** From the absence of forensic counsel for the indigent it follows rather trivially that there is no competition among forensic counselors for their custom.

8) **Public ownership.** Forensic laboratories are publicly owned. In the US often organized under police agencies such as the State police or the FBI.

Each of the features just listed discourages unbiased work of high quality. I review each in turn.

1) **Monopoly.** The monopoly position of forensic labs allows them to do sloppy, biased, even fraudulent work. As Appendix I shows, recent history provides quite a few examples of poor work. The case of Ralph Erdman illustrates how careless forensic work can sometimes be. As I note in Appendix I, this forensic scientist not only faked many of his results, but even managed somehow to loose a human head from a body he was to examine (Kelly and Wearne 1998, p.13). To describe such work as "sloppy" is an understatement.

2) **Dependence**. Dependence creates a pro-prosecution bias. For example, David Williams, an FBI investigator in the Explosives Unit was found to have "tailored" his testimony "to the most incriminating result" in two separate trials, namely, that of World Trade Center Bombing of 1993 and that of the Oklahoma City Bombing of 1995. In the

Oklahoma bombing case, "Williams repeatedly reached conclusions that incriminated the defendants without a scientific basis and that were not explained in the body of the report" (Office of the Inspector General 1997).

3) **Quality control.** Quality control measures tend to be poor, which may easily produce persistently poor work. In Scotland, for example, detective constable Shirley McKie was charged with murder on the basis of a fingerprint identification that has been shown to be false and mistaken. An investigation into the case by the Scottish politician and former criminal defense lawyer Winnie Ewing seems to show that "the system of gathering fingerprints in Scotland by the Scottish Criminal Records Office (SCRO) was less stringent than that used in India" (McBeth, 2004). In other words, the SCRO did not have an effective quality control system. The Houston case, which is discussed in Appendix I, provides another and very dramatic example, as does the Seattle case (Teichroeb 13 March 2004), which I will not examine closely.

4) **Information sharing.** Information sharing between police investigators and forensic scientists creates the strong possibility of unconscious bias (Risinger *et al*., 2002). It also helps dishonest scientists to act on any self-conscious biases they may have. The inappropriate sharing of bias-inducing information might be called "information pollution." I have already mentioned the case of FBI examiner David Williams who identified the explosives used in the Oklahoma bombing of 1995 and the World Trade Center bombing of 1993. Williams' judgment in each case was found to have been determined by the extraneous knowledge of what the suspects had previously purchased and not by legitimate forensic means.

5) **No division of labor between forensic analysis and interpretation.** Forensic error may result from a false interpretation of a test that was properly conducted. In Houston, for example, George Rodriquez was convicted of rape largely on forensic testimony in which a legitimate test was given an illegitimate conclusion. An affidavit by several leading forensic scientists demonstrates that Houston pathologist Jim Bolding interpreted his serological work in ways that were inconsistent with established scientific theory. "Jim Bolding's trial testimony . . . contains egregious misstatements of conventional serology. These statements reveal that either the witness lacked a fundamental understanding of the most basic principles of blood typing analysis or he knowingly gave false testimony to support the State's case against George Rodriguez. His testimony is completely contrary to generally accepted scientific principles" (Blake *et al*. 2004). There seems to have been a similar disparity between test and interpretation in several cases in Cook County Illinois (Mills *et al*. 2004).

6) **Lack of forensic counsel.** The lack of forensic counsel has produced many of the false convictions identified by the Innocence Project in the US and the similar British group, Innocent. (See http://www.innocenceproject.org and http://innocent.org.uk.)

7) **Lack of competition among forensic counselors.** Even if forensic counsel is available, it may not be vigorous or effective in a non-competitive environment.

8) **Public ownership.** As I note in Appendix I, after the DNA work of the Houston Crime Lab (in Texas) was shown to be unreliable, the Houston Police Department began sending all of its DNA work to private labs. This episode, while merely suggestive, nicely illustrates the claim that private labs may have stronger incentives to produce reliable work than do government labs.

The elements of "competitive self-regulation," my suggestion for revising the institutional setting of forensic science, consist in reversing the eight characteristics of the current system that I have noted.

1) **Rivalrous redundancy** should replace monopoly. There should be several competing forensic labs in any jurisdiction. Subject to the constraints of feasibility, some evidence should be chosen at random for duplicate testing at other labs. The same DNA evidence, for example, might be sent to more than one lab for analysis. The forensic worker need not know whether the evidence is examined by another lab. He will know that there could be another lab, and sometimes is.

2) **Independence** should replace dependence. Rivalrous redundancy and privatization would necessarily create independence in at least the formal sense. Competitive self regulation would produce both formal and substantive independence.

3) **Statistical review** would support improved quality control. For example, if a given lab produces an unusually large number of inconclusive findings, its procedures and practices should be examined. Competitive self regulation creates checks and balances.

4) An Evidence Control Officer would substitute **information hiding** for information sharing. Evidence Control Officer would prepare evidence for testing and shield the lab doing a test from all extraneous knowledge of the case particulars. The Evidence Control Officer should use random-number generators to decide which lab gets a given piece of evidence and when to send the same evidence to more than one lab.

5) A **division of labor between forensic analysis and interpretation** should be applied. When this measure is combined with the provision of forensic counsel for the defense, errors of interpretation are less likely to go unchallenged.

6) **Forensic counsel** should be provided for indigent defendants in criminal cases.

7) A system of **forensic vouchers** for indigent defendants would give forensic counselors to the indigent an incentive to provide high-quality services to their clients.

8) **Privatization** should replace public ownership. Private enterprises are subject to civil liability and may be subject to administrative fines for poor performance. They have, therefore, stronger financial incentives than publicly owned enterprises to provide good and reliable work.

Figure 1 summarizes the argument in tabular form.

Current system	**Resulting problem**	**Proposed Solution**
Monopoly	Sloppy, biased, and sometimes fraudulent work	Rivalrous redundancy
Dependence	Bias	Independence
Poor quality control	Persistently poor work	Statistical review
Information sharing	Conscious and unconscious bias	Information hiding
No division of labor between forensic analysis and interpretation	Error from false interpretations of legitimate results.	Division of labor between forensic analysis and interpretation
Lack of forensic counsel	False convictions	Forensic counsel for the indigent
Lack of competition among forensic counselors	Poor quality forensic counsel	Forensic vouchers
Public ownership	Weak financial incentives to provide high-quality work	Privatization

Figure 1

III. Bias

The current organization of forensic work induces bias among forensic workers. I am not arguing that forensic workers enter the system already biased, although that may sometimes happen. Rather, the system induces bias. In some cases induced bias has led to outright fraud. Appendix I contains examples. Perhaps more insidious, however, is the presence of unconscious bias in sincere and conscientious workers.

The proposals suggested later in the paper would have three effects. First, they would reduce the biases discussed in this section. Second, they would mitigate the ill effects of remaining biases. Finally, they would create checks and balances that would use countervailing biases to neutralize one another. The problem is a system that induces biased and sloppy work even if individual forensic workers are sincere and conscientious. The solution is a system that produces unbiased and careful work even if individual forensic workers are disposed toward bias and negligence. [1]

A. The Organization of Forensic Work Creates Bias among Forensic Workers

Most forensic work in the US is performed in police labs, including the FBI lab. Lucas, Leete & Field (1985, p. 72 as cited in Jonakait 1991) report that "about 80%" of US crime labs "are within law enforcement agencies." About 90% of accredited labs in the US today are organized under police agencies (Mills *et al*. 2004). The forensic worker depends on the police (or other law-enforcement agency) for his salary and performance evaluation. This frequently creates a demand for the services of forensic workers who deliver results consistent with the police theory. Forensic workers have often been

[1] I thank Bill Butos with help on the logic of this and the previous paragraph.

former police officers or FBI agents. Until 1994, the FBI crime lab "generally required its examiners to also be FBI agents, except in the Latent Fingerprint section, where the examiners have always been non-agent professional staff." As of September 1996 almost 30% of the lab's 204 examiners were former agents (Office of Inspector General, 1997, part two, section I.C). Risinger *et al*. say, "It appears that the bulk of forensic science examiners began their careers as law enforcement officers" (2002, p. 27, n. 126). This situation is beginning to change, however. In recent years there has been a move toward "civilianization" of forensics in the US (personal conversation with Lawrence Kobilinsky 19 August 2004).

The American Society of Crime Lab Directors/Laboratory Accreditation Board (ASCLD/LAB) accredits many forensic laboratories. It is not clear, however, how meaningful accreditation is. For example, as I note below, their accreditation standards contain no requirements for procedures to reduce the chance that analysts will be biased by extraneous information or suggestive presentation of evidence (Risinger *et al*., 2002, p. 31). Moreover, some labs, including the Houston Crime Lab (McVicker and Khanna, 14 March 2003), are not accredited. (The Houston lab has recently been given the charge to achieve accreditation by September 2005.) The FBI lab was accredited only recently.

Forensic workers tend to identify with the police. They tend, therefore, to seek out evidence supporting the police theory. The comment of "one lab veteran" at the FBI illustrates the point. "People say we're tainted for the prosecution. Hell, that's what we do! We get our evidence and present it for the prosecution" (Kelly and Wearne 1998, pp. 15-16). Kelly and Wearne quote John McDermott, "a senior FBI official," testifying in 1981 before a Congressional subcommittee. The ideal lab specialist, McDermott

explained, “stands in the shoes of the investigator in the field, whom he is serving” (p. 16).

Jack Dillon, then the FBI Firearms-Toolmark Unit chief, told Kelly and Wearne (1998, p. 16), “Sometimes they’re [investigators] pretty confused about what they want, so we’ll call them up to find out what they’re trying to prove. Often we can suggest some better ways of doing it.” As Kelly and Wearne note, this type of “close liaison” between law-enforcement officers and forensic workers creates bias.

A former Firearms-Toolmarks Unit chief at the FBI laboratory named Evan Hodge wrote an article on “Guarding Against Error” in which he relates a particularly revealing story. As Kelly and Wearne (1998, p. 17) retell it, a police inspector took

> a 1911A1-model .45- caliber pistol to a lab for confirmation that it was a murder weapon. “We know this guy shot the victim and this is the gun he used,” the examiner was told. “All we want you to do is confirm what we already know so we can get the scumbag off the street. We will wait. How quick can you do it?” The examiner gave them their instant identification. The suspect confessed and led the police to a second pistol, also a .45, also a 1911A1 model, which lab tests demonstrated was the real murder weapon. “We all do this (give in to investigative pressure) to one extent or another,” Evan Hodge admits, arguing that the only solution is to remove the sources of it from the laboratory completely.

In the current institutional regime, the forensic worker is typically an employee of a law-enforcement agent such as the municipal police or FBI. They are given the evidence in a suggestive manner and try to show that the police theory is true. Their analysis is not likely to be reviewed or questioned by other forensic workers or even most legal defense teams. They are in something of a monopoly position, therefore, with the monopoly franchise being offered by the police.

B. Cognitive Bias in Forensic Analysis

Forensic workers are subject to bias. Relying on the literature in social psychology, Risinger *et al*. provide a long list of cognitive biases that may affect forensic workers. They use the term "observer effects" to embrace the various biases they discuss. These include selective attention, anchoring effects, role effects, experimenter effects, conformity effects, confirmation bias, motivational bias, and bias by examination-irrelevant information or domain-irrelevant information.

Miller (1987) provides an excellent example. He asked a group of 14 students trained in hair analysis to examine four cases each. "The 14 students met the basic requirements for expert testimony on human hair identification in courts of law" (Miller 1987, p. 160). For each student, two cases were presented the usual way. They were given two samples and told that one was from the crime scene and the other from the suspect. The other two cases were presented through a forensic line-up. The known sample from the imaginary crime scene was compared to "five suspect-known hair samples" (Miller 1987, p. 160). In all 56 cases, there were no true matches. The first group of cases showed an error rate of 30.8%. The second group showed an error rate of 3.8%.

Miller's study illustrates Jonakait's overservation that evidence "often is presented to the forensic scientist in a needlessly suggestive manner" (p. 160). The samples are labeled as coming from the defendant or from the victim. The samples are "frequently accompanied by a synopsis of the investigation indicating the reasons that the

investigators believe the suspect is guilty" (160). This creates a bias by suggesting to the forensic worker what result is expected or correct.

Risinger *et al*. note, "If even the mildest of expectations can affect perception, then it is not surprising to find that where an observer has strong motivation to see something, perhaps a motivation springing from hope or anger, reinforced by role-defined desires, that something has an increased likelihood of being 'seen'" (p. 24).

McCabe *et al*. (2001) shows that social context influences cognition.[2] The institutional regime of police forensics influences the preferences of forensic workers and may bias their reasoning. The institutional context creates a kind of "motivated reasoning" among forensic workers. Lodge and Tabor (2000) present a simple model of motivated political reasoning. They list several factors that tend to produce biased judgment. Four of them apply to forensic workers in the current institutional environment, namely, the consequences of being wrong are weak, the judgmental task is complex, evidence is ambiguous, and one is under time pressure (p. 185).

One forensic scientist has told me in a personal conversation that the consequences of error are very high for forensic scientists, making one of Lodge and Tabor's factors inapplicable. An analyst, he has explained, whose work is found to be substandard acquires a damaging reputation that follows him or her. I disagree. The probability of being discovered in an error is relatively low and disciplinary measures are often weak when an error is detected. We have several known cases of substandard work going undetected for years and several cases in which a discredited analyst has suffered little or no adverse career consequences. Recently, for example, Youngstown State

[2] This role of social context is the subject of the long-standing discipline of social psychology. The study of McCabe *et al*. is noteworthy, however, for its link to neuroscience and its use of MRI brain scans.

University, which is located near Cleveland Ohio, hired Joseph Serowik to head its forensic science program. Serowik had been suspended from his job as a lab technician in the Cleveland crime lab after his work had been discredited. His erroneous analysis led to the false conviction of Michael Green for rape. Moreover, "questions about Serowik's competence have been raised in law enforcement circles for more than a decade." Serowik was recommended for the academic post by "Cuyahoga County Common Pleas Judge Timothy McGinty, the former prosecutor who sent Michael Green to prison for rape in 1988 with the help of the now-discredited testimony from Serowik" (Gillispie and Mills 2004).

In addition to the relatively subtle factors listed by Lodge and Tabor, there is a further and more profound reason for motivated bias in forensics. In some cases, the police employ the forensic worker and review his job performance. This police authority creates a strong economic motive to satisfy the police by confirming the police theory. This motive competes with others, such as a desire to see justice done. It is often present, however, and provides another source of conscious and unconscious bias in forensic analysis.

Group-serving bias is another probable source of bias in forensic work.[3] Group-serving bias is created when a person considers himself a member of a "coalitional alliance" (Kurzban *et al*. 2001). A coalitional alliance is characterized by coordinated action toward a common goal.

The psychological roots of group-serving bias were set when humans were evolving from other apes during the Pleistocene epoch. (See Barkow *et al*. 1992 for an

[3] I thank Rob Kurzban for suggesting the term "group-serving bias." I am not aware of any prior uses of it.

explanation of the evolutionary shaping of human psychology.) Psychological identification with the coalition may have reduced the probability of defection and thus increased the individual's long-term value as a coalitional partner. An individual's psychological identification with the group would discourage shirking. It would, therefore, increase the value of the individual's contributions to the coalition, causing an increase in the utility of forming a coalition in the first place. Police and forensic workers are engaged in coordinated action toward a common goal and thus seem to be in a coalitional alliance in the evolutionary sense. The forensic worker and the police are on the same team. They are "us" and suspects are "them." This deep-seated bias in forensic analysis is inconsistent with objective scientific analysis.

Risinger *et al*. paint a vivid portrait of the daily operation of cognitive bias in forensic analyis.

> In light of this, consider the forensic scientist who takes poor notes during an examination and prepares a skimpy report, but then goes back to "spruce them up" shortly before trial. Even assuming the most honest of intentions, that examiner is inviting errors to infiltrate his conclusions and his testimony. The error potential of the original skimpy report, which leaves much to be supplied from memory, facilitates the creation of testimony more consistent with assumptions and later acquired expectations than would be the case with a more detailed and complete contemporaneous account. Reconstructive errors are given room to manifest themselves during the "spruce-up" stage (pp. 16-17).

The OIG report on the FBI crime lab provides an example that fits this portrait perfectly.

> The court asked Rudolph why the diphenylamine test and other tests he described were not documented in his notes. Rudolph responded, When I examine a case I put in my notes things that are important to me when I . . . give testimony. I don't write my notes for the United States Attorney. I don't write my notes for the defense. I write my notes for myself. Rudolph said he had done thousands of tests since 1982 and could not

> possibly remember them all. The court asked, Isn't that one of the reasons you keep notes? (Office of the Inspector General, Part three, section A, subsection II.A).

Risinger *et al*. (2002, pp. 25-26) observe that the cognitive biases they discuss can introduce errors at every stage of analysis. Thus, they conclude, cognitive biases may create:

> Errors of apprehending (errors that occur at the stage of initial perception);
>
> Errors of Recording (errors that creep in at the stage where what is observed is recorded, assuming a record beyond memory is even made);
>
> Errors of Memory (errors that are induced by both desires and the need for schematic consistency, and that escalate over time when memory is relied on);
>
> Errors of Computation (errors that occur when correct observations accurately recorded or remembered are transformed into incorrect results when calculations are performed on them); and
>
> Errors of Interpretation (errors that occur when examiners draw incorrect conclusions from the data).

Only a structural change in the organization of forensic work is likely to greatly reduce cognitive bias in forensic work.

C. Choice of Technique in Forensic Work

Forensic workers have a choice of techniques. They may choose, for example, which of several serological tests to use in matching a suspect's blood to a sample. Jonakait reports that there are no protocols for most forensic procedures (1991, pp. 157-158). Since he wrote, accreditation has somewhat mitigated this problem. Accredited labs in

the US must have protocols, although protocols may vary from lab to lab. But as we have seen, not all labs are accredited. Without protocols, forensic workers have considerable freedom to choose their techniques of analysis.

One author notes, “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” (Pearsall 1989, p 674).

Selective retesting is equivalent to choice of technique. Risinger *et al*. (2002) report on cases in which a forensic worker is asked to re-examine evidence after the expected result failed to appear (pp. 40-42). Selective re-examination introduces bias.

Feigenbaum and Levy (1996) show that choice of technique and selective reporting introduces bias to scientific analysis. The scientist may apply several techniques to a problem and publicly report only those tending to support his preferred theory. Their analysis applies to forensic work as well. The title of their paper, “The Technical Obsolescence of Scientific Fraud,” reveals the danger in leaving forensic workers free to choose. Freedom of choice increases the chances that the worker will be able to produce the result he wants by the use of techniques that are, considered in isolation, perfectly objective and legitimate. He has no need to engage in willful fraud; fraud is obsolete. He has only to apply several tests and report on those that point in the desired direction.

This problem would be serious if only dishonest workers used choice of technique and selective reporting to produce biased results. Unfortunately, however, even honest workers may do the same thing. Scrupulously honest workers may systematically reject unexpected or undesired results and accept expected and desired results. The honest, but unconsciously biased forensic worker will readily seize excuses to cast doubt on tests producing undesired results. He will search for reasons to dismiss doubts about tests producing desired results. The techniques of the sincere and conscientious worker can be almost as biased as those of the unscrupulous cheater.

Risinger *et al*. separate fraud from bias. "We are not concerned here with the examiner who, in light of the other findings, deliberately alters her own opinion to achieve a false consistency. That is the perpetration of an intentional fraud on the justice system, and there are appropriate ways with which such falsification should be dealt (2002, p. 38)." But the line between "honest error" and willful fraud is fluid. On the one hand, outright fraud is technologically obsolete in some circumstances. On the other hand, there are no bright lines as we move from the psychological state of disinterested objectivity to unconscious bias to willful fraud.

IV. Proposals to Improve Police Forensics

The police forensic worker is an autonomous authority with power. Power should be divided and contested among forensic scientists, creating a system of checks and balances. Many observers have recognized that the power of forensic workers can lead to substandard forensics. But as far as I know, no one has suggested fixing the problem by

making that power divided and contested. Such a division of power would create the "checks and balances" called for by the *News-Leader* of Springfield, Missouri (News-Leader, 2004).

Some readers may question the need for any change in institutions. Good lawyering, one might argue, is the cure for bad forensics.[4] This argument overlooks a basic scarcity consideration: High-quality counsel is not a free good. Without constraints on their time or energy, skilled and intelligent lawyers could learn enough about the limits of forensics to persuade judges and juries in those cases in which the forensic evidence presented by the prosecution was deficient; no innocents would be jailed because of forensic error. Good lawyering is a scarce good, however. Most criminal defendants are indigent and must rely on public defenders, who generally lack adequate incentives to perform well (Schulhofer and Friedman 1993) and may also be less skilled than private-practice lawyers specializing in criminal cases.

Even a scientifically well informed defense lawyer may be ineffective. "You can't rely on your own cross-examination of the state's witnesses," according to Kim Campbell, an assistant state's attorney in Illinois. Commenting on a case in which a well-informed lawyer failed in his challenge of an unproved forensic technique, Campbell continued, "You have to have your own expert to say why this kind of science is unreliable. And there was nobody saying that at his trial" (McRoberts *et al*. 2004). Presumably, the difficulty is that even a skilled lawyer has no metaphorical white lab coat creating an aura of scientific authority. Uninformed and boundedly rational jurors and judges may be driven to rely on the scientific credentials of a speaker as a proxy for scientific validity of the speaker's argument.

[4] I thank an anonymous referee for alerting me to this argument.

A. Existing Proposals and Their Limits

The existing literature contains many proposals for improving police forensics. Jonakait (1991) calls for regulation, as does the Forensic Justice Project (http://www.forensicjustice.org). Jonakait's model of successful regulation is the Clinical Laboratory Improvement Act of 1988. The regulation Jonakait calls for would require "inspections, personnel standards, quality control, and external proficiency testing" (1991, p.181).[5] He notes, however, that a full regulatory regime of the sort he desires would be difficult or impossible to institute given that forensic labs are not private, profit-seeking firms (1991, p. 182). At a minimum, in Jonakait's view, forensic labs should be subject to mandatory proficiency testing (Jonakait, 1991, pp. 182-185). While I support proficiency testing, such testing does not create the specific incentives for error detection I discuss below.

Risinger *et al*. (2002) call for several measures, including blind testing and evidence line-ups. Miller (1987) suggested evidence line-ups in the case of hair analysis. These are worthy proposals.

The Ryan Commission Report (Illinois 2002) and Giannelli (1997) call for independence of forensic labs from law-enforcement agencies. The tendency of independence to reduce bias has been questioned in a minority opinion expressed in the Ryan report. "The reality is that no matter how 'independent' this separate state agency

[5] In most industries, "regulation" restricts the contracts private parties can make. I cannot hire a plumber to remove my gall bladder. The principal effect of the "regulation" of forensics would be to impose conditions on persons hired by the government to examine its evidence or to restrict the admissibility in court of certain types of evidence produced through private contracting. Thus, the question of whether such regulation is consistent with a "free market" does not arise.

is, the bulk of its work will still be for police agencies and prosecutors" (Illinois, p. 53). The value of independence depends on other simultaneous factors, such as how forensic labors are divided and whether labs are subject to competitive pressure. In the least favorable conditions, the minority opinion in the Ryan report is probably right. But in more favorable conditions, independence may reduce bias.

Saks *et al*. (2001) propose a state-wide "Forensic Science Service," which would "provide forensic science services to police, prosecutors, defense counsel, judges, and pro se defendants concerned with criminal cases" (Saks *et al*. p.698). With this measure, Saks *et al*. hope "provides forensic science expertise to both the prosecution and the defense on equal terms" (p. 699). The "Commission on Forensic Science Services" would supervise the "Forensic Science Service." They do not say who should guard this guardian. (Juvenal asked, "Who will guard the guardians themselves?") The proposals of Saks *et al*. contain many valuable suggestions. They come well short of competitive self regulation, however. They propose "oversight" (p. 701) of forensics and say their Commission on Forensic Science Services should "strive" to staff laboratories under their supervision with workers who are properly trained and "committed to doing good science." I do not share the faith of Saks *et al*. in the powers of oversight and of command and control.

Thomson (1974) calls for a suite of reforms that is in several ways similar to the set of changes I will propose. Some significant differences exist, however. Thomson calls for 1) "consolidation of forensic facilities," 2) placing forensic labs under the supervision of the courts, 3) accreditation, 4) instituting a mandatory regime of

proficiency testing,[6] and 5) "provision for open access . . . for all parties in a criminal action" (p. 516). Thomson's fifth proposal is similar to my call for forensic vouchers. Thomson even calls for the use of separate facilities by the antagonists in a criminal process (p. 514). He seems to place less emphasis than me, however, on direct competition among labs.

All of the proposals just discussed contain useful ideas. But none of them, with the partial exception of those of Thomson, would adequately address the fact that forensic workers are in a kind of monopoly position with respect to the evidence given to them to analyze. As long as such a monopoly is enjoyed, the forensic worker has an incentive to shirk and to act on any biases he may have. To render power divided and contested, it is necessary to establish competition among forensic workers. Competitive self regulation would not, of course, magically cure all forensic ills. It would, however, induce significant improvements in the quality of forensic work.

B. Competitive self regulation

rivalrous redundancy

There should be several competing forensic labs in any jurisdiction. Subject to the constraints of feasibility, some evidence should be chosen at random for duplicate testing at other labs. The same DNA evidence, for example, might be sent to more than one lab for analysis. The forensic worker need not know whether the evidence is

[6] From the context, I infer that he is imagining a mandatory system.

examined by another lab. He will know that there could be another lab, and sometimes is.

This ***strategic redundancy*** gives each lab an incentive to find the truth and apply rigorous scientific standards. Strategic redundancy should be accompanied by ***statistical review***. For example, if a given lab produces an unusually large number of inconclusive findings, its procedures and practices should be examined. Competitive self regulation creates checks and balances.

It is surprising that the principle of redundancy has not been extensively applied to police forensics. You need redundancy to avoid errors. You get a second opinion when you are sick. Failsafe measures are built into power plants. You keep a spare tire in the trunk. But we have only rather limited forms of redundancy in forensics, such as the "verifications" that may go on within a crime lab.

Strategic redundancy works best if errors and biases are not correlated across labs. If all labs share the same biases, then strategic redundancy is less able to root out error and bias. Indeed, if competing labs all share the same strong bias, then strategic redundancy may make things worse by increasing the seeming legitimacy of what are, in fact, bogus results.[7] It is necessary to create incentives for the discovery of error. The stronger such incentives are the more they will mitigate or overwhelm any biases. Without such incentives we have ***mere redundancy***. When such incentives are in place, however, we have ***rivalrous redundancy***. An example illustrates the point. The example is not realistic, but it illustrates the difference between mere redundancy and rivalrous redundancy.

[7] Several people have pointed out to me that multiple labs may have the same bias, or that we should have specific incentives to error discovery. The list includes Alan Fask, David Friedman, Randy Holcombe, Nelson Pole, and David Porter.

Imagine we have competing forensic labs that are biased in favor of the police theory. Assume we have mere redundancy among forensic labs. If the lab performs a test it receives a money payment. If the police theory is supported, the lab gets a psychic benefit as well. Assume, finally, that the police theory is false in the case under review. In this case, each lab has an incentive to support the police theory. A given lab in this situation may find excuses to construe the evidence in ways tending to incriminate the police suspect or even to simply lie for the police. If the other lab exonerates the suspect, the given lab still has its money payment. But if the other lab also supports the police theory, the given lab enjoys an additional psychic benefit. In the language of game theory, supporting the police theory is a dominant strategy.

Now imagine a special case of rivalrous redundancy. If the labs disagree, there is an infallible adjudication procedure to determine who is right. (Real-world adjudication is, of course, fallible.) The lab that told the truth will collect two money payments, one for performing the test, and one for discovering the other lab's error. The erroneous lab gets nothing. This situation creates an incentive to perform a careful and objective analysis. Each lab would prefer the other to play along by supporting the police theory. On the other hand, each lab always has an incentive to be truthful, either to avoid forfeiting its payment or to get a double payment if the other lab provides a false analysis. In the language of game theory, they are playing a prisoners' dilemma.[8] Mere redundancy will not produce a truth-seeking system, but rivalrous redundancy will. This is a particularly likely outcome under a regime of information hiding.

[8] In a finitely repeated prisoners' dilemma, defection on every round is the only Nash equilibrium. In an infinitely repeated prisoners' dilemma tit-for-tat is a Nash equilibrium. Tit-for-tat is only one of an infinite number of other Nash equilibria, however, including defecting on each round. Overall, it would seem that cooperative outcomes are unlikely in this situation.

Here opens a field of research. What institutional structures induce mere redundancy and what structures induce rivalrous redundancy? How do we successfully adjudicate conflicting claims when all adjudicators are fallible? And so on. The general principle, however, seems to be rather straightforward. It seems perfectly possible to create monetary penalties for deficient laboratories and to create, thereby, a reasonable real-world version of rivalrous redundancy.[9]

Competitive self regulation would create conditions of forensic science similar to the conditions of pure science. In pure science, research results are subject to the discipline of review and reproduction. I propose subjecting forensic scientists to the same discipline of review and reproduction.

Appendix I discusses the case of Josiah Sutton, who was wrongly convicted in Texas on DNA evidence. Sutton's case shows that scientific advances such as new DNA technology will not solve all problems. New techniques will not solve the problem that forensic scientists do not operate in the sort of environment that encourages good science. They face the wrong set of incentives and pressures. New technologies or scientific advances will not solve this problem. The problem and its solution are not a matter of lab science, but of social science. Competitive self regulation puts forensic workers in the right environment to do the right thing.

[9] It is possible that such penalties should be accompanied by information hiding to ensure truthful outcomes. Information hiding makes it difficult or impossible for labs to guess what outcome the police desire; monetary penalties create a specific incentive for error discovery. Whether rivalry can be induced without information hiding is something of a moot point, however, since information hiding is idependently desirable.

evidence control officer and information hiding.

To implement competitive self regulation would require the creation, within each jurisdiction, of an ***Evidence Control Officer***. Such a position would not be entirely novel. Currently, the FBI lab has an Evidence Control Center for internal purposes. Indeed, in most or all jurisdictions an evidence control office exists, sometimes under a different title. But these positions do not serve the functions that I propose be served by a jurisdiction's Evidence Control Officer. Risinger *et al*. also propose creating the position of "Evidence and Quality Control Officer," although they do not propose competitive self regulation. The Evidence and Quality Control Officer, they explain,

> would be responsible not only for coordinating work among examiners in different specialties, but also for being the sole contact point between the entity requesting the test [prosecution or defense] and the laboratory. She would also serve as the filter between each examiner and any information about the case, whether it originated from without or from within the lab. She would decide not only generally what kinds of tests were needed, but what information about the case was needed to perform those tests, and her primary duty would be to maintain appropriate masking between the examiners and all sources of domain-irrelevant information (Risinger *et al*. pp. 46-47).

Risinger *et al*. rightly emphasize the duty of the Evidence Control Officer to engage in information hiding.

In addition to the functions identified by Risinger *et al*., the Evidence Control Officer should use random-number generators to decide which lab gets a given piece of evidence and when to send the same evidence to more than one lab.

The Evidence Control Officer may seem to be in a position to commit the same sorts of conscious and unconscious abuses that many forensic workers have committed.

The evidence control office may look every bit as monopolistic as the forensic worker in the current system.[10] Several considerations suggest, however, that it is easy to structure the job of Evidence Control Office so that the position involves a low incentive to cheat, high costs to being discovered cheating, and a high probability of being caught if cheating is attempted.

First, many of the functions of the evidence control office (in my proposed system) are relatively mechanical. If these functions are witnessed or reviewed publicly, then they are less likely to be improperly carried out. Sandy Ikeda suggested to me, for example, that it may be possible to use lotto numbers for the random numbers that determine who gets what evidence.

Second, it may be desirable to divide the labors of the Evidence Control Officer. For example, the evidence control office might randomly assign one forensic lab to prepare the evidence for a case and a second lab to analyze the evidence prepared by the first lab. Alternatively, evidence preparation might be randomly assigned to volunteers and part-time workers who work as scientists outside the criminal justice system. Even highly educated scientists would, however, have to be specially trained in the tasks of evidence preparation.

Third, the "masking" function of an Evidence Control Officer (information hiding) will be easier to maintain if the Officer is independent of the forensic labs to which he sends evidence. If there is only one lab in the jurisdiction, it becomes more likely that independence will be lost. If there is only one lab in a jurisdiction, the Evidence Control Officer may acquire the psychological attitude of identification with

[10] Thomas McQuade and Richard Adelstein have pointed out to me the risk that an evidence control office may abuse his authority.

the lab. The Officer is subject to the feeling that he is in a coalitional alliance with the police. He may convey this feeling, consciously or unconsciously, to the monopoly lab with which he deals. All of this is less likely to occur if there are many competing labs in a jurisdiction. Note that a jurisdiction may send work to labs that are geographically distant and that a lab may serve several jurisdictions. In the face of competition among labs, the Evidence Control Officer has an incentive to adopt an above-the-fray attitude that helps maintain objectivity and discourage cheating. Moreover, if the Officer should exhibit bias or share inappropriate information, the fact is more likely to be revealed if there are several labs observing the problem. Thus, strategic redundancy is a palliative limiting abuse in the function of the Evidence Control Officer.

The Evidence Control Officer it separated by one step from the tests to be performed on evidence. Thus, if he (or she) wishes to cheat, he would have to coordinate other actors. He would have to create a conspiracy to produce bias. Coordinating a conspiracy is costly; because the conspiracy may fall apart, it is also dangerous. The risk of detection becomes an *ex ante* cost.

Fourth, it seems possible to require the Evidence Control Officer to occasionally send bogus, bias-inducing samples to the labs in its jurisdiction. The lab would be under an obligation to report improper evidence preparation or information sharing. Failure to do so would meet with sanctions. In this context, the Evidence Control Officer would have to fear that any attempt to improperly influence a lab would be discovered.

Finally, it is easy to impose heavy sanctions on Evidence Control Officers who are caught in acts malfeasance.

From the forgoing considerations, it seems easy to ensure that an Evidence Control Officer will not be very willing or able to purposefully subvert the intended functions of his office.

statistical review

The use of multiple labs is strategic redundancy. (When division of the evidence is impossible, of course, only one lab would receive evidence.) With strategic redundancy, a forensic worker must wonder who else is examining the same evidence. The worker's reputation and job prospects will suffer if he or she is found to have provided a false or sloppy analysis. The prospect of such embarrassment gives the worker an incentive to provide a professional, scientific, and objective evaluation of the evidence. The forensic worker has an incentive to do the job right in the lab and to interpret the evidence judiciously in his or her report to the court. When the results of two labs are inconsistent, an adjudication procedure is required to resolve the inconsistency.

Statistical review is the follow up to strategic redundancy. The competing forensic labs in a given jurisdiction should be subject to periodic statistical review. This review consists principally in counting the number of cases falling into various categories. In how many cases was a lab's findings found to be deficient when compared to the contradictory results of competing labs? How many cases led to conviction? How many to exoneration? In how many cases did a lab find the evidence to be inconclusive?

And so on. If a lab is found to have an unusually high or low number of cases in any category, it should be investigated to learn why.

It might seem that there is no reason to look at the number of convictions. The question is how the lab does its work, not who goes to jail. But if the analyses of a given lab correspond to an anomalous number of convictions (whether large or small), then we have reason to inquire if there has been a breach in the wall of separation between the forensics lab and the prosecution or defense.

division of labor with vouchers

In the current system of criminal justice in the US, forensic workers typically conceive of themselves as working for the police or prosecution. As we have seen, this introduces bias. This bias is combined with rules of discovery that make it hard for defense attorneys to challenge the supposed results of forensic tests. The consequence is that the sloppiest work may easily satisfy a jury, who cannot be expected to know about the difficulties of practical forensic science today.

The task of interpreting forensic tests should be divided from the task of performing those tests. Just as the indigent persons on trial are provided an attorney at the state's expense, so too should indigent persons on trial be provided a forensic interpreter at state's expense. Schulhofer and Friedman (1993) argue that public defenders have incentives to go along with the police and prosecutors and thus too easily give up the fight. They propose a voucher system, which would give public defenders an incentive to act in the interests of their clients. I propose that indigent defendants be

given forensic vouchers as well. Saks *et al*. (2001) propose something similar with their Forensic Science Service. Their proposal does not produce rivalrous redundancy and does not does not give the indigent the consumer's power to choose among suppliers of forensic counsel; it is therefore likely to be less effective in removing incentives to biased and sloppy work.

Dividing test from interpretation and providing separate forensic interpreters for both sides would bring forensic evidence into the adversarial system of the courts. The common law system is based on the idea that the truth comes out best in an adversarial process. But, as we have seen, forensic evidence is largely excluded from the adversarial process. This exclusion from the adversarial system is a profound, needless, and inappropriate compromise of one of the most fundamental principles of our the common law system.

Separating out the task of interpretation could also be combined with the creation of standardized reports such that every expert having an ordinary knowledge in the field would be able to reproduce the test and interpret the results.[11] Standardized reports would tend to reduce the unfortunate element of idiosyncrasy that still characterizes much forensic work.

Paul Rubin has asked me whether my proposals might create perverse consequences through feedbacks of the sort Stuntz (1997) identifies. In particular, legislators might reduce funding of defense attorneys to pay for forensic vouchers. I think this result is unlikely for several reasons. As I argue below, my proposals are likely to reduce the costs of police forensics. If anything, this result would tend to increase the spending on defense lawyers. Stuntz argues that legislators may reduce funding of

[11] I thank an anonymous referee, whose very language I have borrowed, for this suggestion.

defense attorneys to "get tough on crime." My proposals would reduce this incentive by increasing the ability of the system to distinguish the guilty from the innocent. Indeed, improved forensics would tend to break the vicious circle Stuntz (1997, pp. 55-56) identifies. Court mandated procedure make procedural arguments more attractive at the margin than material arguments, producing more acquittals on technicalities. Such acquittals induce reduced funding to defense spending, as well as increases in mandatory sentencing and in the number of crimes defined. Improved forensics would reduce the relative price of arguing the facts. Finally, it should be noted that Stuntz does not provide a clear mechanism establishing the links he claims to exist between legislation and the results of criminal procedure. In other words, as Stuntz admits (p. 5), his argument is speculative.

privatization

Finally, competing forensic labs – and interpreters – should be private, profit-making enterprises. Privatization would probably provide cost saving in forensics, just as it has in other areas of the criminal justice system (Benson 1998, pp. 28-34). For example, Oro Valley, Arizona contracted out its police services in 1975 and achieved simultaneous reductions in cost and crime. The arrangement was cut short by a legal challenge (Benson 1998, pp. 28-29). Paraphrasing Benson (1998, p. 1), privatization would have the advantage of turning the entrepreneurial discovery process loose in the forensic laboratories of the criminal justice system.

There is by now a large literature on privatization. (See Megginson and Netter 2001 for a survey.) The general thrust of this literature would tend to support the idea of privatization. As Megginson and Netter note, "privatization tends to have the greatest positive impact . . . in competitive markets or markets that can readily become competitive" (2001, p. 329). Forensics is such an industry.

The current situation is almost the reverse of a natural-monopoly. Currently, a forensics lab's scale of operation is dependent on the size of the jurisdiction it serves. It is thus unable to exploit economies of scale. Under privatization, the same lab may serve many jurisdictions and thus enjoy economies of scale.

Citing Sappington and Stiglitz (1987), Megginson and Netter note that it is generally easier for governments to intervene in state-owned enterprises than private enterprises (2001, p. 330). This is an advantage for privatization if government intervention carries more costs than benefits. Government promises not to intervene in certain ways, for example, may be less credible with respect to state-owned enterprises. Ironically, privatization would improve the ability of national governments to intervene in the operation of forensics labs, as least in the US. Forensic labs are currently under the jurisdiction of local governments, which may adopt policies different from those the national government might choose. Privatization would open the way for national regulations. Privatization would reduce the cost of national regulation and, therefore, of intervention at the national level. Interventions that impose national standards and protocols would be easier under privatization. If interventions at the local level are undesirable in the forensics industry, whereas national regulations are desirable, then

privatization would help create the right set of government regulations of forensic practice.

Some of the incentives to good work are much stronger for private labs than for government labs. A private lab is subject to civil action for false or sloppy work. Private labs can also be made subject to regulation and to administrative fines for bad work. Such fines will be felt directly by a responsible party, namely the firm owner, who is in a position to act to correct problems.

Competitive self regulation will produce cost savings and better results if the competing labs are private, profit making enterprises. The Oro Valley case just discussed illustrates the claim. The incentive to reduce costs is clear. Privatization creates a residual claimant who gains from cost reduction. Quality improvements are likely as well if the privatized enterprise is subject to competition. Benson (1998, p. 35) points to three factors tending encouraging improved quality in privatized enterprises. First, "effective employee monitoring and the development of new technology can simultaneously lower costs and enhance quality." Second, private firms have a reputation to maintain. "A firm with a reputation for providing lower quality than expected may not be in business for very long, if competitive alternatives are available." Third, "in a competitive market the quality of the outcome depends more on demand side incentives than supply side incentives." In other words, if demanders insist on a high-quality product, the market will provide just that. The presence of forensic council for both defense and prosecution gives the authorities an incentive to demand high-quality forensic analysis.[12]

[12] In the current system in the US, the police and prosecution often want not good work, but that their theories be confirmed. The presence of forensic counsel for the defense creates in the prosecution the need

Private labs could easily engage in price competition. Every year a given jurisdiction might sign a one-year contract with several competing labs. The fees for each lab for the ensuing year would be determined annually. This price competition would, of course, tend to produce cost savings for tax payers.

Privatization has at least one more advantage. In the current system, the police in a given jurisdiction have monopsony power in the forensics market. Such power may give them the ability to exercise inappropriate influence on the lab or labs in their jurisdiction. With privatization, some labs could serve several police jurisdictions, including some at long distances. This is not entirely new. The FBI in the US and the FSS in the UK serve many jurisdictions. Privatization would tend to reduce or eliminate the monopsony power of the police.[13]

Poorly designed "privatization" may replace a government bureaucracy with a profit-seeking monopoly (Williamson 1976). This type of privatization should be avoided in forensic reform. If, however, privatization of police forensics is combined with rivalrous redundancy, statistical review, the creation of an Evidence Control Office who hides information, and division of labor with vouchers, then it has considerable potential to raise standards and lower costs.

competitive self regulation

The current system tends to induce biased and sloppy work even among those who enter the system as sincere and conscientious forensic workers. Competitive self regulation

---

for good work that can withstand the hostile scrutiny of the defendant's forensic counsel.

[13] Sandy Ikeda drew my attention to the monopsony aspect of the current system.

would create a system that produces unbiased and careful work even if individual forensic workers are disposed toward bias and negligence. Competitive self regulation would have three effects. First, it would reduce the bias. For example, the use of an "Evidence Control Officer" would reduce presentation bias. Second, it would mitigate the ill effects of remaining biases. For example, rivalrous redundancy increases the chances that a false and biased analysis will be scrutinized and overturned. Finally, it would create checks and balances that would use countervailing biases to neutralize one another. For example, the use of forensic vouchers would counter one bias with another, thereby increasing the chance that all relevant forensic arguments will be presented to a judge or jury.

Competitive self regulation is a principle for the reorganization of forensic laboratories. It is a plan for the institutional structure of forensic science. The plan enables forensic workers to do their work properly. Competitive self regulation uses strategic and rivalrous redundancy, statistical review, the creation of an Evidence Control Officer, division of labor with vouchers, and privatization to induce a competition for excellence among forensic labs.

V. What about Costs?[14]

Competitive self regulation may seem wasteful. Redundancy is costly. A complete answer to this question requires careful empirical work, which is premature at this stage of analysis. Nevertheless, I can provide a rough estimate. Competitive self regulation

[14] Many people have put this question to me, including Mark Campbell, Mark Casson, Bruce Caldwell, Young Back Choi, David Colander, Gloria Gadsden, Randy Holcombe, Chidem Kurdas, Meike Niedbal, and Jonas Prager.

would add less than $300.00 to the costs incurred by the criminal justice system in each investigation and possible trial involving forensics analysis. This estimate is based on annual laboratory budgets and includes, therefore, the expected value of time spent in court as an expert witness. Appendix II explains how I calculated this value and why it is probably too high. The Bureau of Labor Statistics reports that the "Average hourly and weekly earnings of production or nonsupervisory workers on private nonfarm payrolls" in February 2003 was $15.34. At this value of time, the extra forensic analysis required by competitive self regulation would correspond to less than 20 working hours, or the opportunity cost of a day in jail for the average worker. The exaggerated sum of $300.00 is a small fraction of trial costs for the cases that go to trial. A small improvement in the quality of forensic analysis would induce compensating reductions in the social cost of the further crimes of guilty persons not convicted and of the loss of social output from innocent persons wrongly convicted. I believe it is fair to conclude that competitive self regulation is cost effective. Other considerations strengthen this conjecture.

The use of fees will help to reduce the costs of forensics. Saks *et al*. (2001) propose that "fees [be] charged to parties requesting tests and examinations" and that the "schedule of fees shall apply equally to all parties requesting tests and examinations" (p. 703). Privatization would produce this result.[15] Right now, the marginal cost of a forensic test is often zero for the party requesting the test. The government has a third-party payer, the taxpayer. Thus, it is likely that needlessly wasteful tests are being conducted today. Saks *et al*. say, "Because the tests are not without cost to the parties, the requesters will be more thoughtful about the knowledge expected to be obtained from

[15] It is possible that private forensic labs might charge annual subscription fee, thereby reducing the marginal cost of a test to zero. It would be straightforward to prohibit such annual fees.

the costs associated with testing" (p. 703). Thus, the overall result of competitive self regulation might well be a reduction in the costs of forensic testing.

Further cost savings would be realized if the price competition discussed earlier is permitted. While each lab would charge the same fees to all parties, allowable fees would be determined each year by a competitive process.

We already have many forensic laboratories in the U.S. Shipping costs are low. A political jurisdiction may easily rely on geographically distant labs. Indeed, the FBI lab in Washington, D.C. does forensic work for local jurisdictions across the U.S. Thus, competitive self regulation would require little or no additional overhead. Improved forensics would produce fewer costly appeals.[16] The modest increases in the average cost of an individual trial would be more than compensated by a reduction in total number of proceedings.

Finally, we have no adequate measure of the costs of forensic mistakes today. A forensic mistake can put the wrong person in jail. When that happens, we may have one innocent person removed from a productive role in society and another guilty person left free to commit crimes. Each such failure of forensics has a high social cost. It may be that a very small increase in the reliability of police forensics will produce a very large decrease in the social cost of forensic mistakes. Unfortunately, we have no measures of the costs of forensic mistakes in the current system. Given our ignorance in this area, it would be a mistake to dismiss competitive self regulation as "costly" when we have no reliable measure of the costs of the mistakes produced under the current system.

In addition to the social costs of false convictions and false exonerations, poor forensics can be expensive for a city, county, or state that must out-source its forensics

[16] I thank John Schiemann for this point.

and face civil action. DNA work in Harris County, Texas is being sent to private labs in the wake of Houston's crime-lab scandal. In the somewhat similar case in Cleveland involving Joseph Serowik, the city faced a $10 million law suit. The plaintiff settled for $1.6 million, but only as part of an agreement that created a "special master" with extensive powers and discretion to review questionable work in Cleveland's crime lab (Schultz 2004). (A PDF file containing the full text of the agreement between Green and Cleveland is available from the author on request.)

VI. Conclusion

Our knowledge of forensic techniques is running ahead of our knowledge of how to manage and organize forensic labs and workers. This knowledge gap is contributing to sloppy and biased forensics. I have identified a broad set of principles whose skillful application would improve forensic practice.

Competitive self regulation combines rivalrous redundancy with statistical review, information hiding, a division of labor between analysis and interpretation with forensic vouchers for the accused, and "privatization" of forensic laboratories. Under competitive self regulation, each jurisdiction would have several competing forensic labs. Evidence would be divided and sent to one, two, or three separate labs. Chance would determine which labs and how many would receive evidence to analyze.

The strategic redundancy embodied in this proposal is similar to the purposeful redundancies of coding theory in computer programming, failsafe systems in power

plants, and the spare tire in the trunk of one's car. We have redundancy in many systems, but not in forensics. It is high time we begin to study how to apply the salutary principle of redundancy to forensic science.

A large field of inquiry opens before us. The best general term might be "forensic science administration." We need to find sound principles of forensic science administration and learn how to apply them. Some principles are well established, for example evidence lineups and proficiency testing. I have proposed a new set of principles under the label "competitive self regulation." We know very little about how to apply these principles. The field of forensic science administration is largely unexplored. I hope this paper will induce others to contribute to this new field of inquiry.

References

Anez, Bob. 26 August 2004. "Montana High Court Petitioned for Sweeping Crime Lab Inquiry," an Associated Press story found in the *Seattle Post-Intelligencer*.

Barkow, Jerome H., Leda Cosmides, and John Tooby, edited. 1992. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, New York and Oxford: Oxford University Press.

Benson, Bruce L. 1998. *To Serve and Protect: Privatization and Community in Criminal Justice*, New York and London: New York University Press.

Bernstein, David E. 1996. "Junk Science in the United States and the Commonwealth," 21: *Yale Journal of International Law*. 123-182.

Blake, Edward T., Pamela Newall, George Sensabaugh, Robert Shaler, Ronald L. Singer, and Mark D. Stolorow. 2004. "Peer Review Report Texas v. George Rodriguez." Available from the author (Koppl) on request.

Bowers, Michael C. 2002. "Identification from Bitemarks: Proficiency Testing of Board Certified Odontologists," in Faigman, David L. *et* al. *Modern Scientific Evidence: The Law and Science of Expert Testimony*, 2002.

Brave, Ralph. 28 July 2004. "DNA To Go: Legislative Audit Reveals Troubling Problems with Maryland State Police's Handling of DNA Database," *Baltimore Citypaper.*

Bretz, Ronald. 1987. "Scientific Evidence and the Frye Rule: The Case for a Cautious Approach," *Thomas M. Cooley Law Review*, 4.

Browne, M. Neil, Carrie L. Williamson, and Linda L. Barkacs. 2002. "The Perspectival Nature of Expert Testimony in the United States, England, Korea, and France," *Connecticut Journal of International Law*, 18: 55-102.

Douglas, Jim. 24 September 2004. "Expert's Resignation Latest Setback for FWPD Crime Lab," WFAA, downloaded from http://www.wfaa.com 27 September 2004.

FBI Director. 2002. "An Audit of Houston Police Department Crime Laboratory-DNA/Serology Section," December 12-13, 2002.

Feigenbaum, Susan and David M. Levy. 1996. "The Technical Obsolescence of Scientific Fraud," *Rationality and Society*, 8: 261-76.

Friedman, Richard D. 2003. "Squeezing Daubert Out of the Picture," 33 *Seton Hall Law Review*, 33: 1047.

Giannelli, Paul C. 1997. "The Abuse of Evidence in Criminal Cases: The Need for Independent Crime Laboratories," *Virginia Journal of Social Policy & the Law*, 4: 439-478.

Gillispie, Mark and Lila J. Mills. 22 August 2004. "Suspended crime-lab technician lands a job," *The Plain Dealer*.

Greer, Steven. 1994. "Miscarriages of Criminal Justice Reconsidered," *The Modern Law Review*, 57(1): 58-74.

Grieve, David. 1996. "Possession of Truth," *Journal of Forensic Identification*, 46(5): 521-528.

Griffin, Lissa. 2000/2001. "The Correction of Wrongful Convictions: A Comparative Perspective," *American University International Law Review*, 16: 1241-1308.

Horne, Terry. 25 January 2004. "Crime Lab Boss Placed on Leave: Mayor Removes Longtime Director amid Allegations he Helped Cover Up Wrongdoing," *The Indianapolis Star*, as found on http://bioforensics.com/news/crime_lab_boss.html on 11 October 2004.

Illinois, State of. 2002. *Report of the Governor's Commission on Capital Punishment*, State of Illinois, April 15, 2002.

Jonakait, Randolph N. 1991. "Forensic Science: The Need for Regulation," *Harvard Journal of Law and Technology*, 4: 109- 191.

Kaufman, Fred. 1998. *Commission on Proceedings Involving Guy Paul Morin*, Queen's Printer for Ontario, 1998.

Kelly, John F. and Phillip Wearne. 1998. *Tainting Evidence: Inside the Scandals at the FBI Crime Lab*, New York: The Free Press.

Khanna, Roma. 8 February 2003. "DNA from conviction of teen will be retested: 1999 rape case part of HPD crime lab review," *The Houston Chronicle*.

Khanna, Roma. 10 September 2003. "Credentials embellished Transcript: Ex-lab chief told jury he had a Ph.D.," *The Houston Chronicle*.

Koppl, Roger and William Butos. 2003. "Science as a Spontaneous Order: An Essay in the Economics of Science," in Jensen, H. S., Vendeloe, M., and Richter,L., ed. *The Evolution of Scientific Knowledge*, Edward Elgar.

Kurzban, Robert, John Tooby, and Leda Cosmides. 2001. "Can race be erased? Coalitional computation and social categorization," *Proceedings of the National Academy of Science*, 98(26): 15387-15392.

Lodge, Milton and Charles Taber. 2000. "Three Steps toward a Theory of Motivated Political Reasoning," in Lupia, Arthur, Mathew D. McCubbins, and Samuel L. Popkin, editors, *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*, Cambridge: Cambridge University Press.

Lucas, Leete & Field. 1985. "An American Proficiency Testing Program," *Forensic Science International*, vol. 27.

Macfee, Michelle. 15 September 2004. "DNA-Hair Test Results should be Wake-Up Call: Lawyer," downloaded from http://cnews.canoe.ca on 22 September 2004.

McBeth, Jim. 21 September 2004. "Fingerprint row detective cleared by US forensics," *Scotsman.com*. Downloaded from http://scotsman.com on 27 September 2004.

McCabe, K., D. Houser, L. Ryan, V. Smith, and T. Trouard. 2001. "A functional imaging study of cooperation in two-person reciprocal exchange," *Proceedings of the National Academy of Sciences*, 98(20): 11832-11835.

McQuade, T.J. & W.N. Butos (2003). "Order-Dependent Knowledge and the Economics of Science" *Review of Austrian Economics* **16** 2/3: 133-152.

McQuillan, Peter J. 2004. "Forensic Contretemps," downloaded from http://www.innocenceproject.org/dnanews/index.php on 7 October 2004.

McRoberts, Flynn, Steve Mills, and Maurice Possley. 17 October 2004. "Forensics Under the Microscope: Unproven Techniques Sway Courts, Erode Justice," *Chicago Tribune*.

McVicker, Steve and Roma Khanna. 14 March 2003. "D.A. is asked to back call for HPD lab probe: Rosenthal cool to resolution seeking FBI review of cases,'" *The Houston Chronicle*.

Megginson, William L. and Jeffry M. Netter. 2001. "From State to Market: A Survey of Empirical Studies on Privatization," *Journal of Economic Literature*, 39(2): 321-389.

Merton, Robert K. 1957. "Science and the Social Order," in his *Social Theory and Social Structure*, revised and enlarged edition, New York: The Free Press.

Miller, Larry S. 1987. "Procedural Bias in Forensic Science Examinations of Human Hair," *Law and Human Behavior*, 11:157-163.

Mills, Steve, Flynn McRoberts, and Maurice Possley. 20 October 2004. "When Labs Falter, Defendants Pay: Bias Toward Prosecution Cited in Illinois Cases," *Chicago Tribune*.

Mills, Steve and Maurice Possley. 14 January 2001. "Report Alleges Crime Lab Fraud Scientist is Accused of Providing False Testimony," *Chicago Tribune*.

Moenssens, Andre A. 1993. "Novel Scientific Evidence in Criminal Cases: Some Words of Caution," *Journal of Crime Law and Criminology*, 84: 1-?

News-Leader. 19 August 2004. "Changes needed at state crime lab: How big? That should be a matter for debate," *News-Leader* (of Springfield, Missouri).

Office of the Inspector General, United States Department of Justice. 1997. *The FBI Laboratory: An Investigation into Laboratory Practices and Alleged Misconduct in Explosives-Related and Other Cases*, http://www.usdoj.gov/oig/special/97-04a/index.htm.

Peterson, Joseph L., D. Crim and Penelope N. Markham. 1995a. "Crime Lab Proficiency Testing Results, 1978-1991, I: Identification and Classification of Physical Evidence," *Journal of Forensic Sciences*, 40(6): 994-1008.

Peterson, Joseph L., D. Crim and Penelope N. Markham. 1995b. "Crime Lab Proficiency Testing Results, 1978-1991, II: Resolving Questions of Common Origin," *Journal of Forensic Sciences*, 40(6): 1009-1029.

Pearsall, Anthony. 1989. "DNA Printing: The Unexamined 'Witness' in Criminal Trials," *California Law Review*, 77: 665-703.

Polanyi, Michael. 1962. "The Republic of Science: Its Political and Economic Theory," *Minerva* 1: 54-73.

Puit, Glenn. 19 April 2002. "Police Forensics: DNA Mix-Up Prompts Audit at Lab," *Las Vegas Review-Journal*.

Raziq, David and Anna Werner. 2004. "DNA Testing: Study Calls into Question Long-Trusted Lab Results," *The IRE Journal*, 27(1): 13-14, 30.

Risinger, Michael, Michael J. Saks, William C. Thompson, and Robert Rosenthal. 2002. "The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion," *California Law Review* 90: 1-56.

Saks, Michael J. 2003. "Reliability Standards--Too High, Too Low, or Just Right?: The Legal and Scientific Evaluation of Forensic Science (Especially Fingerprint Expert Testimony)," 33 *Seton Hall Law Review*, 49: 1167-1187.

Saks, Michael J. 1998. “Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science,” 49 *Hastings Law Journal*, 49: 1069 -1141.

Saks, Michael J. *et al*. 2001. “Model Prevention and Remedy of Erroneous Convictions Act,” *Arizona State Law Journal*, 33: 665-718.

Salzman, Jonathon. 24 January 2004. “Man Freed in 1997 Shooting of Officer,” *The Boston Globe*.

Schulhofer, Stephen J. and David D. Friedman. 1993. “Rethinking Indigent Defense: Promoting Effective Representation through Consumer Sovereignty and Freedom of Choice for All Criminals,” *American Criminal Law Review*, 31: 71-122.

Schultz, Connie. 08 June 2004. “City to pay $1.6 million for man’s prison time,” *The Plain Dealer*.

Stuntz, William J. 1997. “The Uneasy Relationship Between Criminal Procedure and Criminal Justice,” *Yale Law Journal*, 107: 1-75.

Teichroeb, Ruth. 13 March 2004. “They Sit in Prison – But Crime Lab Tests are Flawed,” *Seattle Post-Intelligencer*.

Teichroeb, Ruth. 22 July 2004. “Rare Look Inside State Crime Labs Reveals Recurring DNA Test Problems,” *Seattle Post-Intelligencer*.

Thompson, William C. 1995. “Subjective Interpretation, Laboratory Error and the Value of Forensic DNA Evidence: Three Case Studies,” 96 *Genetica*, 96: 153 -?

Thompson, William C. 1997. “Accepting Lower Standards: The National Research Council's Second Report on Forensic DNA Evidence,” *Jurimetrics*, 37: 405-?

Thomson, M. A. 1974. “Bias and Quality Control in Forensic Science: A Cause for Concern,” *Journal of Forensic Sciences*, 19: 504-517.

Vaughan, Allen. 17 August 2004. “Crime Lab Official Stole Drugs, Charges Say: The County Dismissed or Declined 284 Cases Because of the Charges,” *News-Leader* (of Springfield, Missouri).

Washington, Michelle. 1 October 2004. “Governor asks Forensics Lab to Test Random DNA Samples,” *The Virginian-Pilot*.

Williamson, Oliver. 1976. “Franchise Bidding for Natural Monopolies -- in General and with Respect to CATV,” *Bell Journal of Economics*, 7(1): 73-104.

Yardley, Jim. 2 May 2001. “Oklahoma Inquiry Focuses on Scientist Used by Prosecutors,” *New York Times*.

Young, Richard and Andrew Sanders. 1994. “The Royal Commission on Criminal Justice: A Confidnence Trick?” *Oxford Journal of Legal Studies*, 14(3): 435-448.

Appenix I

Evidence on the Reliability of Police Forensics

Today we have no adequate measure of the reliability of the forensic analysis actually performed in government crime labs. It is not clear, however, how anyone, including a national government, might measure the forensic error rate. There can be no external, non-forensic test of the correctness of forensic analyses. Nor is it likely that forensic workers could be induced to fully disclose their past errors or even privately recognize them as such. It is clear, however, that forensic analysis is not sufficiently reliable. Most of my evidence for this claim, presented presently, comes from the US. But Europe has experienced similar problems, especially in Great Britain and I will briefly discuss European evidence near the end of this section.

There seems to be more evidence of forensic error in common-law countries, especially the US, than in civil-law countries. It is possible that this difference reflects an advantage of the civil-law system, in which expert witnesses are neutral officers of the court. It is also possible, however, that the adversarial system of common-law countries creates specific incentives for the discovery of forensic error and that, accordingly, more evidence for such error has emerged in them. Thus, the evidence given here may reflect problems fully present, but not fully appreciated in the civil-law countries of Europe. It seems reasonable to conjecture that the problems that seem to affect American forensics are present in at least some significant degree in Europe even beyond the UK.

The case of Josiah Sutton provides dramatic illustration of some current problems with forensics.[17] In 1999 Sutton was convicted of raping a woman. Two men abducted the woman at gunpoint from her Houston apartment complex, raped her, and left her in a field. While driving her car later, the woman saw Sutton with a friend and thought they were her attackers. She notified the police and the two suspects were arrested. Sutton's friend was released after preliminary tests of his body fluids failed to match samples from the victim and her car. Sutton's fluids were reported as a possible match. Sutton was arrested and tried. He was found guilty and sentenced to serve 25 years. He was 16 at the time of his conviction. The case against Sutton was based largely on DNA evidence. A forensic expert from the Houston Crime Lab testified in court that DNA from the semen found on the victim's clothes matched that of the defendant, Sutton. "The testimony strongly implied that this was a unique match, that Mr. Sutton was the only person in the world that would have this DNA pattern, when really thousands and thousands would," according to a DNA expert brought into the case (Khanna, 8 February 2003). Sutton was freed in March 2003 when the original DNA evidence was discredited. Later, new DNA tests proved that he could not have been the source of the semen recovered from the victim and crime scene. He served 4½ years in prison.

The release and subsequent exoneration of Josiah Sutton came in the context of a general review of the DNA/Serology section of the Houston Crime Lab, where the tests were performed (FBI Director, 2002). The report found serious inadequacies in area after

[17] Most of my information on Josiah Sutton and the Houston Crime Lab comes from articles in *The Houston Chronicle* beginning 16 November 2002. As of this writing, they are archived on the paper's searchable website, http://www.chron.com. Credit for breaking the case seems to go to David Raziq and Anna Werner of the Houston television station KHOU (Raziq and Werner 2004). The Houston case has gained increasing attention since an earlier draft of this paper was submitted to this journal. On 25 August 2004 it was the premier example of "faulty forensics" in a television show broadcast nationally in the US on the "Crime TV" cable channel. It was also the subject of a national news story broadcast by *NBC Nightly News* on 5 October 2004.

area. "The laboratory is not designed to minimize contamination due to the central screening area being used by serology, trace, and arson. Better separation of these disciplines is needed. The audit team was informed that on one occasion the roof leaked such that items of evidence came in contact with the water" (FBI Director, 2002, p. 21). Evidence in storage freezers was not properly sealed. It could not be established whether forensic workers wore gloves and lab coats. "Procedures for calibration of equipment have been written," the report indicates, "however, they are not being followed. Logs are not available documenting repair of equipment and calibration prior to being used in casework analysis" (p.34). Lab reports were grossly inadequate. There were no "written procedures for taking and maintaining case notes" (p. 34). Moreover, "Screening notes are very minimal and provide little information. Screening notes do not include a description of the item, what probative stains were identified, how the stains were identified, and what stains were collected" (p. 34). Lab reports were sloppy. They "do not consistently include: case identifier, description of evidence examined, a description of methodology, locus, results and/or conclusions, an interpretative statement, date issued, disposition of evidence (including any depleted samples), and a signature and title of the analyst"(p. 35). Laboratory personnel did not "have the education, training and experience commensurate with the examination and testimony provided" (p.11). The lab even lacked written procedures for the "cleaning of screening areas, common work areas, and equipment" (p. 22). The lab was similarly sloppy regarding the preparation of reagents. "One bottle in the lab had two dates on it and it was unclear which was the date of preparation" (p.30). In June 2002, the lab's director lied about his credentials in court,

falsely claiming that he had a PhD in biochemistry from the University of Texas (Khanna, 10 September 2003).

Risinger *et al*. (2002) say DNA tests "can have surprising problems under some circumstances" (p. 6, n. 17). Citing Thompson (1995 &1997) they claim certain DNA tests "can present highly ambiguous results when mixed samples are involved, which require the same kinds of subjective human interpretation as, say, toolmark or bitemark identification." An expert in DNA typing has told me that this claim about mixed samples holds even for the latest DNA technology. The *Seattle Post-Intelligencer* notes that with the PCR technique, DNA evidence can be contaminated if the technician talks while he works (Teichroeb 22 July 2004). In this context it not surprising that the Houston Crime Lab is not the only laboratory performing unreliable DNA tests.

The case of the Houston Crime Lab is only one of several similar cases in the US and Canada today. Several of them involve DNA testing. Recently problems have been identified in police labs in Baltimore (Brave, 2004), Boston (Saltzman 2004), Cleveland (Schultz 2004), Indianapolis (Horne 2004), Las Vegas (Puit 2002), Manitoba (Macafee 2004), Missouri (Vaughan 2004), Montana (Anez 2004), Fort Worth (Douglas 2004), Virginia (Washington 2004), and Seattle (Teichroeb 13 March 2004).

Proficiency tests give us another source of information on the quality of forensic work. In the US, the Law Enforcement Assistance Administration gave nationwide proficiency tests for forensic labs between 1974 and 1977. The tests are known as the LEAA tests. Over 200 labs volunteered to for a battery of 21 proficiency tests. The results were poor. "Indeed, only one quarter of the participating labs provided entirely

acceptable responses in all cases" (Jonakait 1991, p.110).[18] The "percentage of unsuitable conclusions reached 71% in one blood test, 51% in a paint test, and 67% in a hair test" (Jonakait 1991, p.111).

Jonakait (1991, pp. 114-115) gives three reasons to believe that error rates in the LEAA tests may be less than error rates in daily forensic practice. First, "only volunteer laboratories were tested, and each participated only in the areas of its choice." Second, the labs knew they were being tested; it was not a blind test. Third, "the test samples were much simpler than those a forensic scientist faces in actual casework."

Apparently, there has been only limited improvement since 1977. Peterson *et al*. (1995a, 1995b) published a two-part study of proficiency test conducted by the Forensic Sciences Foundation (FSF) and Collaborative Testing Services (CTS) from 1978-1991. Peterson was involved in the original LEAA tests. Like the LEAA tests, the FSF/CTS tests were voluntary, "open" rather than blind, and characterized by samples that were sometimes less challenging and complex then samples from the field (Peterson *et al*. 1995a, p. 997). In the fingerprint test, however, the testing service "attempted to simulate actual conditions by creating smudged, elongated, compressed and other irregular latent print specimens" (1995b, p. 1010). Peterson and his co-authors directly compare the FSF/CTS and LEAA tests. They find improvement in most areas and a decline the

[18] Jonakait (pp. 110-111) reproduces the following table from the report:

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%
60.0-69.9%	9.4%
50.0-59.9%	3.0%
Below 50%	2.6%

testing of paint. They group forensic techniques into three categories according to the results of the FSF/CTS tests. "Fibers, paints (automotive and household), glass and body fluid mixtures all have improper comparison rates exceeding 10%" as did "animal and human hair" analysis (1995b, p. 1028). In over 10% of the responses on the tests for identification of these areas a positive conclusion was drawn that was, in fact, false and mistaken. The best group includes "finger and palm prints, metals, firearms, and footwear" as well as "bloodstains and drugs" (p. 1028).

Although Peterson *et al*. consider fingerprints a reliable area, the rate of false identifications for the period they cover was 2%. Peterson *et al*. estimate that finger prints appear "in about 7% of the felony case filings" each year (1995b, p. 1028). The Bureau of Justice Statistics reports that "about 924,700 adults were convicted of a felony in State courts in 2000"[19] and another 77,000 in U.S. district courts.[20] These numbers suggest that in the US about a million felony cases are brought to trial and concluded each year, of which 70,000 involve fingerprint evidence. It seems that of these 70,000 cases at least 2%, which is 1,400, involve a false identification. In some cases the error will create a false exoneration. It seems probable, however, that most such errors will produce a false conviction. The identification of a latent print from lifted from a crime scene with an inked print from a suspect or data bank is likely to produce a conviction of the person identified. I do not know how accurate the estimate of 1,400 false convictions a year might be, but there is some reason to suspect it is low. As I have argued, the proficiency tests producing the crucial estimate of a 2% error rate are likely to understate the true fingerprint error rate. The abysmal results of the 1995 test were not included in

[19] http://www.ojp.usdoj.gov/bjs/glance/felconv.htm, viewed 11 October 2004.
[20] http://www.ojp.usdoj.gov/bjs/glance/tables/fedipctab.htm, viewed 11 October 2004.

the studies of Peterson *et al*. These more recent results created "shock" and "disbelief" in the fingerprint community (Grieve 1996, p. 524). Over 30% of answer sheets submitted included at least one false identification and "one if five participants would have provided damning evidence against the wrong person" if the test had been "actual casework" (Grieve1996, p. 526). Whether the rate of false identifications for fingerprints is closer to 2% or 20%, it is (literally) infinitely higher than zero rate sometimes claimed by fingerprint examiners.

Saks (2003) reports on some other disappointing results. Citing Bowers (2002), Saks says, "The first and only published research evaluating the accuracy of forensic dentists revealed an average of 64% false positives and an average of 22% false negatives" (p. 1169). Moreover, "Three similar studies by the forensic odontology community were conducted but not published (because, I am told by Dr. Bowers, those results were deemed unsuitable to be made public)" (Saks, p. 1169, n. 7). Hair analysis is also unreliable. If we consider tests for mitochondrial DNA to be accurate, then the more tradition technique of examination under a microscope "yields a positive result in nearly 35% of the cases in which the facts are negative" (Friedman 2003 as quoted in Saks, p. 1168).

The conclusion Jonakait (1191, pp. 123-124) drew in 1991 seems valid today. "In sum, a review of the data revealed by proficiency studies indicates that lab performance is inadequate and unreliable. The most thorough of the tests, the LEAA study, showed abysmal performances, and all subsequent testing indicates that problems persist."

In 1997 the Justice Department's Office of the Inspector General (OIG) issued a report highly critical of the FBI crime lab. The OIG's investigation was stimulated by the

allegations of a "whistleblower," namely, Supervisory Special Agent Frederic Whitehurst, a Ph.D. scientist employed in the FBI Laboratory.

The OIG report found "significant instances of testimonial errors, substandard analytical work, and deficient practices" (Office of the Inspector General, 1997). David Williams, an FBI investigator in the Explosives Unit was found to have "tailored" his testimony "to the most incriminating result" in two separate trials, namely, that of World Trade Center Bombing of 1993 and that of the Oklahoma City Bombing of 1995 (Office of the Inspector General 1997). In the World Trade Center trial, Williams' identification of urea nitrate as the explosive used in the crime "was based not on science but on speculation based on evidence linking the defendants to that explosive."

> Much the same thing happened in the Oklahoma Bombing case.
>
> His categorical identification of the main charge as ANFO [ammonium nitrate fuel oil] was inappropriate based on the scientific evidence available to him. Here, Williams did not draw a valid scientific conclusion but rather speculated from the fact that one of the defendants purchased ANFO components. His estimate of the weight of the main charge was too specific, and again was based in part on the improper, non-scientific ground of what a defendant had allegedly purchased. In other respects as well, his work was flawed and lacked a scientific foundation. The errors he made were all tilted in such a way as to incriminate the defendants" (Office of the Inspector General 1997).

Concerning the Trade Center trial, the OIG report says, "Ultimately, Williams conceded during our investigation that he had no basis from the crime scene for determining the type of explosive used, acknowledging that based on the crime scene the main charge could have been anything" (Office of the Inspector General 1997).

The case of David Williams is only one of several in which the OIG found problems ranging from substandard work to false testimony. The report found problems of scientifically flawed testimony, inaccurate testimony, testimony beyond the

examiner's expertise, improper preparation of laboratory reports, insufficient documentation of test results, scientifically flawed reports, inadequate record management and retention, failures by management, and a flawed staffing structure of the explosives unit (Office of the Inspector General 1997).

The Williams case is particularly striking. And yet, the report "did not substantiate" allegations "that Laboratory examiners had committed perjury or fabricated evidence" (Office of the Inspector General 1997). Indeed, the OIG report specifically cautions against drawing pessimistic conclusions from the cases it studied. The OIG investigated only certain cases involving three units within the scientific analysis section (SAS), which is, itself, only five sections of the lab. "Our findings and conclusions regarding certain cases in those units," says the report, "should not be imputed to other cases within those units, nor to other units in the SAS or other sections of the Laboratory that we did not investigate" (Office of the Inspector General 1997). It may be that the FBI crime lab has behaved in an exemplary manner in all other cases. But if the cases examined by the Justice Department are representative, the FBI Crime Lab has instead produced much substandard work. The website of the American Society of Crime Lab Directors/Laboratory Accreditation Board now lists the FBI as an accredited lab, although it is probably an open question whether accreditation has improved things.

Further evidence of the inadequacy of current forensic practice exists in the form of a long list of forensic workers whose testimony has been found to be false or inadequate. These include Fred Salem Zain, Ralph R. Erdman, Loise Robbins, Michael West, and Thomas N. Curran.[21] For almost 15 years, beginning in 1979, Fred Zain

[21] Information on these cases can be found in Kelly and Wearne (1998) which is my source for the quotes in this paragraph. Further information can be found in Giannelli 1997.

"testified as an expert in dozens of rape and murder cases [in West Virginia and Texas] about tests he had never done and results he had never obtained." A review of his work in Texas "found rampant fraud and falsification. In one case, Zain had testified about blood evidence when no blood had been found; in other cases he reported performing tests his lab was incapable of doing" (Kelly and Wearne 1998, p.13). Ralph Erdman "faked more than 100 autopsies on unexamined bodies, and falsified dozens of toxicology and blood reports. Dozens of other autopsies were botched. In one case he lost a head" (Kelly and Wearne 1998, p.13). Louise Robbins "claimed the ability to match a footprint on any surface to the person who made it. Robbins appeared as an expert witness for over a decade in more than 20 criminal cases throughout North America before her claims were thoroughly debunked" (Kelly and Wearne 1998, p.13). In the 1990s, Michael West used "long-wave, ultraviolet light and yellow-lensed goggles to study wound patterns on a body." Unfortunately, he was the only one who could detect such patterns (Kelly and Wearne 1998, p.13). A 1975 investigation revealed that FBI Special Agent Thomas N. Curran had "a staggering record of perjury, incompetence and falsification." Like Fred Zain, he gave testimony about tests that were never performed. He "lied repeatedly under oath about his credentials and his reports were persistently deceptive" (Kelly and Wearne 1998, p.14).

More recently, the testimony of Pamela Fish and Joyce Gilchrist has come under scrutiny. In May of 2001 the *New York Times* reported on Joyce Gilchrist who worked on over 3,000 cases from 1980 to 1993. In 2001, the FBI issued a "report that found she had misidentified evidence or given improper courtroom testimony in at least five of eight cases the agency reviewed" (Yardley 2001). A report by Dr. Edward Blake

and criminalist Alan Keel of Forensic Science Associates called the work of Chicago forensic worker Pamela Fish into question. In one case the report found that "Fish's representation of her data . . . can be viewed only as scientific fraud" (Mills and Possley 2001).

The Innocence Project of Yeshiva University's Benjamin N. Cardozo Law School reports that "In twenty-five of the first eighty-two DNA exonerations [that they helped to bring about in the US], scientists and prosecutors presented bad or tainted evidence to the judge or jury" (http://www.innocenceproject.org/causes/junkscience.php).

The situation in the UK seems to be about the same at that in the US. (See Greer 1994, Young and Sanders 1994, Bernstein 1996, Griffin 2000/2001, Browne, Williamson, and Barkacs 2002.) The United Kingdom experienced a series of highly publicized and politically important cases of "miscarriage of justice."

Berstein (1996) notes the case of the forensic worker Dr. Alan Clift, especially in the Preece case. According to Berstein, "investigation revealed that the expert in question, Dr. Alan Clift, had engaged in a wholly inappropriate pattern of advocacy in favor of the prosecution in a series of cases. Clift was forced into early retirement" (p. 167, note 316). Griffin (2000/2001, pp. 1251-1252) reports that exculpatory forensic evidence was suppressed in several of these cases, including the Birmingham Six, the Guilford Four, and the R.V. Kiszko trial. McQuillan (2004) reports that the first person in Scotland to be convicted "solely on the basis of DNA evidence" was wrongly convicted and that the incriminating test was vitiated by cross contamination. (McQuillan cites *The Scotsman* of 23 November 2003.)

The summary assessment of Browne, Williamson, and Barkacs (2002) may also be an apt summing up for this part of the current paper. "Comparative study of the judicial disposition of expertise," they comment, "reveals similar struggles in search of dependable, wise, and 'unbiased' expertise in the United States, England, Korea, and France. In each instance, there is an implicit recognition that in the main, expertise is perspectival. It flows from habits, suppositions, financial and professional gain, and the constraints of professional training in a particular field" (2002, p. 101).

The formal and anecdotal evidence available to us point to an ongoing crisis of forensic science. Saks *et al*. (2001) say, "some forensic science expert witnesses are in a position where they can manufacture evidence merely by wishing it into existence, and evidence suggests that some of them have done precisely that" (p. 700). Forensic work is frequently substandard and sometimes fraudulent. The reasons for this crisis are to be found in how forensic work is organized.

Appendix II

A Cost Estimate

The Bureau of Justice Statistics promises a "Survey of Forensic Crime Laboratories," covering the year 2002, which has not yet been published. Thus, I do not have reliable estimates of the costs of forensics. I can, however, construct a rough estimate of the cost of a DNA test. The *Bureau of Justice Statistics Bulletin* of January 2002 reports that in 2000 the mean annual budget for public DNA labs (when that could be separated from the budget of the general crime lab) was $464,000. It also reported that 110 such labs then existing processed 24,790 cases and 148,347 samples from convicted criminals. The samples were independent of the cases. Cases and samples are not the same. Nevertheless, we can add them to get 173,137 "files," as I will call them. Dividing the number files by the number of labs yields a ratio of 1,574 files per lab. Dividing the workload per lab into the mean lab budget yields $294.80 dollars per file. Presumably, DNA tests are more expensive than most other forensic tests. Thus, it is a conservative estimate to guess that the cost of forensics in an individual case is $300.00. The true number is probably lower. If we make the further pessimistic assumption that competitive self regulation would require one more test per case, we conclude that the extra cost per case would be $300.00.