# Ideas in American Policing

# Evidence-Based Policing

*By Lawrence W. Sherman*

## Abstract

The new paradigm of "evidence-based medicine" holds important implications for policing. It suggests that just doing research is not enough and that proactive efforts are required to push accumulated research evidence into practice through national and community guidelines. These guidelines can then focus in-house evaluations of what works best across agencies, units, victims, and officers. Statistical adjustments for the risk factors shaping crime can provide fair comparisons across police units, including national rankings of police agencies by their crime prevention effectiveness. The example of domestic violence, for which accumulated National Institute of Justice research could lead to evidence-based guidelines, illustrates the way in which agency-based outcomes research could further reduce violence against victims. National pressure to adopt this paradigm could come from agency-ranking studies, but police agency capacity to adopt it will require new data systems creating "medical charts" for crime victims, annual audits of crime reporting systems, and in-house "evidence cops" who document the ongoing patterns and effects of police practices in light of published and in-house research. These analyses can then be integrated into the NYPD Compstat feedback model for management accountability and continuous quality improvement.

*Most of us have thought of the statistician's work as that of measuring and predicting . . . but few of us have thought it the statistician's duty to try to bring about changes in the things that he [or she] measures.*
　　　　　　　　　—W. Edwards Deming

**Lawrence W. Sherman** is professor and chair of the Department of Criminology and Criminal Justice at the University of Maryland at College Park. He was the Police Foundation's director of research from 1979 to 1985.

Of all the ideas in policing, one stands out as the most powerful force for change: police practices should be based on scientific evidence about what works best. Early in this century, Berkeley Police Chief August Vollmer's partnership with his local university helped generate this idea (Carte and Carte 1975), which was clearly derived from that era's expansion of the scientific method into medicine, management, agriculture, and many other fields (Cheit 1975). While science had greater initial impact in those other professions during the first half of the century, policing in recent decades has been moving rapidly to catch up. However, any assessment of this idea in modern policing must begin with an accurate benchmark: catching up to what? More complete evidence on the linkage between research and practice suggests a new paradigm for police improvement and for public safety in general: evidence-based crime prevention.

For years, Sherman (1984, 1992) and others have used medicine as the exemplar of a profession based upon strong scientific evidence. Sherman has praised medicine as a field in which practitioners have advanced training in the scientific method and keep up-to-date with the most recent research evidence by reading medical journals. He has cited the large body of randomized controlled experiments in medicine—now estimated to number almost one million in print (Sackett and Rosenberg 1995)—as the highly rigorous scientific evidence used to guide medical practices. He has suggested that policing should therefore be more like medicine.

Sherman was right about the need for many more randomized experiments in policing, but wrong about how much medicine was really based on scientific research. New evidence shows that doctors resist changing practices based on new research just as much as police do, if not more so. Closer examination reveals medicine to be a battleground between research and practice, with useful lessons for policing on new ways to promote research. Those lessons come from a new strategy called "evidence-based medicine,"[1] "widely hailed as the long-sought link between research and practice" (Zuger 1997) to solve problems like the following (Millenson 1997, 4, 122, 131):

- An estimated 85 percent of medical practices remain untested by research evidence.

- Most doctors rarely read the 2,500 medical journals available, and instead base their practice on local custom.

- Most studies that do guide practice use weak, non-randomized research designs.

---

[1] The term "evidence" in this monograph refers to scientific, not criminal, evidence.

Medicine, in fact, seems just as resistant to the use of evidence to guide practice as are fields with lower educational requirements, such as policing. The National Institutes of Health (NIH) Consensus Guidelines are a case in point. NIH convenes advisory boards to issue to physicians recommendations that are based on intensive reviews of research evidence on specific medical practices. These recommendations usually receive extensive publicity, and are reinforced by mailings of the guideline summaries to some one hundred thousand doctors. But according to a RAND evaluation, doctors rarely change their practices in response to publication of these guidelines (Kosecoff et al. 1987, as cited in Millenson 1997). Thus three years after research found that heart attack patients treated with calcium antagonists were more likely to die, doctors still prescribed this dangerous drug to one-third of heart attack patients. Eight years after antibiotics were shown to cure ulcers, 90 percent of ulcer patients remained untreated by antibiotics (Millenson 1997, 123–25).

## Evidence Cops

The struggle to change medical practice based on research evidence has a long history, with valuable implications for policing. In the 1840s, Ignaz Semmelweiss found evidence that maternal death in childbirth could be reduced if doctors

washed their hands before delivering babies. He then tried to apply this research to medical practice in Vienna, which led to his being driven out of town by his boss, the chief obstetrician. Hundreds of thousands of women died because the profession refused to comply with his evidence-based guidelines for some forty years. The story shows the important distinction between merely doing research and attempting to apply research to redirect professional practices.

One way to describe people who try to apply research is the role of "evidence cop." More like a traffic cop than Victor Hugo's detective Javert, the evidence cop's job is to redirect practice through compliance rather than punishment. While this job may be as challenging as herding cats, it still consists of pointing professionals to practice "this way, not that way." As in all policing, the success rate for this job varies widely. Fortunately, the initial failures of people like Semmelweiss paved the way for greater success in the 1990s.

Consider Scott Weingarten, M.D., of Cedars-Sinai Hospital in Los Angeles. As director of the hospital's Center for Applied Health Services Research, Weingarten is an evidence-cop-in-residence. His job is to monitor what the 2,250 doctors are doing to patients at the hospital and to detect practices that run counter to recommendations based on research evidence. He does this through prodding rather than punishment, convening groups of doctors who treat specific maladies to discuss the research evidence. These groups then produce their own consensus guidelines for practices that become hospital policy. Thirty-five such sets of guidelines were produced in Weingarten's first four years on the job (Millenson 1997, 120).

What NIH, Weingarten, and the 1995 founders of the new journal called *Evidence-Based Medicine* are all trying to do is to push research into practice. Just as policing has become more proactive at dealing with crime, researchers are becoming more proactive about dealing with practice. This trend has developed in many fields, not just medicine.

Increased pressure for "reinventing government" to focus on measurable results is reflected in the 1994 U.S. Government Performance Results Act (GPRA), which requires all federal agencies to file annual reports on quantitative indicators of their achievements. Education is under growing pressure to raise test scores as proof that children are learning, which has led to increased discussion of research evidence on what works in education (Raspberry 1998). And the U.S. Congress has required that the effectiveness of federally funded crime prevention programs be evaluated using "rigorous and scientifically recognized standards and methodologies" (House 1995, sec. 116). All this sets the stage for a new paradigm for making research more useful to policing than it has ever been before.

## Key Questions

In suggesting a new paradigm called evidence-based policing, there are four key questions to answer: What is it? What is new about it? How does it apply to a specific example of police practice? How can it be institutionalized?

### What is it?

Evidence-based policing is the use of the best available research on the outcomes of police work to implement guidelines and evaluate agencies, units, and officers. Put more simply,

evidence-based policing uses research to guide practice and evaluate practitioners. It uses the best evidence to shape the best practice. It is a systematic effort to parse out and codify unsystematic "experience" as the basis for police work, refining it by ongoing systematic testing of hypotheses.

Evaluation of ongoing operations has been the crucial missing link in many recent attempts to improve policing. If it is true that most police work has yet to go "beyond 911" (Sparrow, Moore, and Kennedy 1990), the underlying reason may be a lack of evaluation systems that clearly link research-based guidelines to outcomes. It is only with that addition that policing can become a "reflexive" or "smart" institution, continuously improving with ongoing feedback.

The basic premise of evidence-based practice is that we are all entitled to our own opinions, but not to our own facts. Yet left alone to practice individually, practitioners do come up with their own "facts," which often turn out to be wrong. A recent survey of 82 Washington State doctors found 137 different strategies for treating urinary tract infections (Berg 1991). No doubt the same result could be found for handling domestic disturbances. A study evaluating the accuracy of strep throat diagnoses based on unstructured examination by experienced pediatricians found it

far inferior to a systematic, evidence-based checklist used by nurses. The mythic power of subjective and unstructured wisdom holds back every field and keeps it from systematically discovering and implementing what works best in repeated tasks.

A prime example of the power of systematic, ongoing evaluations comes again from medicine. In 1990, the New York State Health Department began to publish death rates for coronary bypass surgery grouped by hospital and individual surgeon. This action was prompted by research showing that while the statewide average death rate was 3.7 percent, some doctors ran as high as 82 percent. Moreover, after adjusting for the risk of death by the pre-operation condition of the patient caseload, patients were 4.4 times more likely to die in surgery at the least successful hospitals than at the best hospitals. Despite enormous opposition from hospitals and surgeons, these data were made

public, revealing a strong practice effect: the more operations doctors and hospitals did each year, the lower the risk-adjusted death rate. Using this clear correlation to push low-frequency surgeons and hospitals out of this business altogether, hospitals were able to lower the death rate in these operations by 40 percent in just three years (Millenson 1997, 195).

Evidence-based policing is about two very different kinds of research: basic research on what works best when implemented properly under controlled conditions, and ongoing outcomes research about the results each unit is actually achieving by applying (or ignoring) basic research in practice. This combination creates a feedback loop (fig. 1) that begins with either published or in-house studies suggesting how policing might obtain the best effects. The review of this evidence can lead to guidelines taking law, ethics, and community



*Figure 1. Evidence-Based Policing.*
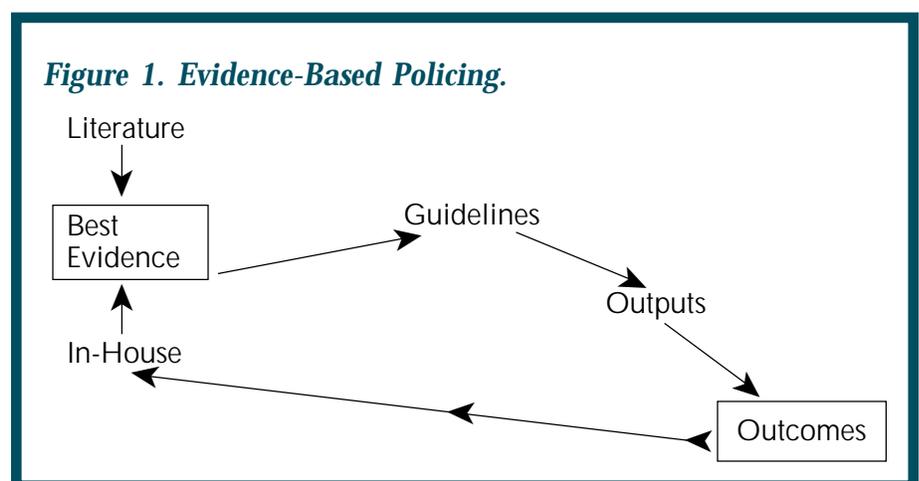
culture into account. These guidelines would specify measurable "outputs," or practices that police are asked to follow. Their varying degrees of success at delivering those outputs can then be assessed by tracking risk-adjusted "outcomes," or results over a reasonably long follow-up period. These outcomes may be defined in several different ways: offenses per 1,000 residents, repeat victimizations per 100 victims, repeat offending per 100 offenders, and so on. The observation that some units are getting better results than others can be used to further identify factors associated with success, which can then be fed back as new in-house research to refine the guidelines and raise the overall success level of the agency. Such research could also be published in national journals or at least kept in an agency database as institutional memory about success and failure rates for different methods.

### What is new about it?

Skeptics may say that there is nothing new in evidence-based policing, and that other paradigms already embrace these principles. On closer examination, however, we will see that no other paradigm contains the principles for its own implementation. No other paradigm contains a principle for both changing practices and measuring the success of those changes with risk-adjusted

**Evidence-based policing is clearly different from, but very helpful to, all three present paradigms of policing.**

outcomes research (like bypass surgery death rates). No other paradigm—not even NYPD's Computerized Crime Comparison Statistics (Compstat) strategy (Bratton with Knobler 1998)— uses scientific evidence to hold professionals accountable for results in peer-reviewed and even public discussions of outcomes evidence.

Evidence-based policing is clearly different from, but very helpful to, all three present paradigms of policing. Incident-specific policing, or 911 responses, currently lack any outcomes measure except time out of service. Police officers who take too much time to handle a call are sometimes accused of shirking and are urged by supervisors to work faster.[2] But no one tracks the rate of repeat calls by officer or unit to see how effective the first response was in preventing future problems. Evidence-based policing could

use such outcomes to justify longer time spent on each call on the basis of an officer's average results, rather than issuing a crude demand that he or she stay within an average time limit. It could also place much more emphasis on learning how to deal with each call most effectively and preventively, a question that currently gets little attention.

Community policing, however defined, is not clearly linked to evidence about effectiveness in preventing crime. It is much more about how to do police work—a set of outputs— than it is about desired results, or outcomes. Working with the community and listening to and respecting community members are all important elements of the paradigm. But that paradigm alone has been easy for many officers to ignore. Adding the accountability systems from the paradigm of evidence-based policing could actually make police far more active in working with the community.

Problem-oriented policing is clearly the major source for

---

[2] This sounds oddly like the pressure for drive-in, drive-out childbirth health insurance now barred by federal law.

evidence-based policing. Herman Goldstein's writings (1979, 1990), as well as John Eck and William Spelman's SARA model (1987), clearly emphasize assessment of problem-solving responses as a key part of the process. Yet there is no clear statement about the use of scientific evidence either in selecting strategies for responding to problems or in monitoring the implementation and results of those strategies (Sherman 1991). Reports on problem-oriented policing have so far produced little evidence either from controlled tests or outcomes research. Because the paradigm stresses the unique characteristics of each crime pattern, problem-oriented policing has not been used to respond to highly repetitive situations like domestic assaults or disputes. Few comparisons of different methods for attacking the same problem have been developed. Few officers are even held accountable for not implementing a problem-solving plan they have agreed to undertake. Problem-oriented policing has clearly revolutionized the way many police think about their objectives, moving them away from a narrow focus on each incident to a broader focus on patterns and systems. But in the absence of pressure from an evidence-based approach to evaluating success and management accountability, problem-oriented policing has been kept at the margins of police work.

NYPD's Compstat strategy (Bratton with Knobler 1998) has pushed the results accountability principle farther than ever before, but it has not used the scientific method to assess cause and effect. Successful managers are rewarded, but successful methods are not pinpointed and codified.

What evidence-based policing adds to these paradigms is a new principle for decision making: scientific evidence. Most police practice, like medical practice, is still shaped by local custom, opinions, theories, and subjective impressions. Evidence-based policing challenges those principles of decision making and creates systematic feedback to provide continuous quality improvement in the achievement of police objectives (see Hoover 1996). Hence the inspiration for this paradigm is not only medicine and its randomized trials, but also the principles of quality control in manufacturing developed by Walter Shewhart (1939) and W. Edwards Deming (1986). These principles were initially rejected by U.S. business leaders, but were finally embraced in the 1980s after Japanese industries used them to far surpass U.S. manufacturers in the quality of their products.

What makes both policing and medicine different from manufacturing, of course, is the far greater variability in the raw material to be processed—human beings. That is what gives the gold standard of evaluation research, the randomized

controlled trial, both its strength and its limitations. The strength of the research design, pioneered in policing by the Police Foundation, is its ability to reduce uncertainty about the average effects of a policy on vast numbers of people. The limitation of the research design is that it cannot escape variability in treatments, responses, and implementation.

The variability of treatments in policing is much like that in surgery, which stands in sharp contrast to pharmaceuticals. While the chemical content of medical drugs is almost always identical, the procedural content of surgery varies widely. Similarly, the style and tone each officer brings to a citizen encounter varies enormously and can make a big difference in the outcome of a specific case. Dosage, timing, and follow-up of both drugs and police work can vary widely in practice.

Even holding treatment constant, there is evidence that both patients and offenders respond to treatments with wide variations. Some of these responses, allergic reactions, can kill some people with treatments that cure most others. Offenders are known to vary in their responses to police actions by individual, neighborhood, and city. And implementation of new practices based on controlled experiments in both medicine and policing varies according to how well research is communicated, how much information is created

about whether practices actually change, and how much reinforcement there is for the change, both positive and negative.

Evidence-based policing assumes that experiments alone are not enough. Putting research into practice requires just as much attention to implementation as it does to controlled evaluations. Ongoing systems for researching implementation can close the feedback loop to create the principle of industrial quality improvement.

### How does it apply to a specific example of police practice?

The policing of domestic violence offers a clear illustration of what is new about the evidence-based paradigm. Domestic violence has been the subject of more police practices research than any other crime problem. The research has arguably had little effect on police practice, at least by the new standards of evidence-based medicine. Yet the available evidence offers a fair and scientifically valid approach for holding police agencies, units, and officers accountable for the results of police work, as measured by repeated domestic violence against the same victims.

The National Institute of Justice (NIJ) and the Police Foundation have provided policing with extensive information on what works to prevent repeated violence. The

research has also shown that, like surgery, police practices vary greatly in their implementation. These variations in practice cause varying results for repeat offending against victims. Even holding practice constant, responses to arrest vary by offender, neighborhood, and city. Finally, research shows very poor compliance with mandatory arrest guidelines after they are adopted (Ferraro 1989).

There are many varieties of arrest for misdemeanor domestic violence. The offender may or may not be handcuffed, arrested in front of family and neighbors, given a chance to explain his version of events to the police, or

treated with courtesy and politeness. Do these variations on the theme of arrest make a difference? They should, according to the "defiance" theory of criminal sanction effects (Sherman 1993). And they did in Milwaukee, according to Raymond Paternoster and his colleagues (1997). The Milwaukee evidence reveals that controlling for other risk factors among some 800 arrested offenders, those who felt they were not treated in a procedurally fair and polite manner were 60 percent more likely to commit a reported act of domestic violence in the future (fig. 2). This finding suggests three ways

### Figure 2. Repeat Domestic Violence and Police Fairness.

Source: Paternoster, et al.
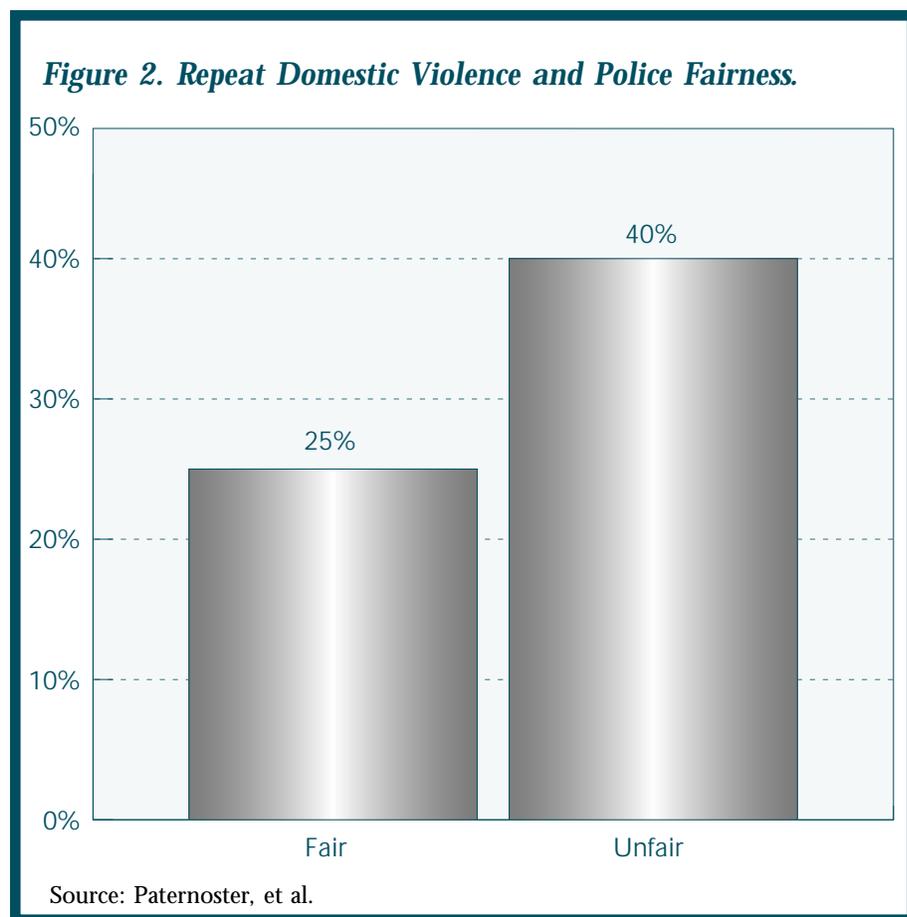
to push research into practice: 1) change the guidelines for making domestic violence arrests to include those elements that would enable offenders to perceive more "procedural justice"; 2) hold police accountable for using these guidelines by comparing rates of repeat victimization associated with different police units; and 3) compute these rates using statistical adjustments for the pre-existing level of recidivism risks.

The NIJ research provides other evidence for ways that police can reduce repeat offending in misdemeanor domestic violence. Rather than a one-size-fits-all policy, the evidence suggests specific guidelines to be used under different conditions. Offenders who are absent when police arrive—as they are in some 40 percent of cases—respond more effectively to arrest warrants than offenders who are arrested on the scene (Dunford 1990). Offenders who are employed are deterred by arrest, while offenders who are unemployed generally increase their offending more if they are arrested than if they are handled in some other fashion (Pate and Hamilton 1992; Berk et al. 1992; Sherman and Smith 1992). Offenders who live in urban areas of concentrated poverty commit more repeat offenses if they are arrested than if not, while offenders who live in more affluent areas commit fewer repeat offenses if they are arrested (Marciniak 1994). All of these

findings could be changed by further research, but for the moment they are the best evidence available.

This research evidence could support guidelines for policing domestic violence that differed by neighborhood and absence or presence of the offender. It could also support guidelines about listening to suspects' side of the story before making arrest decisions and generally treating suspects with courtesy. Other evidence, such as the extremely high-risk period for repeat victimization in the first days and weeks after the last police encounter (Strang and Sherman

1996), could be used to fashion new problem-oriented strategies. Most important, the existing research can be used to create a fair system for evaluating police performance on the basis of risk-adjusted outcomes. That evidence (fig. 3) shows that the likelihood of a repeat offense is strongly linked to the number of previous offenses each offender has.

Once the risk of repeat offending can be predicted with reasonable accuracy, it becomes possible to use those predictions as a benchmark for police performance. Just as in the bypass surgery death rates in New York, the outcomes of policing can be
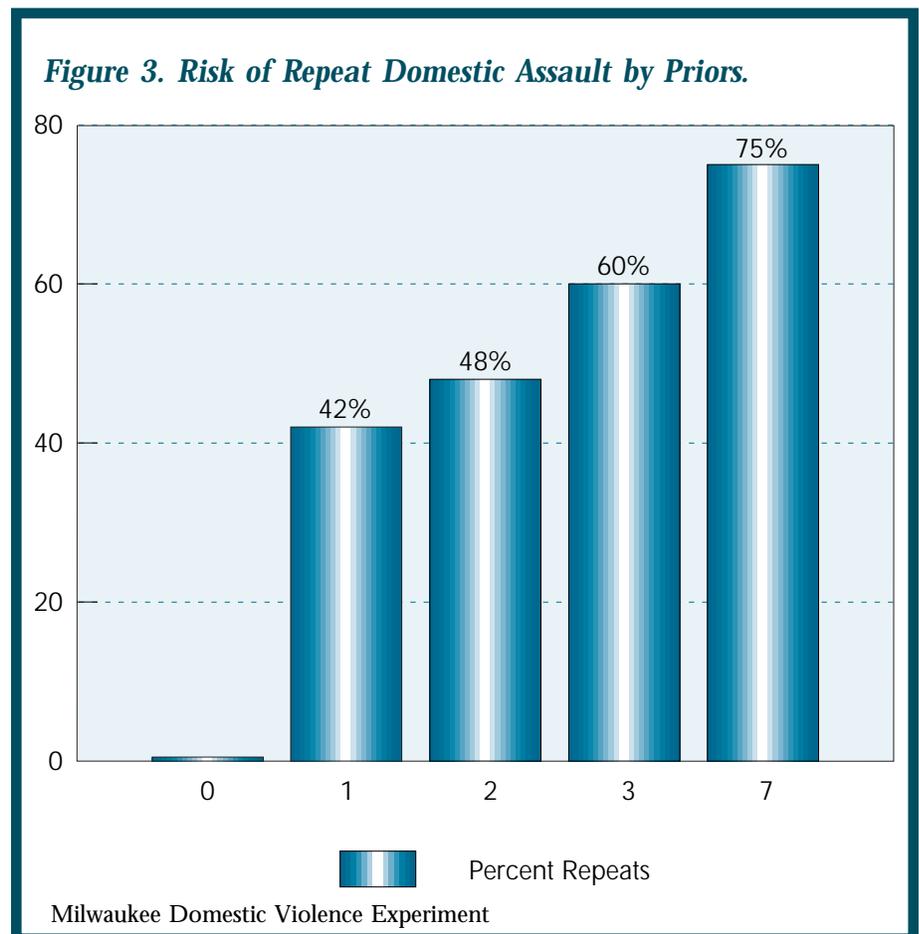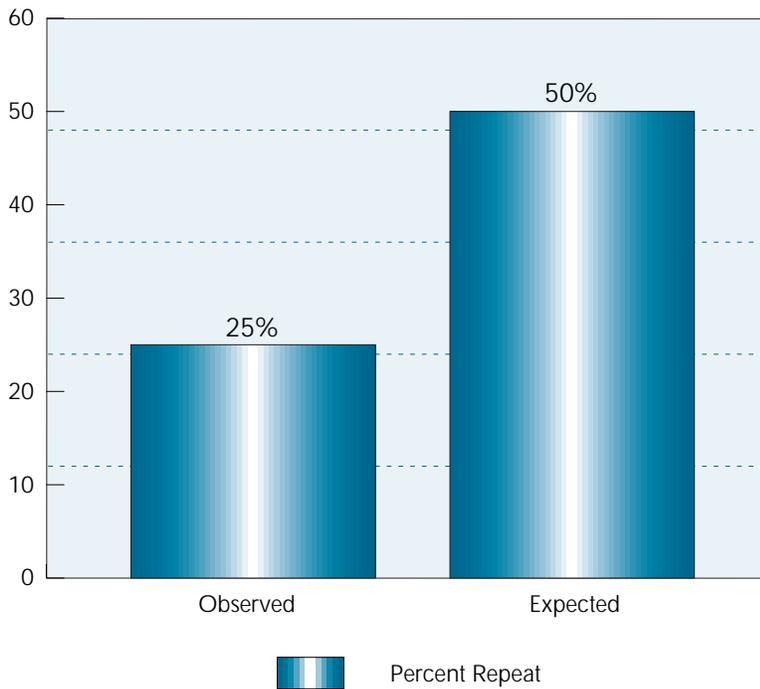
**Figure 3. Risk of Repeat Domestic Assault by Priors.**

Percent Repeats

Milwaukee Domestic Violence Experiment

**Figure 4. Observed vs. Expected Risk of Repeat Domestic Violence.**



By constructing information systems for this kind of outcome research, police departments can focus on an objective that has only previously been measured in major experiments. Making the goal of policing each domestic assault the *outcome* of a reduced repeat offending rate rather than the *output* of whether an arrest is made would have several effects. One is that crime prevention would get greater attention than retribution for its own sake. While not everyone would welcome that, it is consistent with at least some police leaders' view of the purpose of the police as a crime prevention agency (Bratton with Knobler 1998). Another effect would be to seek out and

controlled for the risk level inherent in the caseload they face. Using a citywide database of all domestic assaults, now running over ten thousand cases per year in cities like Milwaukee, a model can be constructed to assess the risk of repeat offending in each case. The overall mix of cases in each police precinct or for each officer can generate an average risk level for that caseload. Each police patrol district can then be evaluated according to the actual versus predicted rate of repeat offending each year (fig. 4). All patrol districts in the city can then be compared on the basis of their relative percentage difference between expected and actual rates of repeat domestic assault (fig. 5).

**Figure 5. Observed vs. Expected Ranking by Precinct.**
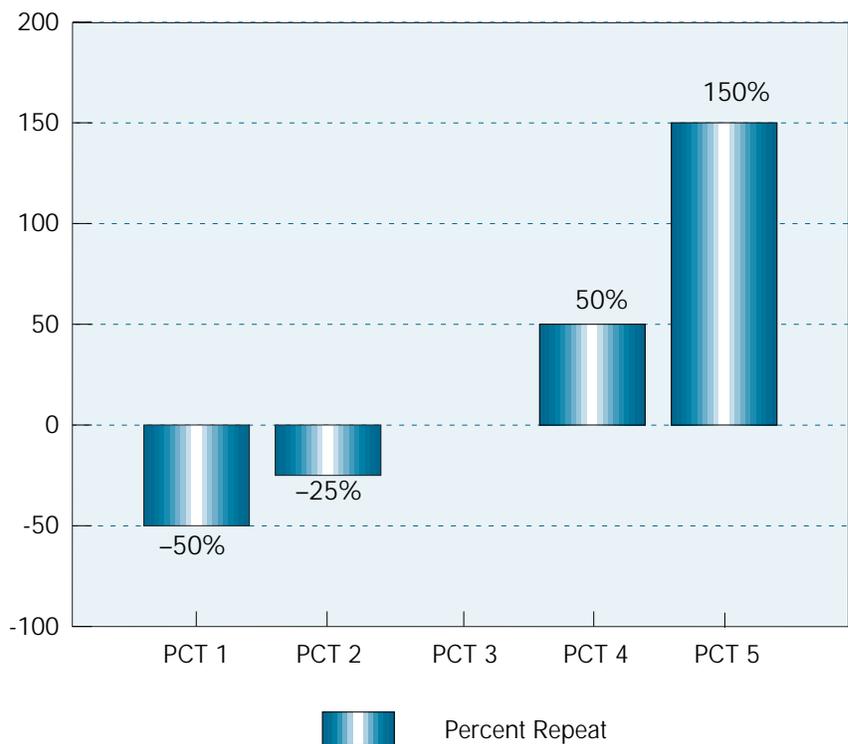
> *The strongest claim about evidence-based policing is that it contains the principles of its own implementation.*

even initiate more research on what works best to prevent domestic violence. In the world as we now know it, no one in policing—from the police chief to the rookie officer—has any direct incentive to reduce repeat offending against known victims. No one in policing is held accountable for accomplishing, or even measuring, that objective. As a result, no one knows whether repeat victimization rates get better or worse from year to year. Using outcomes evidence to evaluate performance would make police practices far more victim-centered, the top priority being that of preventing any further assaults.

### How can it be institutionalized?

The strongest claim about evidence-based policing is that it contains the principles of its own implementation. The principles of using evidence both to change and evaluate practice can be applied to a broad institutional analysis of implementation. Thus while the changes described

above would have to occur one police agency at a time, there are certain national forces that can help start the ball rolling. This can be seen, for example, in national rankings of big-city police agencies, as well as national mandates for improving police data systems to provide better evidence. Yet even such external pressures will not succeed without internal evidence cops to import, apply, and create research evidence.

No institution is likely to increase voluntarily its accountability except under strong external pressure. It is unlikely that evidence-based policing could be adopted by a police executive simply because it appears to be a good idea. The history of evidence-based medicine and education strongly suggests that professionals will only make such changes under external coercion. Nothing seems to foster such pressure as much as performance rankings across agencies (Millenson 1997; Steinberg 1998). Just as various public performance measures

allow stockbrokers to rank publicly-held corporations and provide those companies with strong incentives for better results, public information about police performance would create the strongest pressure for improvement.[3]

One example of how the major city police departments could be ranked on performance can be found in their homicide rates, which already receive extensive publicity. What these statistics lack, however, is any scientific analysis of expected risk. Police performance has nothing to do, at least in the short run, with the social, economic, demographic, and drug market forces that help shape a city's homicide rate. While police performance may also affect those homicide rates, the other factors must be taken into account. Using risk-adjusted homicide rates provides one indication of how well a police department may be doing things like confiscating illegal weapons, patrolling hot spots, regulating violent taverns and drug markets, and monitoring youth gangs. While the basic research literature would increasingly provide a source of guidance for taking initiatives against homicide, a

---

[3] The 1919 results of the first national rankings of hospitals were deemed so threatening that the American College of Surgeons decided to burn the report immediately in the furnace of New York's Waldorf-Astoria Hotel (Millenson 1997, 146).

risk-adjusted outcomes analysis (fig. 6) would indicate how well that research had been put into practice.[4]

If a credible national research organization would produce such "league rankings" among big-city police departments each year (like the *U.S. News & World Report* rankings of colleges and universities), the predictable result in the short term would be attacks on the methodology used. That is, in fact, what continues to go on in New York with the death rates in surgery. But the New York rankings have spread to other states, and consumers have found them quite valuable. Doctors—and police—may also find rankings very valuable in the long run. Both professions should enjoy greater public respect as they get better at producing the results their consumers want.
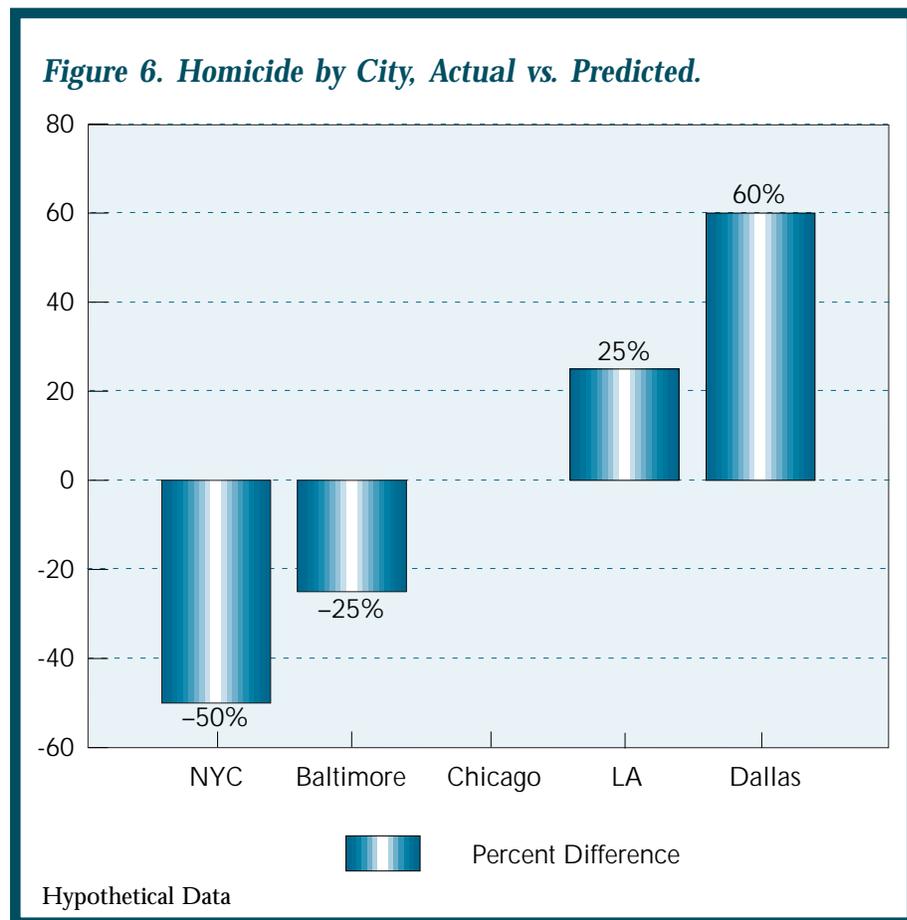
The more seriously performance indicators influence the fate of organizations, the more likely they are to be subverted. Recent examples include the U.S. Postal Service in West Virginia, where an elaborate scheme to defeat the on-time mail delivery audit was recently

alleged (McAllister 1998). Other examples include teachers helping students to cheat on their answers to national achievement tests and, of course, police departments under-reporting crime. The New York City police have removed three commanders in the past five years for improperly counting crime to make their performance look better (Kocieniewski 1998), and several chiefs of police elsewhere have been convicted on criminal charges for similar conduct.

Quite apart from pressures to corrupt data, criminologists have long known that police crime reporting is not reliable, with the possible exception of homicide. No two agencies classify crime the same way. The same event may be called an aggravated assault in one agency and a "miscellaneous incident" in another. The recent FBI decision to drop Philadelphia from the national crime reporting program was not an isolated action. In 1988, the FBI quietly dropped the entire states of Florida and Kentucky. Since the FBI lacks resources to do on-site audits in each police agency every year, these examples are just the tip of a very big iceberg. There are already rising suspicions of police manipulation of crime data as

---

[4] While many of the basic risk factors would be computed from Census data that could be out of date by the middle of each decade, other risk data can be derived from annually updated sources, such as the NIJ ADAM data on drug abuse among arrestees. Unemployment, school dropout, teen childbirth, and infant mortality data are also available annually for each city and could help predict the expected rate of homicide.



**Figure 6. Homicide by City, Actual vs. Predicted.**

Hypothetical Data

Percent Difference

NYC: −50%
Baltimore: −25%
LA: 25%
Dallas: 60%

crime rates fall in many cities. More serious pressure from national rankings would threaten data integrity even more.

One viable solution to this problem is a federal requirement for police departments to retain CPA firms to produce annual audits of their reported crime data. This requirement could be imposed as a condition for receiving federal funds, just as many other federal mandates have already done. Anticipating court challenges about unfunded mandates (such as the Brady Bill), Congress could also provide funds to pay for the audits. Crime counting standards could be set nationally by the accounting profession in collaboration with the FBI. Alternatively, each state legislature could require (or even fund) these audits as a means of assuring fairness in performance rankings of police departments within the state. State agencies such as the criminal justice statistical centers could also produce such rankings as a service to taxpayers. States already have the option of spending federal funds on such a purpose under the broad category of evaluation funds.

In the process of revitalizing crime data integrity, there would be great value in reorganizing police data systems. Most important would be the creation of a "medical chart" for each crime victim. Like computerized patient records, this chart would show the diagnosis (offense

*In the process of revitalizing crime data integrity, there would be great value in reorganizing police data systems.*

description) for each incident a victim presents to a police agency, perhaps anywhere in the state. The chart would also show what police did in response, everything from taking an offense report to arresting an offender whose release date from prison is also kept, updated, in the computerized victim chart. This information tool could help develop many proactive police methods for preventing repeat victimization. Allowing officers to use these data to keep their own private "batting averages" for repeat victimization (even without adjusting for risk) may encourage them to become involved and committed to doing a better job at preventing crime.

Better records are also needed about what police do about crime according to certain patterns of offenses. "Medical charts" for violent taverns, frequently robbed convenience stores, and other hot spots where most crime occurs would be very useful for ongoing problem-oriented policing

attempts to reduce repeat offending at those places. Similar records could be kept about a pattern of crimes spread out across a wider area, such as automatic teller machine robberies. If officer teams or units identify these places or patterns as crime targets and designate a control group, these medical charts can become the basis for estimating how much crime each police unit has prevented.

Computers can also help police officers to implement practice guidelines. Medical computer systems now offer recommended practice guidelines in response to a checklist of data, as well as warning when drug prescriptions fall outside programmed parameters of disease type and dosage. The use of hand-held computers to advise officers in the field and to provide instant quality control checks may not happen soon, but the growth of police research may make it inevitable in the long run. Doctors are not expected to keep

large amounts of research data in their heads, nor even medical guidelines for each diagnosis. Computers will not replace good judgment, but they can clearly enhance it.

Federal rules could also require police departments to appoint a certified police criminologist (either internally or in partnership with a university or research organization), who would become the agency's evidence cop. Like Scott Weingarten of Cedars-Sinai, the departmental criminologist would be responsible for putting research into practice, then evaluating the results. Whether the criminologist is actually an employee or a university professor working in partnership with the police may not matter as much as the role itself. The criminologist could help develop more effective guidelines for preventing repeat offending, and could develop expected versus actual repeat offending data by offense type for each police district or detective unit. A criminologist could add the scientific method to the NYPD Compstat process (Bratton with Knobler 1998), providing statistics at each meeting on each patrol district's crime trends and patterns (or even its complaints against police officers) in relation to the district's risk level. Building the capacity to import, apply, and create evidence within each police agency may be an essential ingredient in the success of this paradigm.

We may also find that the traditional distance between researchers and police officials shrinks when researchers provide more immediate managerial information. Criminologists have long refused to provide police managers with data on particular officers, deeming it contrary to the ethics of basic research (Hartnett 1998). By finally providing the data in a scientifically reasonable format, criminologists may become far more effective at pushing research into practice.

Criminologists can also act on the finding that doctors tend to change practices based on personal interaction and repeated computerized feedback, and not from conferences, classes, or written research reports (Millenson 1997, 127–30). Similar findings have been published about the effectiveness of agricultural extension services,

*The empirical question for research is, what practices work best to change practices?*

in which university scientists visit farms and show farmers new techniques for improving their crop yields. They echo a Chinese proverb: Tell me and I will forget; show me and I will remember; involve me and I will understand.

The one test of this principle in policing to date is Alex Weiss's (1997) research on how police departments adopt innovations. Based on a national survey of police chiefs and their top aides, Weiss discovered that telephone calls from agency to agency played a vital role in spreading new ideas. While written reports may have supplemented the phone calls, word-of-mouth seems to be the major way in which police innovations are communicated and adopted.

Weiss's study suggests the great importance of gathering more evidence on evidence. The empirical question for research is, what practices work best to change practices? This inherently reflexive posture may lead us to empirical comparisons of the effectiveness of, for example, NIJ conferences, mass mailings of research-in-brief reports, or new one-on-one approaches. One example of the latter would be proactive telephone calls to police agencies around the U.S. made by present or former police officers; callers could be trained by research organizations to describe new research findings. If national consensus guidelines for practice were developed by panels of police executives and

researchers, the callers could communicate those as well. Other approaches worth testing might include field demonstrations in police technique. This training would not be based on experience, as is the current Field Training Officer system, but rather it would be based on evidence that the method being demonstrated has been proven effective in reducing repeat offending.

## Conclusion

The test of this paradigm's results is not whether it is adopted this year or in twenty years. As Lord Keynes has suggested, the influence of ideas may be far more glacial than volcanic. The pressure for better measures of results is in the spirit of the age, and police cannot long escape it. All this paper does is add one inch to the glacier, so that we can say of policing what Dr. William Mayo of the Mayo Clinic said of his profession almost a century ago: "The glory of medicine is that it is constantly moving forward, that there is always something more to learn."

## References

Berg, A.O. 1991. Variations among family physicians' management strategies for lower urinary tract infections in women: A report from the Washington Physicians' Collaborative Research Network. *Journal of the American Board of Family Practice* (September–October): 327–30.

Berk, Richard A.; Alec Campbell; Ruth Klap; and Bruce Western. 1992. The deterrent effect of arrest in incidents of domestic violence: A Bayesian analysis of four field experiments. *American Sociological Review* 57: 698–708.

Bratton, William, with Peter Knobler. 1998. *Turnaround: How America's top cop reversed the crime epidemic.* New York: Random House.

Carte, Gene, and Elaine Carte. 1975. *Police reform in the United States: The era of August Vollmer.* Berkeley: University of California Press.

Cheit, Earl. 1975. *The useful arts and the liberal tradition.* New York: McGraw-Hill.

Deming, W. Edwards. 1986. *Out of the crisis.* Cambridge: Massachusetts Institute of Technology, Center for Advanced Engineering Study.

Dunford, Franklyn. 1990. System-initiated warrants for suspects of misdemeanor domestic assault: A pilot study. *Justice Quarterly* 7: 631–53.

Eck, John, and William Spelman. 1987. *Problem-solving: Problem-oriented policing in Newport News.* Washington, D.C.: Police Executive Research Forum.

Ferraro, Kathleen J. 1989. Policing woman battering. *Social Problems* 36: 61–74.

Goldstein, Herman. 1979. Improving policing: A problem-oriented approach. *Crime and Delinquency* 25: 236–58.

———. 1990. *Problem-oriented policing.* New York: McGraw-Hill.

Hartnett, Susan. 1998. Address to the Third National Institute of Justice Conference on Police-Research Partnerships, February.

Hodge, Melville H. 1990. Direct use by physicians of the TDS Medical Information System. In *A History of Medical Informatics*, edited by Bruce I. Blum and Karen Duncan. New York: ACM.

Kosecoff, Jacqueline, et al. 1987. Effect of the National Institutes of Health Consensus Development Program on physician practice. *Journal of the American Medical Association* 258 (November 20): 2708–13.

McAllister, Bill. 1998. A "special" delivery in West Virginia: Postal employees cheat to beat rating system. *Washington Post*, 10 January, A1.

Marciniak, Elizabeth. 1994. Community policing of domestic violence: Neighborhood differences in the effect of arrest. Ph.D. diss., University of Maryland.

Millenson, Michael L. 1997. *Demanding medical excellence: Doctors and accountability in*

*the information age.* Chicago: University of Chicago Press.

Office of Technology Assessment of the Congress of the United States. 1983. *The impact of randomized clinical trials on health policy and medical practice.* Background paper OTA-BP-H-22. Washington, D.C.: Government Printing Office.

Pate, Antony M., and Edwin E. Hamilton. 1992. Formal and informal deterrents to domestic violence: The Dade County Spouse Assault Experiment. *American Sociological Review* 57: 691–98.

Paternoster, Ray; Bobby Brame; Ronet Bachman; and Lawrence W. Sherman. 1997. Do fair procedures matter? Procedural justice in the Milwaukee Domestic Violence Experiment. *Law and Society Review.*

Raspberry, William. 1998. Tried, true and ignored. *Washington Post*, 2 February, A19.

Sackett, David L., and William M.C. Rosenberg. 1995. On the need for evidence-based medicine. *Health Economics* 4: 249–54.

Sherman, Lawrence W. 1984. Experiments in police discretion: Scientific boon or dangerous knowledge? *Law and Contemporary Problems* 47, no. 4: 61–81.

———. 1992. *Policing domestic violence: Experiments and dilemmas.* New York: Free Press.

———. 1993. Defiance, deterrence and irrelevance: A theory of the criminal sanction. *Journal of Research in Crime and Delinquency* 30: 445–73.

——— and Douglas A. Smith. 1992. Crime, punishment and stake in conformity: Legal and informal control of domestic violence. *American Sociological Review* 57.

Shewhart, Walter A. 1939. *Statistical methods from the viewpoint of quality control.* Edited by W.E. Deming. Lancaster, Pennsylvania: Graduate School of the U.S. Department of Agriculture.

Sparrow, Malcolm; Mark Moore; and David Kennedy. 1990. *Beyond 911: A new era for policing.* New York: Basic Books.

Steinberg, Jacques. 1998. Public shaming: Rating system for schools: Some states are finding that humiliation leads to improvement. *New York Times*, 7 January, A19.

Strang, Heather, and Lawrence W. Sherman. 1996. Predicting domestic homicide. Paper presented to the American Association for the Advancement of Science, January, in Baltimore, Maryland.

U.S. Congress. House. 104th Congress, 1st sess., H. Rept. 104-387, sec. 116.

Weiss, Alexander. 1992. Diffusion of innovations in police departments. Ph.D. diss., Northwestern University.

Zuger, Abigail. 1997. New way of doctoring: By the book. *New York Times*, 16 December, C1.

## ABOUT THE
## POLICE FOUNDATION

The Police Foundation is a private, independent, not-for-profit organization dedicated to supporting innovation and improvement in policing through its research, technical assistance, and communications programs. Established in 1970, the foundation has conducted seminal research in police behavior, policy, and procedure, and works to transfer to local agencies the best new information about practices for dealing effectively with a range of important police operational and administrative concerns. Motivating all of the foundation's efforts is the goal of efficient, humane policing that operates within the framework of democratic principles and the highest ideals of the nation.

---

## POLICE
### FOUNDATION

**1201 Connecticut Avenue, NW, Washington, DC 20036**
**(202) 833-1460 • Fax: (202) 659-9149 • e-mail: pfinfo@policefoundation.org**