

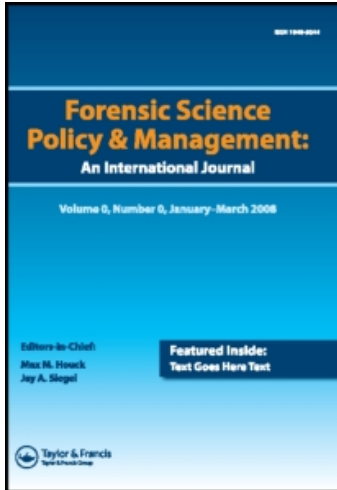
This article was downloaded by: [King, William R.][Sam Houston State University]

On: 9 March 2010

Access details: Access Details: [subscription number 918619870]

Publisher Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Forensic Science Policy & Management: An International Journal

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t790627369>

Assessing the Performance of Systems Designed to Process Criminal Forensic Evidence

William King ^a; Edward Maguire ^b

^a Sam Houston State University, Huntsville, TX, USA ^b American University, Washington, D.C., USA

Online publication date: 23 February 2010

To cite this Article King, William and Maguire, Edward(2009) 'Assessing the Performance of Systems Designed to Process Criminal Forensic Evidence', *Forensic Science Policy & Management: An International Journal*, 1: 3, 159 – 170

To link to this Article: DOI: 10.1080/19409041003611143

URL: <http://dx.doi.org/10.1080/19409041003611143>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Assessing the Performance of Systems Designed to Process Criminal Forensic Evidence

William King¹ and Edward Maguire²

¹Sam Houston State University, Huntsville, TX, USA

²American University, Washington, D.C., USA

Abstract *This paper examines methods that can be used to assess the performance of public organizations and systems designed to process criminal forensic evidence. There is considerable literature devoted to applying scientific and technical knowledge to criminal forensics techniques, such as DNA, AFIS fingerprint systems, and IBIS ballistics systems. However, the forensic science literature tends to overlook the nature and dynamics of organizations and systems responsible for finding, gathering, transporting, and processing physical evidence. How should the performance of these organizations and systems be measured, and what benchmarks exist for comparing their performance relative to their peers? We present a framework for establishing performance measures for organizations and systems that process physical evidence. These measures can serve as a valuable tool for enabling managers to assess the performance of organizations and systems and to evaluate the impact of reforms.*

Keywords Performance measurement, forensic, evidence, organizations, systems

Introduction

Various technical methods of processing physical evidence collected from crime scenes have been developed since the pioneering work of Hans Gross and Edmond Locard more than 100 years ago (Saferstein, 1998). Over the years an industry has developed for gathering and processing physical evidence. We refer to it here as the forensic evidence processing industry, which encompasses *police agencies, coroners, medical examiners, and crime laboratories* that gather and process physical evidence. The industry also includes the agencies, units, and personnel charged with securing, transporting, and storing this evidence.

As defined above, the forensic evidence processing industry is large. In the U.S. alone there are approximately 389 publicly funded crime labs employing 11,900 full-time equivalent employees, and costing nearly \$1.2 billion to operate annually (Durose 2008). There are more than 17,000 publicly funded police agencies in the U.S., employing nearly 1.1 million full-time employees (Maguire et al. 1998; Reaves 2007) and costing more than \$43.3 billion

annually (Pastore and Maguire, 2003: Table 1.43.2003). Finally, there are more than 2,300 death investigation jurisdictions in the U.S., served by coroners and medical examiners (Committee on Identifying the Needs of the Forensic Science Community, 2009, p. 246). These coroners and medical examiners may also submit evidence to crime labs, and they generate information that may contribute to criminal investigations.

Each police agency is capable (at least in theory) of collecting physical evidence, although some do not. As we will explore shortly, in some instances these local agencies, primarily larger agencies, will process this physical evidence in their own lab. In other instances they will either store the evidence or submit it to an external crime lab. Not all physical evidence collected at crime scenes is submitted to crime labs for analysis, although we don't know of any studies that assess the percentage of evidence *not* submitted. Some police agencies adopt a hybrid approach, analyzing certain types in-house (like fingerprints or ballistic evidence) and submitting other types (like DNA) to an external crime lab. Likewise, coroners and medical examiners may submit evidence such as projectiles and serological samples to crime labs.

In sum, the forensic evidence processing industry is a large, expensive undertaking and it is an important part of the criminal justice system. The industry has come

Received 2 October 2009; accepted 7 January 2010.

Address correspondence to William King at Sam Houston State University, College of Criminal Justice, Box 2296, Huntsville, TX 77341-2296. E-mail: william.king@shsu.edu

Table 1. Percent of Police Agencies with Primary Responsibility for Lab Functions

Population served	Fingerprint processing	Crime lab services	Ballistics testing
All sizes	25%	4%	2%
1,000,000 or more	100%	81%	75%
500,000-999,999	96%	50%	50%
250,000-99,999	91%	63%	48%
100,000-249,999	87%	38%	10%
50,000-99,999	67%	21%	3%
25,000-9,999	51%	12%	2%
10,000-24,999	39%	5%	2%
2,500-9,999	27%	2%	1%
Under 2,500	12%	2%	1%

Source: Hickman & Reaves, 2001, p. 7.

under increased attention in recent years after a number of well-publicized scandals in crime labs (Difonzo 2005). A distinguished panel assembled by the National Academy of Sciences recently concluded that the industry needs “an upgrading of systems and organizational structures, better training, the widespread adoption of uniform and enforceable best practices, and mandatory certification and accreditation programs. The forensic science community and the medical examiner/coroner system must be upgraded if forensic practitioners are to be expected to serve the goals of justice” (Committee on Identifying the Needs of the Forensic Science Community, 2009, p. 15).

Similarly, another recent report from the Innocence Project (2009, p. 3) concluded: “Nearly five years after Congress passed legislation to help ensure that forensic negligence or misconduct is properly investigated, extensive independent reviews show that the law is largely being ignored and, as a result, serious problems in crime labs and other forensic facilities have not been remedied.” Though some view these critiques as a “wrongful conviction” of the industry, these crises clearly illustrate the need for forensic systems to develop methods for measuring and documenting their performance (Collins and Jarvis 2009). These measures will enable systems that perform well to demonstrate their performance to key stakeholders, and systems that perform poorly to diagnose their weaknesses as an important first step toward improving performance.

Police agencies, crime laboratories, coroners, and medical examiners primarily do the work of forensic evidence processing; public organizations so far have tended to avoid or resist the establishment of agency-level performance measures or organizational report cards (see Gormley and Weimer 1999; Maguire 2005; Moore 2002). We contend that there are two major roadblocks to the establishment of quality performance measurement in forensic systems. First, there is substantial heterogeneity in the forensic evidence processing industry, including variabil-

ity in the structure of the organizations and systems that do the work, as well as in the occupational composition of the industry. Second, the industry already has in place a series of quality-control mechanisms that may appear on the surface to constitute performance measures, but in reality are poor substitutes. These quality controls may have value in their own right, but they do not constitute effective performance measures for reasons that we will explain shortly.

Industry Heterogeneity

One issue that complicates the development of performance measurement is the sheer degree of heterogeneity in the forensic evidence processing industry. Here we focus on heterogeneity in both the structure and the occupational composition of the industry. In the United States, forensic evidence processing is accomplished by a fragmented and heterogeneous network of arrangements between publicly funded police agencies and crime laboratories as well as private laboratories to which some work (primarily DNA analysis) is outsourced (Childs, Witt, and Nurtegin 2009). The majority of crime labs operate at the state or regional level, and therefore many of them receive inputs (evidence that needs to be processed) from multiple agencies. A 2002 study found that among publicly funded crime labs in the United States, 57.8% were state or regional labs, 18.5% were county, 14.2% were municipal, and 9.4% were federal labs (Peterson and Hickman 2005). Some laboratories were independent agencies, some were contained within police agencies, and others were part of a larger system of laboratories. Many of them reported outsourcing to private laboratories for a variety of services, primarily DNA analysis (Durose 2008; Peterson and Hickman 2005; Steadman 2002). Moreover, even *within* police agencies and crime laboratories, the responsibility for forensic evidence processing is often assigned to many different functional niches or special units, some of which may communicate and work with one another seamlessly, and others not. Little is known about the interorganizational arrangements for forensic evidence processing both within and between agencies.

Fortunately, recent surveys of police agencies and crime labs in the United States provide a basic glimpse of the structure of the industry (Childs, Witt, and Nurtegin 2009). Only 25% of local police departments in the United States reported in a 1999 survey that they have primary responsibility for fingerprint processing (Hickman and Reaves 2001). The vast majority of police agencies in the United States are very small, and the size of the organization (and the size of the population it serves) appears to be a key determinant of whether police agencies have responsibility for processing fingerprints. Only 12% of local police agencies serving populations less than 2,500 have

responsibility for processing fingerprints, compared with 87% of agencies serving populations of 100,000 to 250,000, and 100% of agencies serving more than a million people. Only 4% of local police departments provide crime laboratory services, and only 2% provide ballistic testing. But a similar pattern holds with regard to organizational size and population served: the larger the police agency, the more likely that it has primary responsibility for the function. Table 1, drawn from Hickman and Reaves (2001, p. 7), illustrates these patterns in more detail. In general, large police agencies in the United States manage their own crime laboratories in-house; smaller agencies must secure forensic services from external crime laboratories.

In another national survey of local police agencies in the United States in 2000, 20% of respondents reported having access to automated fingerprint identification systems (AFIS); 5% reported having “exclusive or shared ownership,” while another 15% reported having only a “remote access terminal” (Hickman and Reaves 2003, p. 24). Similar patterns held with regard to the size of the organization or its jurisdiction. All of the agencies serving populations over a million reported having access to AFIS (whether owned or remote); only 85% of agencies serving populations of 100,000 to 250,000 had AFIS access, and only 13% of agencies serving populations under 2,500.

Surveys of publicly funded crime laboratories also reveal interesting patterns useful for understanding heterogeneity in the forensic industry in the United States. Although there is a tendency to view all crime laboratory personnel as the same—the people wearing the white lab coats—the forensic sciences are comprised of a number of different scientific disciplines, many with their own educational requirements, professional associations, and certification standards. A 2005 survey of crime labs revealed that 58% of employees were analysts or examiners, 13% were managers, 10% were technical support personnel, 8% were clerical support personnel, 6% were crime scene technicians, and 5% held other roles (Durose 2008). As of April 2009, the Forensic Specialties Board had accredited eight different forensic specialty organizations, though many others exist (Forensic Specialties Accreditation Board Inc. 2009). The functions performed by these laboratories also vary widely. While 89% analyzed controlled substances, only 53% performed DNA analysis, and only 12% handled computer crimes (Durose 2008). Table 2, taken from Durose (2008, p. 3), illustrates the variety of functions performed by crime laboratories in the United States. Taken together, these findings suggest that crime laboratories are complex organizations both occupationally and functionally differentiated in a variety of ways.

Like crime labs and police agencies, the system in place for death investigation in the U.S. is also fragmented, consisting primarily of two types of structures: coroners and medical examiners. According to a recent report by the

Table 2. Percent of Publicly Funded Crime Labs Performing Various Functions

<i>Function</i>	<i>%</i>
Controlled substances	89%
Firearms/toolmarks	59%
Biology screening	57%
Latent prints	55%
Trace evidence	55%
DNA analysis	53%
Toxicology	53%
Impressions	52%
Crime scene	40%
Questioned documents	20%
Computer crimes	12%
Total labs	351

Source: Durose, 2008, p. 3.

National Academy of Sciences (Committee on Identifying the Needs of the Forensic Science Community 2009, p. 245–246):

In total, there are approximately 2,342 separate death investigation jurisdictions. Of 1,590 coroner offices in the United States, 82 serve jurisdictions with more than 250,000 people; 660 medium-sized offices serve between 25,000 and 249,999 people; and 848 offices serve small jurisdictions of fewer than 25,000 people. The hodgepodge and multiplicity of systems and controlling statutes makes standardization of performance difficult, if not impossible.

That same report concludes that fragmentation in medico-legal death investigation systems limits interagency communication, effectiveness, and the adoption of best practices.

The heterogeneity and fragmentation of the forensic evidence processing industry complicates the development of performance measures. Whatever measurement system is adopted will need to be sufficiently flexible to accommodate a wide variety of systems, structures, and occupations. While the complexity inherent in the system for processing criminal forensic evidence in the United States makes developing performance measures challenging, we contend that the challenge is both worthwhile and surmountable.

Current Quality Control Measures

The agencies comprising the forensic evidence processing system conduct important work and there have been many attempts to improve their performance. Most of these attempts have been directed toward crime labs and their employees and have focused on credentialing and process improvement. Comparatively little effort has

been devoted to understanding and improving the role of police agencies in gathering, securing, processing, or storing forensic evidence. Efforts to improve the investigation of death in the U.S. have focused on replacing untrained or poorly trained coroners with medically trained medical examiners. Few of the existing quality-control measures have focused on system-wide outcomes or the overall performance of forensic systems. In the parlance of performance measurement, existing measures focus much more on processes and outputs than on outcomes (Gormley and Weimer 1999; Hatry 1999; Maguire 2005).

This myopic focus on process improvement and credentialing is not surprising, given that between 1972 and 2005 the flagship journal for forensic scientists (*Journal of Forensic Sciences*) only published one article with the word “organization” in the title (an article on the Italian judicial police; Pisano 1979), two original articles investigating “management” (Dibdin 2001; Nelson 1990), and zero articles on leadership in crime labs or police agencies. Topics like organization, management, leadership, and system performance have been largely ignored in the forensics literature (for recent exceptions, see Houck et al. 2009; Speaker 2009a, 2009b). Little effort has been expended in devising comprehensive, outcome-based performance measures. Contemporary discussion of quality-control measurement in forensic evidence processing is based on four main topics: proficiency testing, accreditation of crime labs, certification of individuals, and the development of industry guidelines. Sometimes these quality control mechanisms are mistaken for performance measures. We will explain why that although these mechanisms may be very useful, they do not constitute effective performance measures.

Proficiency Testing

Most labs undergo periodic proficiency testing of laboratory techniques to see if they are conducting analyses accurately (Peterson et al. 2003; Field 1976). A 2002 census of publicly funded crime laboratories in the U.S. found that 97% engaged in proficiency testing (Peterson and Hickman 2005). Almost all of these used “*declared tests*, a type of test in which the examiner knows he/she is being tested” (Peterson and Hickman 2005, p. 11). Fifty-four percent of labs used *random case reanalysis*, “where examiners’ completed prior casework is randomly selected for reanalysis by a supervisor or another examiner” (Peterson and Hickman 2005, p. 11). Only 26% of labs used *blind tests* in which “the examiner doesn’t know the sample being analyzed is a test sample” (Peterson and Hickman, 2005, p. 11).

Although proficiency testing is important, it only focuses on the technical proficiency of individuals, not other important elements of laboratory or system behavior. For instance, one recent study found that a

laboratory produced meaningful results from its IBIS (integrated ballistics imaging system) analysis of ballistic evidence, but the results of these analyses routinely failed to be delivered to the detectives who could use the information (King and Maguire 2009). Similarly, many jurisdictions have amassed enormous backlogs of unprocessed evidence. Thus, performing analysis in a technically proficient manner is just one part—albeit an important part—of crime laboratory performance. As Hadley and Fereday (2008, p. 8) argue, proficiency testing provides a limited snapshot of performance but “far too much can be read into the apparent assurance such schemes provide.” Proficiency tests can serve as an important component of a more comprehensive suite of performance measures, but alone they are incomplete measures of an organization’s or a system’s performance.

Accreditation of Crime Labs

Crime labs may seek voluntary accreditation by the American Society of Crime Lab Directors–Lab Accreditation Board (ASCLD–LAB) or by other accrediting organizations. Accreditation is growing rapidly in the crime lab industry. A 2002 survey found that 71% of publicly funded forensic crime laboratories were accredited, 61% by ASCLD–LAB and another 10% by other organizations (Peterson and Hickman 2005). By 2005, 82% of labs were accredited, 78% by ASCLD–LAB and another 3% by other organizations (Durose 2008).

The accreditation process assesses the procedures and operations of a lab (by reviewing procedures, past cases, and proficiency testing), as well as the lab’s facilities and equipment. Accreditation is voluntary and may take a number of years to complete. Labs must prepare before initiating the accreditation process; if the accreditation team notes deficiencies, the lab has up to one year to remedy the deficiencies. Accreditation lasts for five years, after which the lab must apply for re-accreditation.

Accreditation is a limited measure of performance because it focuses heavily on processes rather than outcomes. Moreover, it is dichotomous: either a lab is accredited or it is not. Performance measures need to be able to capture a much broader range of variation, preferably at the ordinal or interval level. A dichotomous “measure” like accreditation doesn’t allow for a distinction in performance between accredited agencies. Accreditation may be a useful tool for changing operations and procedures, but it is not a textured measure of performance capable of enabling the comparison of labs to one another or tracking a lab’s performance over time.

Certification of Individual Examiners

Although there is a tendency among outsiders to view all crime laboratory personnel as the same, the forensic sciences are comprised of a number of different scientific

disciplines, many with their own professional associations and certification standards. Several organizations, like the American Board of Criminalists (ABC), certify individual examiners. This is a certification of proficiency in performing specific technical tasks, not overall employee performance. Police officers in some jurisdictions must meet similar certification standards in specific tasks like shooting accuracy or pursuit driving. These are important tasks to be sure, but meeting these minimum standards says little about the overall quality of the officer. Individual certification has not diffused widely among criminalistics technicians. For instance, only 19% of fire and explosion debris forensic analysts reported being certified by ABC in a 1999 survey (Allen, Case, and Fredrick 2000). While individual certification is an important component of improving quality control in crime laboratories, it is not a substitute for more comprehensive outcome-based performance measurement systems.

Guidelines and Standards

According to some, guidelines and standards are “the key to everything associated with ensuring competent performance” in crime labs (Hadley and Fereday 2008, p. 10). A number of industry guidelines and standards have been established, such as those developed by the FBI’s various Scientific Working Groups, such as the Scientific Working Group for Material Analysis and the Technical Working Group for Fire and Explosives. While the development of industry guidelines and standards is important and potentially useful, these mechanisms remain weak unless they are coupled with some type of enforcement mechanism.

While all of these quality control measures have value in their own right, they do not constitute a coherent outcome-based performance measurement strategy for systems designed to process criminal forensic evidence. One of the biggest problems with existing quality control approaches is that they are not *system-focused*; they are heavily focused on crime labs alone. The forensic evidence processing system in the United States is heterogeneous and fragmented. In the journey from crime scene to courtroom, criminal evidence is located, secured, packaged, transported, analyzed, and reported on by personnel from multiple agencies and multiple units within those agencies (Childs, Witt, and Nurtegin 2009). Crime labs are only as good as the evidence submitted to them. Improperly collected evidence, evidence that is contaminated or mislabeled, or evidence that lacks clear directions from the police (as to the question at hand) all hamper the quality of outputs generated by crime labs. Similarly, police and prosecutors are only able to make effective use of criminal evidence that is processed accurately and within a sufficient time frame. The results of analyses must then be routed to the investigators who can act upon these results

to further their investigations. Although it seems implicit that all of these steps will occur, social science research shows that sometimes they don’t (King and Maguire 2009; Schroeder and White 2009). Therefore it is vital to look at the *systems* that process physical evidence, not just individuals, units, or organizations.

In addition, these quality control methods focus largely on processes and not outcomes. Accreditation, certification, standards, and guidelines are all based on the assumption that these mechanisms will improve or sustain performance. Collecting outcome-based performance measures will allow the industry to begin testing these assumptions. Proficiency testing does hold some promise as one among many potential sources of performance measurement data. However, declared tests, which are the most widely used kind of proficiency tests, are of little value relative to random case reanalysis or blind tests. Declared tests are useful for learning whether analysts know how to use the appropriate procedures, but they are not useful for knowing whether analysts use those same procedures when nobody is watching over them.

Organizational Performance Measures

All organizations are designed to produce a product or provide a service. For example, auto manufacturers produce cars and school systems educate children. The easiest way to quantify organizational performance is to measure outputs, such as the number of cars produced or children educated. Although outputs are often simple to count, many output-based measures do not truly capture the performance of an organization (Gormley and Weimer 1999; Hatry 1999; Mohr 1973). The cars may be lemons and the graduates may be illiterate. Moreover, outputs are often susceptible to manipulation by organizations seeking to improve their image without actually improving their performance. For instance if school leaders know their performance will be measured using graduation rates, they can inflate their graduation rates by awarding diplomas to students regardless of their literacy.

An alternative to measuring *outputs* is to measure *outcomes*, or the extent to which an organization accomplishes its mission and goals. (Mohr 1973; Simon 1964). Instead of counting the number of graduates, we should measure their academic proficiency, their job placement rates, or their college acceptance rates. Measuring organizational outcomes is not simple, however. First, organizations seek to achieve multiple goals and thus have multiple outcomes worthy of measurement (Mohr 1973; Simon 1964). Thus, it is better to measure multiple outcomes, especially in public organizations. Additionally, organizational outcomes can often be influenced by a wide range of variables, some of which may not be within the control of the organization. For instance, crime is not solely within

the province of the police; it is affected by family structure, the economy, demographics, and other factors. In general, the more important the outcome, the greater the likelihood that it will be influenced by the work of multiple organizations and social forces, and the more difficult the task of isolating the impact of a single organization or other causal influence. Some outcomes, which we refer to as *proximal* outcomes, can be attributed to the efforts or effects of a specific organization. Other outcomes are further removed from an organization, and the direct lineage or causal linkage between the organization and these outcomes is less clear. We call these *distal* outcomes. For example, high schools hope to produce literate, knowledgeable graduates, a proximate outcome when measured at or near the time of graduation. However, we also expect these graduates to become productive members of society, a distal outcome that is both difficult to measure and to attribute to the efforts of a specific high school.

Because data on outcomes can be difficult to find or expensive to collect, proxy measures are often used. For example, we could measure the performance of a professional sports team using proxies like the number of famous athletes the team employs, the previous record of the athletes, the amount of revenue generated by ticket and merchandise sales, or attendance at home games. But all of these performance measures are crude and indirect when compared to more direct outcome-based measures of performance like the team's win-loss record. The selection of appropriate measures of performance is vital. Focusing on the wrong performance measures can lead an organization to deploy its resources and energies on the wrong things, and can harm performance. Careful analysis of outcome-based indicators of performance led to a revolution in how Major League Baseball teams in the United States are managed. This analytic insight, called *sabermetrics*, allowed the Oakland A's baseball team to produce a highly successful season against teams spending more than two times as much money on players' salaries (Lewis 2003). Similar results have been discovered in other industries that have adopted outcome-based performance measures, including health care, education, air travel, and others (Gormley and Weimer 1999).

Despite challenges involved in crafting outcome-based measures of performance, these measures can be very useful for improving performance. They provide managers with feedback useful for making decisions about how to change operations and structures to improve performance. Crime labs and police agencies serve vital and noble roles in society, and we should endeavor to measure the effectiveness of these (and other) organizations in carrying out their roles. Though challenging, measuring performance outcomes is better than the alternative, muddling along with no clear picture of organizational performance.

The Rationale for Better Performance Measures

Accreditation, credentialing, guidelines and standards, and declared proficiency tests are all important components of creating a high-performance system. But none of these methods is able to measure directly the extent to which an organization or system is successful at meeting its goals. Creating more textured and detailed performance measures is important for three reasons. First, if we want to conduct comparative analyses of crime labs, either to see which labs perform better than others, or to determine the causes or correlates of success, we need better measures of performance. Process-based indicators of performance, such as accreditation and certification, are not sufficiently nuanced. For example, lab accreditation is a dichotomous "measure": labs are either accredited or not. A recent study found that 82% of publicly funded labs in the U.S. were accredited by 2005 (Durose 2008). Yet, not all accredited labs are performing optimally or at equivalent levels. It would be useful to be able to distinguish different levels of performance among accredited labs. Moreover, process-based measures have not been subjected to empirical scrutiny. Although there is a strong belief within the field that accreditation standards are associated with improved performance, other fields have struggled to demonstrate the linkage between accreditation and beneficial outcomes (e.g., Chen et al. 2003; Miller et al. 2005; Thompson et al. 2009). Accreditation and certification are sometimes pursued for their symbolic value rather than (or in addition to) their substantive value for improving performance. Adopting outcome-based performance measures will provide a useful opportunity to test the association between performance and process-based measures like accreditation and certification.

Second, more nuanced measures of performance can serve as a valuable management tool for tracking performance *within* an organization or system. For instance, performance likely varies across different segments within a single lab; among the different lab sections, such as firearms and DNA, but also across competencies, such as writing reports and working closely with the district attorney's office and police agencies. More detailed indicators of performance can help managers determine what factors promote or impede high performance. Moreover, better performance indicators can serve as an early warning system to help labs and police agencies identify performance issues before they grow into much larger crises. In 2008, the Detroit Police Department disbanded its crime lab after allegations of mishandled analyses surfaced in the firearms section. Lab employees contended that the problems were isolated to the firearms section and were not generalizable to other lab sections, but the scandal became so volatile that it resulted in the disbanding of the entire lab (Hunter 2008; Kusluski 2008; Williams 2008).

It is difficult to imagine that the performance deficiencies leading to the disbanding of the lab emerged all of a sudden. Effective performance measures might have been useful for revealing this crisis in the making before it spun out of control (King and Maguire 2009).

Third, organizations and systems without valid performance measures are at risk of having reforms and/or performance measures foisted upon them externally from sources that may not understand what they do or the context in which they operate (Field 1976, p. 53-54). Externally imposed measures tend to originate in the aftermath of a scandal, crisis, or embarrassing media report. During the 1990s, a number of U.S. police agencies (such as Los Angeles and Cincinnati) were subjected to consent decrees imposed upon them by the U.S. Justice Department following highly publicized scandals and crises. Externally imposed performance measures can sometimes do more harm than good. Well-intentioned (but uninformed) outsiders may propose or mandate reforms that may not improve system performance or that might generate unintended consequences. Outsiders may not understand the deeper, more complex processes or problems that produced the initial error or scandal. The grist for outsiders bent upon reforming forensic systems appears regularly in the media. For example, news coverage of falsified forensic reports or inaccurate analyses have emerged in a number of locales, including Douglas County, NE (Ferak 2009), Detroit, MI (Williams 2008), and Houston, TX (Houston Police Dept. 2009). In the absence of appropriate performance measures, forensics managers are at risk of having reforms and less than optimal performance measures foisted upon them. We now propose various performance indicators for systems that process forensic evidence.

Potential Performance Indicators for Forensic Systems

We outline three domains of performance indicators for systems that collect and process forensic evidence. These indicators can be used to compare the performance of peer systems, to track performance of systems over time, to evaluate the effectiveness of reforms or innovations, and to justify additional funding and resources. We discuss these performance indicators in a roughly chronological order in the progression of forensic evidence from crime scene to courtroom. Table 3 summarizes the three domains and the indicators within each domain.

Performance Indicators for Crime Scene Processing

Agencies, units, and individuals differ in the skill with which they find, secure, and process crime scenes. In some jurisdictions, police officers are responsible for processing crime scenes. In others, civilian crime scene

Table 3. Potential Performance Indicators for Forensic Systems

1. Performance Indicators for Crime Scene Processing and Evidence Storage:
<ul style="list-style-type: none"> • Ability to find, secure, and process crime scenes • Ability to locate and package properly physical evidence • Ability to document crime scenes (sketches, notes, photographs, etc.) • Ability to submit properly physical evidence for analysis or storage • Ability to properly store and secure evidence (prevent alteration, destruction, or theft of evidence while stored) • Ability to dispose of/destroy physical evidence when appropriate • Proper use of forensics processing by investigators and police
2. Performance Indicators for Analyzing Evidence:
<ul style="list-style-type: none"> • Speed of analysis • Size and age of backlogs • Accuracy of analysis • Workable ways to expedite and triage cases and analyses • Ability to properly store and secure evidence (prevent alteration, destruction, or theft of evidence while stored) • Ability to dispose of/destroy physical evidence when appropriate
3. Performance Indicators for Information Dissemination, Usage, and Utility
<ul style="list-style-type: none"> • Dissemination of information from forensic analyses to investigators and prosecutors • Understandability of information for investigators and prosecutors • Utility of forensics information for cases, prosecutions, and clearances • Availability of forensics information for investigators and prosecutors • Overall satisfaction of information consumers (investigators, prosecutors, judges) with the forensics information they receive

technicians or laboratory analysts play a role as well. The performance of crime scene personnel during the early hours of an investigation is crucial and in some cases their actions determine, in part, the solvability of the case (Wellford and Cronin 1999). Crime scene personnel should be assessed based upon their ability to find, secure, and process crime scenes. Processing crime scenes entails locating, documenting, and packaging physical evidence, and ensuring that the process is both legally and scientifically sound. For example, the proper submission of evidence requires documenting the chain of custody and ensuring the evidence is not contaminated, altered, or pilfered between the scene and the lab or evidence storage facility. The skills that crime scene personnel exhibit in completing each of these tasks could be gauged by raters with expertise in criminal investigation and the forensic sciences. For example, these raters could periodically attend a random sample of crime scenes and grade the performance of the police and other crime scene personnel. Actors at later steps in the process could also be enlisted to review the performance of personnel at crime scenes. For example, interviews or surveys of crime lab personnel, criminal investigators, and prosecutors could be used to reveal the extent to which crime scenes

are properly secured or processed or evidence properly packaged. Some agencies now outsource evidence storage to outside vendors. Since outsourcing is merely a contractual arrangement for accomplishing the same set of tasks, the public agencies entering into these agreements should still be held accountable for them.

The police and crime labs serve an important role as custodians of physical evidence. Usually evidence is stored in property lockers located in secure facilities. Physical evidence must be properly packaged and maintained to preserve it. For example, evidence cannot get wet or be subjected to excessive heat and humidity, and it must be protected from rodents. Furthermore, evidence must be protected from illicit alteration, pilfering, and theft. Theft from police evidence storage areas is not unheard of. During the 1970s, 193 lb. of cocaine and heroin were stolen from a New York City Police Department property room (Murphy and Plate 1977). Nearly every major city in the U.S. has experienced at least one publicized instance of evidence missing from police property rooms. Finally, some evidence must eventually be destroyed, and agencies should be assessed on how well they perform this task. Property lockers might be searched by auditors to ensure that items slated for destruction have been properly disposed. Auditors can also document cases where evidence is recorded as destroyed (such as firearms) but later “reappears” on the street. In sum, all agencies in the forensic evidence processing system—including police, coroners, medical examiners, and crime labs—should be assessed on their ability to properly store, secure, and dispose of evidence.

The police and investigators could also be judged based on their use of the forensics processing system. Some police agencies may have such difficulty in collecting evidence that they submit little or none, although they may attend crime scenes containing potentially meaningful physical evidence that ought to be collected and processed. Others may be reluctant to submit evidence for processing due to backlogs, processing delays, or back pressure from overburdened labs. It is important to identify systems in which evidence is not routinely collected and submitted for processing. In the 1970s, the California Justice Department sent short questionnaires to police officers who had submitted evidence to a crime lab. The questionnaires asked whether lab personnel were cooperative and whether the analyses were completed in a timely and satisfactory manner. The goal of the survey was to identify and rectify any problems or misunderstandings between police and lab personnel (Smith, 1976, p. 14). A study of twelve U.K. police forces found substantial variation in the proportion of cases in which forensic evidence was submitted to the crime lab for processing (Tilley and Ford 1996). A recent study of homicide investigations in New York City revealed that even when biological material was gathered from

crime scenes, detectives rarely requested DNA analysis (Schroeder and White 2009). Research that seeks to shed light on the decision-making calculus of detectives in deciding what evidence to collect, what evidence to submit for processing, and what information to request from the lab would make a useful contribution.

It seems reasonable to assume that the more forensic information the police request, the better the results of their investigations. For instance, a recent field experiment conducted in five U.S. cities concluded that “property crime cases where DNA evidence is processed have more than twice as many suspects identified, twice as many suspects arrested, and more than twice as many cases accepted for prosecution compared with traditional investigation” (Roman et al. 2008, p. 3). But we know relatively little about the manner in which investigators use forensic information. It is possible that investigators with cases that are essentially solved (e.g., a suspect has confessed, there are no loose ends in the case, and the D.A. thinks the case will be quickly resolved via plea bargain) will *not* request forensics analysis. Perhaps detectives rely on forensic analysis for tougher cases or to bolster cases where the suspect is uncooperative. These are all empirical questions worthy of research because they have significant implications for how forensics evidence is used in reality.

Performance Indicators for Analyzing Evidence

Crime labs, coroners, and medical examiners usually conduct the analysis of physical evidence. In some instances, certain types of evidence are processed “in-house” by special units within police or law enforcement agencies. Two common examples are ballistics and fingerprints. Moreover, some analyses are outsourced to contractors—some public and some private—paid by the originating jurisdiction to handle certain types of analyses like DNA or toxicology. Regardless of which entity carries out the analyses, we suspect there is tremendous variation in the skill, quality, and accuracy with which analyses are carried out. Measuring these differences—either comparatively (comparing peer forensic systems to each other) or over time within a single agency, section, or unit (to detect increases or decreases in performance, or to evaluate the effectiveness of reforms)—is a useful exercise for gauging and improving performance.

First, physical evidence should be processed in a timely manner. The elapsed time between when physical evidence is submitted for analysis and a report is produced is a useful indicator of performance (Smith 1976, p. 14). Elapsed time could be calculated for different sections or units. Another option is to compute a weighted mean for an entire lab. Any comparative measure of lab backlog should be sensitive to the fact that analysis time is dependent on the type and complexity of the analysis.

Thus, a lab that performs mostly routine analyses that can be accomplished quickly, such as controlled substances and toxicology (Peterson and Hickman 2005, p. 8), should not be compared directly to a lab that handles more elaborate, time-intensive analyses such as DNA or computer forensics (Peterson and Hickman 2005, p. 9) unless adequate statistical controls are included in the analysis (Gormley and Weimer 1999; Maguire 2005). The literature on organizational performance measurement is sensitive to the need to select comparable “peer” agencies and to use “risk adjustment” procedures to ensure fairness in comparisons. For example, hospitals use risk-adjusted mortality rates to account for the fact that they may serve clientele with differential risks of death (Gormley and Weimer 1999). Similarly, in comparing the performance of police agencies across cities, criminologists use risk adjusted homicide rates to account for the fact that some cities are at greater risk for violent crime than others, independent of police activities (Maguire 2005).

Speed of analysis is related to another vital measure of performance: the size and age of backlogs. Backlogs are ubiquitous in forensics because evidence cannot be processed instantaneously and many labs have insufficient budgets and personnel (Committee on Identifying the Needs of the Forensic Science Community 2009). More importantly backlogs differ in their size and how long the evidence has been awaiting analysis (the time-depth). Backlogs are defined by the Bureau of Justice Statistics and West Virginia University’s FORESIGHT project (a forensic laboratory improvement project) as evidence that has not been processed within 30 days (Durose 2008; Houck et al. 2009; Peterson and Hickman 2005).

Another cluster of performance indicators deals with the accuracy of analysis, a topic of great interest recently (Committee on Identifying the Needs of the Forensic Science Community 2009). Proficiency testing offers a feasible method for determining whether examiners in labs perform their work accurately (Peterson et al. 2003). Declared (or open) proficiency testing is less optimal than blind or double-blind proficiency testing or random case reanalysis (Koehler 2008; Lentini 2003; Peterson et al. 2003). The testing should be wide enough in scope to provide an accurate snapshot of examiners and labs. For example, if we wanted to measure the performance of the firearms sections at three crime labs using proficiency testing, we would need to adopt a careful and comprehensive approach that includes all the examiners (or a representative sample of them) and draws on multiple tests with different levels of difficulty. Just as in other statistical quality control efforts, much thought would need to be invested in selecting the appropriate random sampling methods.

Labs could also be assessed on their ability to manage their caseloads by triaging cases or parts of an analysis, dropping cases from their workload, and expedit-

ing cases. Some labs, for example, will not analyze evidence unless the case has a suspect or is likely to be prosecuted. Likewise, some labs permit an agency to cancel a planned analysis if that analysis no longer fits the agency’s needs. These policies represent attempts by labs to conserve their resources and allocate the efforts of their employees more judiciously (Durose 2008, p. 7). Further, some labs have established methods for expediting important cases to improve the effectiveness of other system agencies like the police and prosecutors. In the absence of a workable way to expedite cases, labs run the risk of being forced to accommodate less effective systems like responding to crises (“putting out fires”) or to those who make the most compelling demands (“the squeaky wheel gets the grease”). Labs could first be assessed on the presence of a system, such as policies or guidelines for triaging, dropping, and expediting cases. But this would be a process-based measure. The mere presence of a policy is one thing, but an effectively implemented policy is something different. Thus, labs could also be assessed on the efficiency of their systems for triaging, dropping, and expediting cases. Lab employees could be surveyed about how workable they find their agency’s triage and expediting policies. This determination of efficiency would help distinguish labs with unworkable or cumbersome policies from those whose policies improve their efficiency. A sample of triaged and expedited cases could be reviewed or analyzed to determine if the policy is actually triaging or expediting the intended types of cases. It is possible, even in the presence of a formal policy, that cases appropriate for triage and expediting are handled routinely, just like any other case. Finally, investigators and prosecutors could be asked about the track record of labs in triaging, dropping, and expediting cases. This measure has the benefit of taking into account the systemic nature of forensic evidence processing.

Performance Indicators for Information Dissemination and Usage

No matter how accurate the analyses produced by crime labs or other entities, the analyses are not useful until the findings are disseminated to, and used by, those who need the information to conduct investigations and adjudicate offenders. The dissemination and usage of information is vital to the success of the criminal justice process. Therefore, a valuable set of performance indicators is the extent to which forensics information is transmitted to the agencies and individuals who need it and the information is digestible or “user friendly” (Smith 1976). One recent study found that IBIS correlations from homicide and shooting cases were not regularly transmitted to the appropriate police detectives because the internal police system for routing this information did not work properly

(King and Maguire 2009). One source of performance data could be a survey of investigators that asks to what extent they receive forensics information when they request it, how useful the information is, and whether they are able to resolve their questions or concerns with forensic examiners. A similar survey of prosecutors could also yield useful performance information concerning the utility of forensics information.

Once transmitted, the information must be readily accessible to anyone involved in the investigation. Some agencies attempt to ensure that information is available by requiring that investigators attach copies of reports and lab results to case folders. Thus, one performance indicator would involve auditing case folders to ensure that officers include relevant reports. Asking officers and investigators about their patterns of forensic information usage to understand what information is most helpful to investigators and how that information is used could develop another indicator. Furthermore, these queries could gauge the utility of information provided by forensic examinations; in other words, is the information useful and understandable to investigators? For example, investigators could be queried about specific examples where they received forensics reports and/or information from forensic analysts. Possible questions could include to what extent the investigator understood the results of the analysis, whether the analysis was appropriate for the case at hand, and the contribution of the information to the case. These queries could be structured as assessments of the overall utility of forensics information provided to investigators, or the queries could probe deeper by asking about a specific case or cases, perhaps to focus on specific analyses provided by a lab, coroner, or medical examiner.

Finally, we should assess the utility of forensic information for identifying and locating suspects, making arrests, and producing successful prosecutions. One hypothesis is that a greater amount of forensics information will result in the identification of more suspects, more successful prosecutions, or higher clearance rates. But not all forensic analyses are equal in their effects on case solvability. For instance, one recent study found that DNA evidence more than doubled the number of suspects identified and arrested. That same study found that blood evidence resulted in “better case outcomes than other biological evidence” (Roman et al. 2008, p. 3). One set of performance indicators could include the proportion of cases in which a suspect is identified, arrested, and convicted, each categorized by the types of evidence available.

Analysts routinely produce reports that document the findings from their analyses. These reports serve as “inputs” for other criminal justice agencies, such as the police, prosecutors’ offices, and courts. Performing timely and accurate analyses is one step, but converting these analyses into useful information for other agencies is yet another source of potential performance indicators for

crime labs. Using a panel of experts to review a random sample of reports produced could assess the quality of information produced by forensic analyses. One could also survey the consumers of this information, such as police officers, investigators, prosecutors, and judges, to gauge their satisfaction with the information they receive from the lab. These are precisely the kinds of performance indicators that emerge when we think of forensic evidence processing from an interagency or “systems” perspective rather than the much more limited (and common) approach of thinking only about agencies and individuals.

Moderators of Performance

Considerable debate in performance measurement circles has focused on the extent to which agencies and systems should collect data not only on outcomes, but also on factors that are likely to *influence* outcomes. We refer to those factors as “moderators” of performance. In the context of police agencies, such moderators might include the extent to which police officers are trained properly on how to secure a crime scene, package evidence, or interpret the results of forensic analyses. In the context of crime labs, such moderators might include staff morale and turnover, certifications among analysts, or the quality of the work environment (Dale and Becker 2004; Smith 1976). In each instance, however, it is vital to remember that these moderators of performance do not constitute outcomes. The performance measurement literature does draw a distinction between *primary* (or “end”) and *secondary* (or “intermediate”) outcomes. Primary outcomes are the *raison d’être* of an organization—its overall mission and goals. Secondary outcomes are not based directly on the organization’s mission and goals, but they are still valuable in and of themselves. Examples of secondary outcomes might include job satisfaction, employee health and safety, or adoption of measures to protect the environment. It may be valuable to measure secondary outcomes either because they are intrinsically valuable to the organization or because they are moderators of primary outcomes.

Discussion & Conclusion

The forensic evidence processing system in the United States is fragmented and heterogeneous. Cities, counties, and states have a patchwork set of interagency and intergovernmental agreements in place for processing forensic evidence. Moreover, some of this work is outsourced to private contractors. The work of this system is done by multiple agencies and by individuals with widely varying degrees of competence, training, education, experience, and professional credentials. The complexity inherent in the system makes it difficult to devise and implement outcome-based performance measures.

As a result, the industry relies on other quality control mechanisms—declared proficiency tests, accreditation, certification, and development of guidelines and standards—to improve performance.

When performance measures are used, they tend to be dominated heavily by measures of inputs, processes, and outputs, not *outcomes*. For instance, one state crime lab system has adopted seven performance measures, only one of which includes an outcome: the percent of cases more than 30 days old. The remaining measures may be meaningful to lab managers or budgetary analysts, but they are not direct measures of performance outcomes. For instance, one of the performance measures is the number of cases, a classic measure of *input*. This measure of workload may be useful in the numerator or denominator of ratios intended to measure efficiency (such as the number of cases completed per analyst), but alone it reveals nothing about a lab's performance. Other indicators focus on the number of new positions added, the percent of obsolete equipment replaced, and the proportion of employees attending training sessions. These are all measures of inputs and processes, but they are not outcome-based performance measures. They may be *related* to performance but they do not, by themselves, constitute performance. Unfortunately, measures of inputs, processes, and outputs represent the standard approach in the industry for measuring the performance of forensic processing systems.

Our proposed performance measures can be implemented at various levels and in various segments of forensic evidence processing systems. Organizations and subunits within organizations could measure some of these indicators. For example, a police agency that processes crime scenes and submits and stores physical evidence (but does not perform forensic analyses) could assess its ability to secure and process crime scenes; submit useful, well-preserved, and properly packaged physical evidence to a crime lab; and store evidence securely before trial. Likewise, sections within a crime lab, such as toxicology, could assess their performance in processing samples quickly and accurately. Governments at the city, county, or state levels could use these performance measures. Government agencies, accreditation bodies, or professional associations may be in the best position to assess performance across multiple organizations in one (or more) jurisdiction(s), such as the police, coroners, or medical examiners, crime labs, and prosecutors. These performance indicators could be used to establish benchmarks for various components of the system. These benchmarks could, in turn, be used to identify low-performing outliers that could then receive remediation. Or benchmarks could be used to identify high-performing units and organizations so the high performers could be rewarded appropriately and serve as a resource for other organizations. Researchers and policy makers could use these performance indicators to identify the causes and corre-

lates of strong (or weak) performance, and to assess the impact of changes in policy, training, budgets and resources, and other procedures. Once these performance indicators have been tested and validated, they could also be integrated into accreditation procedures, a step already taken in the health care and education sectors (Braun, Koss, and Loeb 1999; Elmore, Abelmann, and Fuhrman 1996).

Several complexities and constraints will make it challenging to implement outcome-based performance measurement in forensic systems. Yet many public agencies struggle with these same kinds of issues. Performance measurement in the public sector is a vital tool for determining whether the public is getting the high-quality service that it deserves. Agencies and systems that resist developing their own performance measures run the risk of having such measures imposed upon them externally. Good managers demand high-quality performance measures. Weak or shortsighted managers fear performance measures. The forensic system is in a time of crisis, and results-based performance measurement is one potential tool for diagnosing and improving performance.

References

- Allen, S., S. Case, and C. Frederick. 2000. Survey of forensic science laboratories by the Technical Working Group for Fire and Explosions (TWGFEX). *Forensic Science Communications* 2(1): <http://www.fbi.gov/hq/lab/fsc/backissu/jan2000/allen.htm> (accessed October 2, 2009).
- Braun, B. I., R. G. Koss, and J. M. Loeb. 1999. Integrating performance measure data into the joint commission accreditation process. *Evaluation and the Health Professions* 22: 283–297.
- Chen, J., S. S. Rathore, M. J. Radford, and H. M. Krumholz. 2003. JCAHO accreditation and quality of care for acute myocardial infarction. *Health Affairs* 22: 243–254.
- Childs, R., T. Witt, and K. Nurtegin. 2009. Survey of forensic science providers. *Forensic Science Policy and Management* 1: 49–56.
- Collins, J. and J. Jarvis. 2009. Wrongful conviction of forensic science. *Forensic Science Policy and Management* 1: 17–31.
- Committee on Identifying the Needs of the Forensic Science Community. 2009. *Strengthening forensic science in the United States: A path forward*. Washington, DC: The National Academies Press.
- Dale, W. M. and W. Becker. 2004. A case study of forensic scientist turnover. *Forensic Science Communications* 6(3): http://www.fbi.gov/hq/lab/fsc/backissu/july2004/research/2004_03_research04.htm (accessed October 2, 2009).
- Dibdin, J. 2001. Continuous quality improvement as a management concept for death investigation systems. *Journal of Forensic Sciences* 46(1): 94–97.
- Difonzo, J. H. 2005. The crimes of crime labs. *Hofstra Law Review* 34: 1–11.
- Durose, M. 2008. *Census of publicly funded forensic crime laboratories, 2005*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.
- Elmore, R. F., C. H. Abelmann, and S. H. Fuhrman. 1996. The new accountability in state education reform: From process to

- performance. In *Holding schools accountable: Performance-based reform in education*, H. F. Ladd, ed., chapter 3, pp. 65–98. Washington, DC: The Brookings Institution.
- Ferak, J. 2009. CSI director to stand trial. *Omaha World-Herald*, July 23. <http://www.omaha.com/article/20090723/NEWS01/707239839> (accessed August 6, 2009).
- Field, K. S. 1976. Quality assurance through proficiency testing and quality control programs. In *Crime laboratory management forum 1976*, R. H. Fox, and F. H. Wynbrandt, eds. Rockville, MD: The Forensic Sciences Foundation Press.
- Forensic Specialties Accreditation Board. 2009. Accredited boards. <http://www.thefsab.org/accredited.html> (accessed August 3, 2009).
- Gormley, W. T. Jr. and D. L. Weimer. 1999. *Organizational report cards*. Cambridge, MA: Harvard University Press.
- Hadley, K. and M. J. Fereday. 2008. *Ensuring competent performance in forensic practice*. New York: CRC Press.
- Hatry, H. P. 1999. *Performance measurement: Getting results*. Washington, DC: Urban Institute Press.
- Hickman, M. J. and B. A. Reaves. 2001. *Local police departments 1999*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.
- Hickman, M. J. and B. A. Reaves. 2003. *Local police departments 2000*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.
- Houck, M. M., R. A. Riley, P. J. Speaker, and T. S. Witt. 2009. FORESIGHT: A business approach to improving forensic science service. *Forensic Science Policy and Management* 1: 85–95.
- Houston Police Department. 2009. Crime lab: Frequently asked questions. <http://www.houstontx.gov/police/crimelab.faq.htm> (accessed August 6, 2009).
- Hunter, George. 2008. Workers rip crime lab's closing. State cops' audit a grab for funding, lab biologist says, though city officials ordered the shutdown. *Detroit News*, October 1.
- Innocence Project. 2009. *Investigating forensic problems in the United States: How the federal government can strengthen oversight through the Coverdell grant program*. New York: Cardozo School of Law.
- King, W. R. and E. R. Maguire. 2009. Systems designed to fail: The case of forensic evidence processing. (in press).
- Koehler, J. J. 2008. Fingerprint error rates and proficiency tests: What they are and why they matter. *Hastings Law Journal* 59:1077–1100.
- Kusluski, Michael. 2008. Don't degrade Detroit's crime lab. *Detroit Free Press*, November 13.
- Lentini, J. J. 2003. What you don't know can hurt you: How do you know your lab has it right? *Fire and Arson Investigator* 53(3):.
- Lewis, M. M. 2003. *Moneyball: The art of winning an unfair game*. New York: W.W. Norton & Co.
- Maguire, E. R. 2005. Measuring the performance of police agencies. *Law Enforcement Executive Forum* January: 1–30.
- Maguire, E. R., J. B. Snipes, C. D. Uchida, and M. Townsend. 1998. Counting cops: Estimating the number of police departments and police officers in the United States. *Policing: An International Journal of Police Strategies and Management* 21: 97–120.
- Miller, M. R., P. Pronovost, M. Donithan, S. Zeger, C. L. Zhan, Morlock, and G. S. Meyer. 2005. Relationship between performance measurement and accreditation: Implications for quality of care and patient safety. *American Journal of Medical Quality* 20: 239–252.
- Mohr, L. B. 1973. The concept of organizational goal. *The American Political Science Review* 67: 470–481.
- Moore, M. H. 2002. *Recognizing value in policing: The challenge of measuring police performance*. Washington, DC: Police Executive Research Forum.
- Murphy, P. V. and T. Plate. 1977. *Commissioner: A view from the top of American law enforcement*. New York: Simon and Schuster.
- Nelson, L. 1990. Management analysis: A mathematical model for determining the staffing requirements of a forensic science laboratory system. *Journal of Forensic Sciences* 35(5): 1173–1185.
- Pastore, A. L. and K. Maguire, eds. 2003. *Sourcebook of criminal justice statistics*. <http://www.albany.edu/sourcebook> (accessed September 15, 2006).
- Peterson, J. L. and M. J. Hickman. 2005. *Census of publicly funded forensic crime laboratories, 2002*. Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics.
- Peterson, J. L., G. Lin, M. Ho, Y. Chen, and R. E. Gaensslen. 2003. The feasibility of external blind DNA proficiency testing. *Journal of Forensic Sciences* 48: 21–31.
- Pisano, V. 1979. The organization and responsibilities of the Italian judicial police. *Journal of Forensic Sciences* 24(1): 221–226.
- Reaves, B. A. 2007. *Census of state and local law enforcement agencies, 2004*. Washington, DC: Bureau of Justice Statistics.
- Roman, J. K., S. Reid, J. A. Reid, Chalfin, W. Adams, and C. Knight. 2008. *The DNA field experiment: Cost-effectiveness analysis of the use of DNA in the investigation of high-volume crimes*. Washington, DC: Urban Institute.
- Saferstein, R. 1998. *Criminalistics: An introduction to forensic science*, 6th edition. Upper Saddle River, NJ: Prentice Hall.
- Schroeder, D. A. and M. D. White. 2009. Exploring the use of DNA evidence in homicide investigations: Implications for detective work and case clearance. *Police Quarterly* 12: 319–342.
- Simon, H. A. 1964. On the concept of organizational goal. *Administrative Science Quarterly* 9: 1–22.
- Smith, W. C. 1976. Overview of crime laboratory management. In *Crime laboratory management forum 1976*, R. H. Fox, and F. H. Wynbrandt, eds. Rockville, MD: The Forensic Sciences Foundation Press, pp. 5–21.
- Speaker, P. J. 2009a. Key performance indicators and managerial analysis for forensic laboratories. *Forensic Science Policy and Management* 1: 32–42.
- Speaker, P. J. 2009b. The decomposition of return on investment for forensic laboratories. *Forensic Science Policy and Management* 1: 96–102.
- Steadman, G. W. 2002. *Survey of DNA crime laboratories, 2001*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.
- Thompson, M., K. Mathieson, L. Owen, A. P. Damant, and R. Wood. 2009. The relationship between accreditation status and performance in a proficiency test. *Accreditation Quality Assurance* 14: 73–78.
- Tilley, N. and A. Ford. 1996. *Forensic science and crime investigation: Crime detection and prevention series paper 73*. London: Home Office Police Research Group.
- Wellford, C. and J. Cronin. 1999. *An analysis of variables affecting the clearance of homicides: A multistate study*. Washington, DC: Justice Research and Statistics Association.
- Williams, Corey. 2008. Detroit mayor, police chief close city's crime lab after errors found in evidence collection. *Chicago Tribune*, October 14.