New Perspectives in Policing

JANUARY 2011

HARVARD Kennedy School
*Program in Criminal Justice
Policy and Management*

NIJ
National Institute of Justice

Governing Science

Malcolm K. Sparrow, Ph.D.

Introduction

A favorite family pastime is to discover road signs that, either through ambiguous wording or lack of punctuation, lend themselves to multiple different interpretations. One of my daughter's favorites has always been:

If read as presumably intended, the sign warns motorists there might be children nearby, and SLOW is a command. But my daughter points out that SLOW could be an adjective instead: Perhaps motorists should allow more time for the (slower) children to get out of the way. And, she says, if this appeared outside the headmaster's office rather than by the side of the road, it might be a reminder to the children themselves to conduct themselves with decorum and not go tearing around.

Likewise, the title of this paper — *Governing Science* — could suggest three different meanings in the context of policing. Perhaps the word *governing* is an adjective, in which case it would be the *science* that is doing the governing. Then the

object might be to explore and define the science that should govern police as they consider how to conduct their business.

A second interpretation could be that *governing science* is a job to be done, with perhaps a hint that science (or scientists) might need to be controlled or restrained. In the context of policing, a discussion along these lines might set appropriate limits for the role of science and the influence of scientists.

A third possible interpretation arises from reading the phrase *governing science* the same way we read the phrases *fishing tackle* and *climbing gear.* There is a challenge to be met: to catch fish, or to conquer mountains, or to provide quality democratic governance. Through years of accumulated experience and often painful experimentation, those facing the challenge develop a sense of *needs.* To meet those needs, they invent or design various types of *tackle* (for fishing), *gear* (for climbing), or *science* (for governing) to help get the job done. If we wanted to know what the *science for governing* was that might improve the quality of life in a democracy, we would first focus on clarifying the role for police within the broader frame of democratic governance; second, we would define the types of science and areas of application that might best serve in support.

David Weisburd and Peter Neyroud have presented a paper in this series, titled "Police Science."[1] Their subject, broadly viewed, covers the merits of closer collaboration between the fields of policing and scholarship. Anyone who cares about policing cherishes that collaboration enormously. It has already delivered considerable benefits for policing and is poised to deliver many more. Everyone should want

that relationship to flourish. However, at this time, the relationship remains fragile, and much harm might be done if we accept a vision for the future of the relationship that is somehow misguided, inappropriate or off-base.

The Evidence-Based Policing Movement

Despite their very broad title, Weisburd and Neyroud (hereafter W&N) echo many of the familiar themes of the *evidence-based policing* movement (hereafter EBP), which espouses a very particular vision of how the relationship between scholars and police should work. In presenting their diagnosis of how and why the relationship currently fails, W&N emphasize the following five major points:

1. They observe a "fundamental disconnect between science and policing."[2] The "evidence-based model for developing practices and policies has not been widely adopted by police agencies."[3]

2. Why not? In their view, mostly because "police agencies have little interest in using scientific methods to evaluate programs and practices,"[4] and police generally implement strategies, therefore, "with little reference to research evidence."[5] W&N describe "the lack of value of science in much of the policing industry"[6] and suggest that, in the future, police "will have to take science seriously."[7]

3. W&N contrast the police profession with other professions — particularly medicine and public health — that have huge research infrastructures and substantial levels of government-funding to support research. As a result of these disparities

in attention to science, W&N state that medical practice is now based soundly on scientifically validated practices, whereas policing is not.

4. W&N focus mainly on two types of science relevant to policing. One is *social science* research (which includes *criminology*), and the other involves *new technologies* arising from advances in natural sciences and engineering (e.g., DNA, computer forensics, and surveillance).[8] They regard the police profession as insufficiently concerned with social science research and overly eager to adopt new devices and technologies, even without properly evaluating their efficacy.[9]

5. W&N conclude that "a radical reformation of the role of science in policing will be necessary if policing is to become an arena of evidence-based policies."[10] Their proposed solution is a "shift in ownership of police science from the universities to police agencies,"[11] and they provide various suggestions as to how that might happen. One thing they suggest is a "committed percentage of police spending devoted to research, evaluation and the development of the science and research base … ."[12]

W&N acknowledge some divergence of interests between the partners (scholars and police) in terms of their areas of interest. They observe that police need to act quickly and care about issues such as finance and efficiency as well as effectiveness. Academia, by contrast, is often slow to reach any conclusions, often too late to be operationally relevant and, in studying crime prevention, focuses on programs of not much interest to police (such as early childhood interventions and their effects on delinquency or criminal propensity later in life).[13] W&N also suggest that part of the underlying problem might be that police do not study scientific methods and scientists do not know much about operational policing, so the relationship has a rocky start, lacking common ground or shared experience.

W&N's proposed remedy involves repositioning the relationship, changing the sources of impetus and support, and thus shifting the balance of ownership between the partners. They focus less on the underlying defects in the relationship and why it is not working. They certainly admit that it is not working; indeed, that is what motivates their paper. They express concern that investments made during the 1990s have since dwindled, police science having failed to establish itself or produce many results of value or relevance to police agencies.

This paper focuses on the underlying assumptions of W&N's paper rather than on its particular suggestions. Many social scientists would not relish their suggestion — that police take charge of the research agenda — fearing a drop-off in the quality of scholarship. In that respect, W&N's conclusions are unusual. However, their underlying assumptions — that policing should be *evidence-based,* and that you can't know what works unless you take scientific research seriously — align closely with the foundations of the evidence-based policing movement.[14] This paper examines the underlying assumptions of that broader EBP movement, as what EBP proposes requires some counterbalance and caution, particularly at this time in the development of policing.

Evidence-based policing rests on an underlying assumption that the only way for police to know what works is for them to allow social scientists — the professional evaluators — to make determinations for them, and that social scientists, being trained in statistical and empirical methods (whereas police generally are not) can offer their "high science" of controlled experiments and sophisticated program evaluation methods. Police ought then to be keenly interested in and grateful for the truths that social science methods make available. Furthermore, the champions of EBP propose that police should subsequently limit themselves to using only those programs that the scholarly community has been able to establish as effective. In other words, *science should govern policing.* Thus, the central message in the EBP movement aligns quite well with the first of the three possible interpretations of *governing science.*

Lawrence Sherman, describing the underlying theory of EBP in 1998, proposes:

> 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."[15]

Police practitioners might bristle at the notion of being herded (like cats) by social scientists. However, Sherman pushes further, proposing that police be *evaluated* on the basis of whether they conform to what the researchers have recommended:

> 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. ...

> Evidence-based policing is about two very different kinds of research: basic research on what works best ... and ongoing outcomes research about the results each unit is actually achieving by applying (or ignoring) basic research in practice.[16]

This kind of language infuriates police practitioners. Should police managers — who carry all of the responsibility for day-to-day policing and suffer directly the consequences of failure — be chastised by social scientists (who carry none of the responsibility) simply because they chose to ignore a published research finding, or executed an untested or unproven strategy? The idea that science should guide and govern policing in such a way — so that scientists discipline practitioners who don't comply with scientific guidelines — seems ridiculous to practitioners and completely inappropriate to many academics as well. But exactly why the relationship should *not* be structured this way is a serious enough question, which this paper seeks to answer.

Many of us are more attracted to the third interpretation of *governing science* that, by exploring

the police role in the context of democratic governance, emphasizes multiple dimensions of performance and value, and embraces a range of operational styles that move considerably beyond the replication of a small number of "proven" or approved programs. This third interpretation of *governing science* also seems most neutral on the question of which partner (police or science) is supposed to govern the other. It suggests a more healthy collaboration in the long term, with each party delivering their appropriate and respective contributions in support of democratic policing.

Given the more aggressive claims of some of EBP's champions, there is also some serious work to be done along the lines of the second interpretation. Police themselves need to do some *governing*. The police profession needs:

- A more comprehensive view of the range of scientific methods relevant to policing.

- A proper understanding of where different types of science belong.

- Confidence to specify the investments in science that they most need.

- A clear sense of what might be at risk when scholars claim too much or stray beyond their proper role.

Periodic Reminders for Social Scientists

Social scientific research methods have their place, of course, in adding to knowledge. The evidence-based policy movement in general emphasizes program evaluation techniques and concentrates on determining causation. Many

of the relevant research techniques require analytical sophistication. Valid experiments take considerable care and skill to design, conduct and evaluate. The tools of EBP are expensive, but anyone who values knowledge should surely value methods that can help to produce it. Reliable findings about what works, and what doesn't, can help avoid the perpetuation of useless practices and can prevent police officials or politicians from making bogus claims about their achievements or perpetuating useless programs for personal or political reasons. Police managers should surely take note of experimental results and research findings that impinge on operational decisions they need to make. Not to do so would be professionally irresponsible.

From time to time, though, it seems that social scientists need to be reminded of a few things:

- They have no monopoly on useful knowledge or on useful methods for acquiring it.

- Experience and skills count too; there are myriad ways of discovering useful truths without the elaborate machinery of social science evaluations.

- The majority of scientific advances benefitting humankind have arisen and become firmly established without their help.

- "Lay inquiry," with its messier methods and iterative ad hoc experimentation, contributes mightily to the development of knowledge.

- Program evaluation comes very late in a long process of research, problem identification,

diagnosis and policy development. All of the earlier stages — spotting problems in the first place, scoping them, figuring out their structure and dynamics, and designing a set of plausibly effective interventions — all require analytic support, too, but not normally of the specific types offered by the conventions of social science research.

In 1990, Charles Lindblom, a professor of political science at Yale University, published *Inquiry and Change: The Troubled Attempt to Understand and Shape Society.* Lindblom set out to examine "how people in contemporary industrialized societies, competently or not, go about gathering and analyzing information in grappling with social problems."[17] Lindblom's "people" (who go about this task) include politicians, citizens, natural scientists, social scientists, practitioners, and ordinary but curious folk — whom he labels "lay inquirers." By no means did Lindblom set out to attack the social sciences, but in the process of evaluating relative contributions from different types of inquiry and groups of inquirers, he does end up giving social scientists a very hard time. They make the mistake, he says, of overvaluing their own highly technical approaches to the acquisition of knowledge and of presuming that opinions reached any other way must stem from unfounded beliefs or foolishness:

> To be sure, many social scientists and other commentators on social problem solving have fallen into believing that decision makers can approach problems in only one of two ways: either technically, as means to ends, or with all the rigidities, obfuscations, and imprecisions of ideology. But a third

option is available: selective and varied probing of both ends and means, as well as of other values.[18]

Mark Moore (2006) also comments on the challenge that a continuum of knowledge poses, and the perils of ignoring everything between the extremes:

> [B]oth the research and the practice field in policing face the important question of how far down the path of scientific sophistication they should go in their combined efforts to establish a firm experiential and empirical basis for policing. More provocatively put, they have to decide what to do with the knowledge that lies between mere opinion on one hand, and results established through randomized trials on the other.[19]

A 1995 paper by Moore, titled "Learning While Doing," examines the linkages between knowledge and policy formulation, specifically in the context of community policing and violence prevention in the United States.[20] Moore recognizes, of course, the value of social science research methods and acknowledges their place in policy development, but, like Lindblom, he warns against giving them too central a role in policy development:

> Let me hasten to say that I don't think that social scientists are wrong to want knowledge to guide policy. Indeed, it would be irresponsible not to use thought, evidence and experience to guide policy makers when they commit substantial public resources to a particular goal. Instead, I

think their mistake lies in having too narrow a view of what constitutes knowledge valuable enough to use in confronting public problems, too rigid an idea about where and how useful knowledge accumulates in the society, and too unrealistic a view of how knowledge might best be diffused and deployed in aid of both immediate action and continued learning.[21]

Not Just Another Periodic Reminder

My purpose here is not just to issue yet another periodic reminder. Others have done that job quite thoroughly elsewhere and continue to do it in a variety of fields, whenever social scientists exaggerate their own contributions or attempt to exert control over practitioners (i.e., to *govern* policymaking). The contention of this paper is stronger, more particular, and timely, I hope. I believe that we are in a particularly important period in the development of police science, requiring enriched and productive relationships between police and academia. I also believe that much harm might result if we give EBP a dominant position in the context of that relationship.

Why Police Should Govern the Role of Science

Here are three reasons why the police profession should work particularly hard to govern science at this time.

1. The methods championed by proponents of EBP are fundamentally incompatible with the operational realities of problem-oriented policing. Although many departments have made some progress in learning some particular forms of the problem-solving method, relatively few have developed the kind of versatility that Herman Goldstein originally envisaged. Fewer still have developed the range of analytic techniques, organizational fluidity, and related managerial skills that would enable them to work effectively on problems of all shapes and sizes. The maturing of the problem-solving approach remains a priority for the profession, particularly as the range of threats confronted by police expands beyond those that are neighborhood or place-based. EBP represents a potential threat to, and a diversion from, the styles of scientific inquiry needed to advance the art of problem-oriented policing. Social scientists championing the cause of EBP, if given their head at this particular point in time, could unwittingly obstruct the maturation of the problem-solving strategy.

2. The social scientific research methods embraced by EBP represent a tiny fraction of the scientific methods relevant to policing. They should therefore represent a small portion of the relevant investment portfolio, and should garner a relatively small fraction of the attention given to science. Giving too much attention to EBP at this time necessarily means giving too little attention to a much broader range of scientific inquiry methods that deserve higher priority. Equating EBP with science is grossly misleading.

3. The form of the relationship between police and academia envisaged by EBP is unstable and unsustainable. There is too much in it for the social scientists, and almost nothing in it for the police. That is precisely why the champions of

EBP press so hard, and why police continue to show so little interest and remain largely unaffected. It is not so much that the relationship needs to be relocated (as W&N suggest); it needs to be redefined. The prescription is wrong. If EBP is given a central place in the relationship, the relationship may in fact be damaged, and many other opportunities for productive collaboration may be lost as a result.

The following sections examine each of these three arguments in more detail.

Why Evidence-Based Policing Is Fundamentally Incompatible With Problem-Oriented Policing

EBP is incompatible with POP for the following seven reasons:

1. EBP is too slow in making determinations to support operational problem-solving. The problems that spawned the interventions have themselves long since passed, or morphed into another form, by the time the interventions can pass through the elaborate experimental and evaluative procedures espoused by EBP. EBP may eventually produce dependable results with high levels of confidence, but these typically arrive between 3 and 5 years after the development of an intervention. This makes EBP findings relevant to operations only when it evaluates programs that are permanent or long-standing and change very little over time or across jurisdictions. Such programs are not the focus of problem-oriented policing, which seeks ad hoc and sufficient solutions for the problems of the day and then moves on quickly to the problems of tomorrow, expecting that those will be different.

2. EBP produces no new solutions and may even narrow the range of solutions available. Proponents of EBP suggest or imply that police should only use those methods that EBP scholars have already been able to validate. Problem-oriented policing, by contrast, encourages creativity and rapid experimentation, thus dramatically expanding the range of techniques and methods available. Ceding too much influence to EBP may therefore produce a bias against action and too narrow a search for solutions.[22]

3. Social scientists focus on subtle effects at high (aggregate) levels; problem-solving focuses on much more obvious effects but at lower levels. Social scientists (and economists) have tended to conduct macro-level analyses on aggregate data sets. They like to use sophisticated statistical methods on large data sets to reveal subtle correlations and causations between factors and outcomes. Inheriting these tendencies, EBP emphasizes the importance of evaluating the effect that particular programs (e.g., DARE, early childhood intervention programs, or random patrols) might or might not have on overall crime rates or on some major category of crime rates (e.g., violence), delinquency rates, or addiction rates later in life.

Problem-solving, as taught by Goldstein, emphasizes careful *disaggregation* of broad crime categories, following the intuition that major crime problems have many parts (lower-level components) and that, usually, the various parts each behave differently and depend on different factors. Once the lower-level objects have been found (often through analysis), then each one can be studied and "unpicked." In *The Character of Harms,* I have

described how the art of navigating these lower level strata of *problems* or *harms* is emerging as a vital professional skill for regulators and law enforcement:

> The habits of mind ... have something in common with the skills involved in a relatively mundane task: the undoing of knots. Give a knotted mass of string to an adult, who has developed all of the relevant cognitive skills (and maybe had some experience too), and watch how they behave. Notice how they hold the whole object up to the light, and look at it this way, then that way, turning it around and around, examining it diligently from all sides — careful all the time not to pull or tug or to make matters worse — until they begin to understand the *structure of the thing itself.* As the structure of the knot becomes clearer, so the components or stages of a plan begin to form in their minds. ... If they understood the structure correctly, and fashioned a plan accordingly, the knot eventually falls apart, and is no more.
>
> In the regulatory field, we have a growing list of harms undone, knots untied, risk-concentrations eliminated or substantially mitigated. Invariably, the knots undone by regulators, or others who act in this vein, are not broad, general phenomena (at the level of "air pollution," or "corruption," or "motor vehicle accidents"). Nor are they minutiae, representing single incidents (of crime, or injury, or death). These knots untied, these harms undone, all lie *in between,* where the object of study is larger than a single incident or event, but smaller than a general class of harms. It is in this *in-between* realm where much exciting work seems to take place, amid the complex and multi-layered texture that connects individual incidents at the bottom to entire classes of risk (with their one or two word descriptions) at the top.[23]

The impetus for problem-oriented policing arises in part from the realization that it makes little sense to focus on general programmatic treatments for general crime categories if the texture beneath is in fact highly complex, variegated, and populated by many unlike objects. Problem-oriented policing is born from a conviction that working in the textured layers beneath (rather than at the level of generalities or major crime categories) offers greater promise and quicker results.

4. Ironically, greater influence for EBP may *reduce* the rate of experimentation in policing. Professional researchers, as masters of experimental design and evaluation, regard themselves as the authority on what constitutes a "proper" experiment. Thus, police agencies where the *evidence cops* hold sway might be less inclined to proceed with any experimentation that falls short of scholarly standards. In particular, such agencies might be less inclined to proceed with the type of iterative, developmental and exploratory experimentation that characterizes problem-solving.

EBP proponents want valid controls as well as crystal-clear specification of the intervention being tested. Their design purpose is to establish causal connections. However, problem-solvers' purposes and methods are different. They seek to quickly generate creative, plausibly effective solutions, which are worth trying just because there is a chance they might fix the problem.[24] Problem-solvers certainly want to see problems reduced or eliminated and should be methodologically rigorous when it comes to monitoring the abatement of the specific problems addressed so they can tell when progress is being made (hence, Goldstein's strong emphasis on measurement and monitoring).[25] However, they are not so concerned about proving causality. Consequently, problem-solving does not normally impose the additional methodological constraints that would support determinations of causality. Problem-solvers use iterative techniques, short-cycle development and rapid, early assessments of impact, followed by ad hoc and multiple adjustments — all of which confound the technical methods of social science evaluation. As John Eck has pointed out, "Rigorous evaluations are an awkward, inefficient, and unnatural way to learn about what works when we are interested in small-scale, small-claim, discrete interventions."[26] Hence the danger: If EBP is allowed to set the standards for police experimentation, then much valuable experimentation might be curtailed.

5. EBP may reinforce and perpetuate the program-centric mindset in policing, which problem-oriented policing was supposed to dispel. The entire motivation for problem-solving — not just in policing but also across the whole field of social regulation — is to help public agencies understand the deficiencies of a functional or programmatic view of their work, and discover what it means to be task-based rather than tool-based.[27] Skilled craftsmen do not spend the day staring at the array of tools hanging from the workshop wall, contemplating which ones work and which ones don't; rather, the craftsman stands at the task bench and focuses on what must be accomplished. Problem-solving represents a fundamental departure from a tool-centric or program-centric approach, because it recreates the experience of the craftsman in his shop, standing at the task bench, studying the task, facing the dawning and uncomfortable realization that "I don't have a tool for this"; at which point the successful craftsman invents and fabricates a new tool tailor-made for the job.

Proponents of EBP argue that they, too, realize that programs should not be mindlessly copied from one jurisdiction to another. They acknowledge the need to anticipate adjustments and refinements based on local conditions when replicating successful programs. However, this is a tiny move and not enough to restore the appropriate frame of mind for problem-solving. Make some minor adjustments to a hammer and it is still fundamentally a hammer. Adjust your saw blade, and it still only makes cuts. A tool-focus is what we were trying to escape. An adjustable wrench is still a wrench, and no amount of fiddling with it will help if the task is to retrieve a loose screw lodged deep in an engine crankcase, and the craftsman has no suitable tool for that. Making tools adjustable might make them more broadly useful. Nevertheless, focusing first on programs is still a fundamentally different frame of

mind than focusing first on problems; these two mindsets lead to entirely different organizational behaviors and responses.

6. With its reliance on statistical techniques, EBP may not recognize or reward the best problem-solving performance. In any risk-control or harm-reduction setting real success means "spotting emerging problems early and suppressing them before they do much damage."[28] Sophisticated analysis and pattern recognition capabilities, along with bristling intelligence antennae and other forms of alertness and vigilance, can help an agency spot emerging problems earlier rather than later. The earlier the *spotting,* the less noticeable (in a statistical sense) will be the *suppressing.* The problem itself and the effects of any intervention will each be less discernible through quantitative analysis if the action was early and swift. By contrast, problems that have been allowed to grow hopelessly out of control, and which are then dramatically reduced through some sizeable effort, are much more likely to show up as demonstrable successes through the evaluative lenses of EBP. EBP's methods will mostly recognize only bigger, later suppressions and may not be able to discern or appreciate the deftness and nimbleness that constitutes real problem-solving success. Allowing EBP to arbitrate *what works* could have the perverse effect of leading the profession to celebrate only those crime-reduction successes that had been preceded by substantial failures.

7. EBP focuses only on specific interventions and pays little attention to the development of an agency's problem-solving capacity and skills. Problem-oriented policing has profound implications for almost every aspect of a police department's operations:

- It requires new sets of skills for officers engaging in it.

- It requires extensive analytic support at several different stages of the problem-solving process.

- It makes senior officers responsible for tackling a portfolio of problems or risks rather than managing a portfolio of programs or functions.

- It severely stretches the internal fabric of an agency because the majority of problems simply don't fit neatly within existing organizational units.

- It plunges the agency into a constellation of complicated inter-agency and cross-sectoral partnerships, simply because real-world problems don't respect agency boundaries either.

EBP focuses closely on the evaluation of specific interventions and very little, if at all, on the development of agency competencies. Even interventions that failed — in the narrow sense of having produced no measurable impact on levels of crime or disorder — may nevertheless have contributed to agency experience, developed the capacity and confidence of its officers, enriched important partnerships with other parts of government, and strengthened community engagement through collaborative efforts. For problem-oriented policing to mature, the

profession must pay significant attention to all of these other forms of progress, which EBP tends to overlook.

Evidence-Based Policing Fights Back

Several of these arguments have been made before, and some of the more enlightened advocates for EBP seem prepared to acknowledge many or all of them. But the EBP movement seems unwilling to let problem-oriented policing alone or to recognize it as an area where EBP's preferred methods might have severely limited value. Curiously, as if problem-solving represents some kind of threat to the status of social science, EBP seeks to reassert control, and its supporters appear to have pursued two particular strategies for this purpose.

Evaluating Problem-Oriented Policing as a General Strategy

The first involves moving to a higher level. EBP may concede that social science research methods cannot keep pace with operational policing, and might be too expensive and elaborate to apply to low-level and short-term problem-solving efforts, but they can surely *evaluate the overall strategy of problem-solving!* This represents an attractive proposition for the scholars, if only it were possible. They might be able to establish that problem-solving actually works to reduce crime and disorder, in which case EBP could share the credit for anything that problem-solving subsequently accomplished. Alternatively, perhaps scientific research might demonstrate conclusively that problem-oriented policing doesn't work at all, in which case all of the threats to the scientists' *right to govern* policing, laid out earlier, would simply fizzle away.

As a theoretical matter, evaluating an overall strategy (such as *problem-oriented policing*) is quite different from evaluating a set of particular interventions that the strategy has produced.[29] As a practical matter, there is no way that the efficacy of problem-oriented policing, as an overall strategy, could be determined through formally structured experiments or evaluations. There are simply too many different forms of it, many of them deemed "shallow" one way or another by the scholars,[30] and too little maturity in terms of the broader versatility originally envisaged. The prospect of finding even 50 departments who operate the same version of problem-solving, and another 50 who clearly do not (for the sake of providing a suitable control group), seems extremely remote.

A recent study led by David Weisburd illustrates the difficulties involved in trying to evaluate problem-oriented policing as an overall strategy. It also provides a wonderful illustration of the consequences of focusing first on quality of evidence rather than on a broader search for operational insights. Four researchers set out to conduct a "Campbell Systematic Review" of existing literature in order to determine "whether POP is effective in reducing crime and disorder."[31] Following protocols established by the Campbell Collaboration,[32] these researchers first conducted a massive troll of the research literature, uncovering no less than 5,500 relevant articles and reports. They applied the standard methodological threshold tests and concluded that only 10 of these studies (those that involved randomized or well-matched comparison groups) made the cut.

Weisburd and his fellow researchers then combined the findings from these 10 studies, using meta-analysis techniques, and arrived at the conclusion that POP seemed to have some modest, but nevertheless perceptible effect.[33] However, the researchers noted that, if they had chosen to use a different method of combining the results from these 10 studies (a method called *vote-counting*[34]), then the conclusion would have been entirely different (i.e., "no discernible effect").[35] After all that effort, their eventual determination of whether POP has any effect at all hinges on the researchers' choice among available methods for combining the results.

There was potentially more encouraging news from the second part of this study. The authors noted that, by relaxing their methodological standards somewhat (admitting studies that had pre/post data but lacked control or comparison groups), they could bring in a further 45 studies from the remaining pool. The combined results from this broader collection were "overwhelmingly in favor of POP effectiveness."[36] However, the authors then noted that combining the effects from a broad collection of problem-solving interventions, each aimed at quite different types of problems, seemed problematic. Indeed, it does. After all, the idea was to test the *overall strategy* of problem-oriented policing, not to try to combine a set of miscellaneous but particular *interventions* that problem-oriented approaches had produced. Using statistical aggregation techniques to combine outcomes from interventions focused on quite different types of problems seems vaguely bizarre. It is like posing some general and high-level question such as "Do drugs work?" and then trying to answer that question by combining studies involving quite different drugs, applied to patients with quite different conditions. Normally meta-analysis techniques are used to combine results from several implementations of the *same program*. Cognizant of this difficulty, Weisburd and colleagues add an appropriately cautious rider to these (initially more encouraging) results: "this diversity of programs and approaches also should bring caution to any conclusions drawn from our study."[37]

The net result? A mammoth undertaking, involving the review of 5,500 articles and reports, rejection of all but a handful of them because the evidence they contained was deemed not of sufficient quality, and sophisticated meta-analysis of the few that did clear the threshold, yielding highly tenuous conclusions that readers are advised to treat with "caution." For professional social scientists, this is a veritable *tour de force* demonstrating the highest levels of technical and methodological sophistication. And for operational policing? Probably nothing much useful: no new insights or ideas, and no reliable conclusions. No wonder that scholars across many policy domains are now asking, "What is it about experimental evaluation, or … quasi-experimental evaluation, which leads even the very best of it to yield so little?"[38]

Of course, had this review uncovered hundreds or thousands of properly conducted experiments, rather than only 10, then the results might have been more conclusive. Weisburd and his colleagues are quick to observe the general absence

of such studies, concluding that "the evidence base in this area is deficient given the strong investment in POP."[39] How should we remedy that deficiency? Weisburd and colleagues offer the standard EBP proposal that "a much larger number of studies is needed to draw strong generalizations regarding the possible effectiveness of POP."[40]

Of course, there might be some other ways to remedy the situation. One might pay more attention to other forms of evidence or ponder, at least for a moment, the insights and wisdom contained in the other 5,445 reports.

Gilles Paquet, former President of the Royal Society of Canada, describes a variety of "blockages" to the production of knowledge suitable for informing public policy and aims squarely at the evidence-based policy movement generally:

> The second family of blockages pertains to the notion of evidence. It stems from a tendency of the fundamentalists to summarily reject a whole range of types of knowledge as irrelevant, if not meaningless, if that knowledge does not originate from the credentialized tribe and is not the result of work done according to certain prescribed protocols.[41]

Proponents of EBP have set the bar for *knowing* so high, and made the means for generating knowledge so particular, that they end up knowing relatively little. Operational police need to know much more, just well enough and much sooner, in order to keep up with the pace and variety of the challenges they face.

Focusing on Place-Based Problem-Solving Interventions

EBP will probably never manage to produce a convincing evaluation of problem-oriented policing at the level of a departmental strategy. Perhaps recognizing this, the EBP movement makes a second attempt to re-insert itself firmly into the problem-solving arena. If the research scientists can't keep pace with individual problem-solving projects, and they have little hope of evaluating the overall strategy, then maybe they can find some particular *version* of problem-solving that can act as a proxy for the overall strategy and which they can actually evaluate. EBP does seem to have found one: the use of *place-based interventions.* Much of the current energy in the EBP movement seems to be gravitating to this area — testing the effects of order maintenance and other localized interventions — and confirming for us what must have seemed intuitively obvious to police executives for decades: Place-based problems tend to have place-based solutions.

It seems somewhat curious that EBP, in trying to offer some insight on the efficacy of problem-oriented policing, would end up focusing on such an old and familiar police tradition, one that actually predates Goldstein. Perhaps EBP focuses on place-based interventions because place-based experiments are relatively easy to design and conduct. The data required to identify spatial (or temporal) concentrations already exists. The analysis required to identify geographic clusters is straightforward and familiar. Furthermore, *places,* when divided into treatment and control groups, don't complain, call their lawyers, or

lodge constitutional objections about unequal treatment.[42]

Organizing experiments around *other* dimensions may be more difficult. Substantial ethical difficulties arise and potential legal challenges may result whenever randomized controlled experiments are organized around pervasively criminal families, classes of victims, or different cohorts of schoolchildren drawn into gang-related activity — where substantial groups of *people* end up getting quite different treatments.

"But, in medicine, they do that all the time," some may object. "They conduct experiments on issues of life and death, with human control groups, all day and every day." True. However, medical experimentation is based on *informed consent* and *voluntary participation* — features of the experimental environment that policing seldom enjoys.

One of the broader and more sophisticated inquiries into the efficacy of problem-solving was conducted recently by Anthony Braga and Brenda Bond, working with the Lowell, Mass., Police Department.[43] Through analysis, they identified 34 crime hot spots in Lowell and allocated 17 of them to a treatment group and 17 to a control group, using a matching procedure. Three types of problem-solving interventions were applied within the treatment group: (1) sustained programs of misdemeanor arrests, (2) other "situational" (i.e., place-based) strategies, and (3) some "social service" strategies (referrals and other services offered to specific individuals).

Braga and Bond's analysis of the experiment, which employed *mediation analysis* and other highly sophisticated statistical methods, enabled them to draw two main conclusions: (a) a collection of interventions, "focused at specific high-activity crime and disorder places in the city," can generate crime prevention gains;[44] and (b) "the strongest crime prevention benefits were driven by situational strategies that attempted to modify the criminal opportunity structure at crime and disorder hot-spot locations," with misdemeanor arrest strategies and social-service-type interventions scoring less well.[45]

Should we therefore conclude that situational crime prevention techniques are hereby validated and that the alternate (people-based) strategies should continue to be regarded with continuing skepticism? I think not. I have complete confidence in these two authors' analytic skills, experimental disciplines, and the diligent cooperation of the Lowell Police Department under their Chief at the time, Ed Davis. However, I have a strong suspicion that the conclusions the researchers could draw as a result of this experiment are not surprising and are largely determined by the way the experiment was designed. The crime concentrations selected as the foundation for the experiment were *spatial*. Experience with problem-solving in a broad range of other domains teaches us that the dimensions in which a problem or risk is concentrated are often (but not always) closely related to the dimensionality of the solutions.[46] Place-based problems are more likely to have place-based remedies. Family-centered problems are more

likely to respond to family-centric interventions. Social-needs-based problems are more likely to benefit from the provision of social services. Thus, it is not fair to compare three classes of intervention, each organized around different dimensions, starting with only place-based crime concentrations. One might expect, or might even predict, that place-based strategies would come out on top.

It may be that criminologists conduct place-based experiments simply *because they can.* (In Weisburd and colleagues' Campbell Systematic Review, they found only four randomized studies among the 5,500 POP-related articles, and *all four* involved place-based experiments.[47]) Researchers may therefore be quicker to confirm the efficacy of place-based strategies than other types of problem-based interventions. The danger, of course, is that the audience for these evaluations might imagine this actually teaches us about what works and what doesn't in policing. What EBP can actually "prove" has as much to do with the limitations and feasibility of their own research methods as it has to do with what actually works. Perhaps this is why the list of approved interventions remains so short. The shortness of the list might have much less to do with the effectiveness of policing strategies, and much more to do with the limitations of EBP's approved methodologies, and the difficulties of applying them in the policing environment.

A Broader Range of Scientific Methods

The social sciences have an older brother, the natural sciences, with a better established and more robust record of accomplishment. Natural scientists not only look into different areas (physics, biology, chemistry, astronomy, engineering) but also tend to inquire in different ways.

Social science experimental techniques tend to treat complex systems (e.g., communities, families, school populations, and even crime organizations) as black boxes. Researchers can control the inputs, testing them in various combinations; and they can monitor what comes out at the other end of the box some time later (e.g., delinquency rates, crime rates, addiction rates, or propensity for violence). They can then apply sophisticated statistical techniques to their accumulated data about inputs and outcomes, and draw causal inferences in some cases.

Natural scientists tend to have different instincts. They lift up the lid of the box and peer inside. They poke and prod around, not knowing at the outset what they expect to find, open to all sorts of possibilities, not yet knowing what tools they will need to probe further. Their inquiry methods are *reflexive,* which means that, as Gilles Paquet explains, "knowledge acquired gets integrated during the process; it influences the design and thereby modifies the outcome."[48] They do not emphasize any particular or preferred toolkit, nor do they have ingrained in their consciousness any formally approved hierarchy of evidence. They *explore.* They inspect mechanisms up close, rather than observing inputs and outcomes in the aggregate and from a distance. As Pawson and Tilley observed, very few experiments in natural science use experimental/control-group logic.[49]

Different Scientific Traditions

I remember a recent day-long meeting at Harvard University's School of Law that drew faculty from

several of Harvard's schools and from many disciplines. The subject was addiction and addictive behavior, particularly among juveniles, and the effects that various early childhood programs might have on addictive behavior exhibited later in life.[50] For the first hour or so of the meeting, the social science researchers held sway, describing this study and that one, and what they could and couldn't tell from the collection of available studies (which were contradictory in some areas, and generally inconclusive in the aggregate). The moderator invited Jack Shonkoff (Professor of Child Health and Development, and Director of the Center on the Developing Child at Harvard University), who had been quiet until that point, to comment. His first words were:

> I wouldn't start with program evaluation. Nor would I start by talking about early preventive programs. I'd start with the science, and what we know about early brain development.

Professor Shonkoff and a colleague, Charles Nelson (Professor of Pediatrics), proceeded to explain to the group what they knew about the plasticity of the brain and the effects of toxic levels of stress during early childhood. Through intensive use of brain scans, the pediatric neuroscience community had been able to watch over time the different effects of too much stress, too little stress, and healthy levels of stress during the early years of childhood, when the patterns of synapses within the brain are still being formed. Natural scientists and medical experts know the value of program evaluation, but they draw on a much broader repertoire of inquiry techniques.

Ernest Nagel, in *The Structure of Science,* points out just how much has been learned by the human race through lay inquiry, careful observation, creativity, exploration, experimentation, trial and error, and incremental adjustment.

> Long before the beginnings of modern civilization, men acquired vast funds of information about their environment. They discovered the uses of fire and developed skills for transforming raw materials into shelters, clothing, and utensils. They invented arts of tilling the soil, communicating, and governing themselves. Some of them discovered that objects are moved more easily when placed on carts with wheels, that the sizes of fields are more reliably compared when standard schemes of measurement are employed, and that the seasons of the year as well as many phenomena of the heavens succeed each other with a certain regularity.[51]

Charles Lindblom pushes a little harder and questions whether we actually need social science at all. The accomplishments of the *natural* sciences and engineering, he proposes as a stark contrast, are many and obvious:

> Yet the troubling possibility persists that with no or only a few exceptions, societies could perhaps continue to go about these and other activities if social scientists vanished, along with their historical documents, findings, hypotheses, and all human memory of them. …

The disappearance would presumably in some ways render social tasks more difficult, but perhaps in no case render any existing social task impossible, as would the disappearance of any one of many contributions from natural science and engineering. The value of social science to social problem-solving remains clouded to a degree that would shake any social scientist's complacency.[52]

My purpose in quoting these rather pointed arguments is not to dismiss the relevance of social science research methods to policing but, rather, to press the point that social scientific experiments and evaluation constitute a relatively small and very particular subset of the relevant inquiry toolkit.

We should at least consider which *natural science* inquiry methods might turn out to be relevant or important for policing. A great many of them, I would suggest. Most of what we know about social problems and most of the knowledge already accumulated by police stems from the mindset and methods of natural science inquiry — observation, inspection, investigation and diagnosis, leading to the development of ideas about the scope, nature, and dynamics of various dysfunctions and breakdowns in the social order. Even in policing, natural science inquiry methods have a better established and more robust record of accomplishment than social science's experimental methods.

Some sociologists and criminologists might complain that this is unfair and might protest that they themselves use many of the methods of natural science inquiry, even when examining social issues.

Indeed, some of them do. Many social scientists engage in field research, case studies, observation and reporting, synthesis, evaluation, hypothesis development and testing. Many of them have an attitude of professional curiosity, conduct careful observations, compile descriptions, construct stories and derive meaning, offering insights that others may then accept or reject.

However, an *elite* emerges within the discipline: the *randomistas,* as they are known in the field of development economics.[53] They argue that one cannot possibly know anything for sure without a randomized, controlled experiment. They set the standards for professional inquiry so high, and focused on such particular methods, that they then become the ones uniquely qualified to make determinations. They explain carefully to their peers, and to the rest of the world, why more casual or unstructured methods provide no substitute, and how most people therefore really don't know anything for sure.

In this sense, regrettably, EBP is in danger of developing as an *elite* science. Many of its proponents are thinly disguised *randomistas,* and some have no disguise at all. They focus on the most demanding levels of proof, view lay inquiry as poorly structured and therefore invalid, and claim the monopoly right to govern operational decisions in policing. Whatever progress had been made — when social scientists learned to embrace a broader range of natural science methods — is swiftly undone when the *randomistas* produce their hierarchy of evidence and draw threshold lines across it. They leave virtually all of the natural science inquiry methods below the line, effectively demoting them

to the unacceptable category, for which there is no place within their "elite (social) science."

EBP's Scientific Methods Scale

The EBP movement has developed a five-level hierarchy, which they call a *scientific methods scale.*[54] Randomized controlled experiments belong at the highest level (tier 5), whereas mere correlations belong at the lowest level (tier 1). The threshold for acceptability is drawn at tier 3, where experimental designs include "moderate statistical controls" such as comparisons between control and treatment groups and between pre- and post-treatment:

> Programs coded as working must have at least two "level-3" to "level-5" evaluations showing statistically significant and desirable results and the preponderance of all available evidence showing effectiveness.[55]

Hence, police programs will only be deemed *proven* if multiple independent studies have confirmed their effects. To be valid, the contributing "experiments and quasi-experiments should include large samples, long follow-up periods, follow-up interviews, and provision for an economic analysis."[56] EBP has also declared some willingness to consider findings from *meta-studies,* which compile volumes of data from multiple sources as an alternative to designing new experiments from scratch. To be acceptable, such studies must be extensive and suitably sophisticated. Such stringent specifications will surely have the effect of keeping "acceptable methods"

beyond the capabilities of ordinary mortals and thereby guaranteeing a stream of social science research funding for decades to come. EBP has set its thresholds, and the vast majority of ordinary "lay inquiry" and natural science methods fall short of it.

Above EBP's threshold line (in terms of acceptable methods for establishing program effectiveness) lie controlled experiments (preferably randomized), meta-studies, and a miscellaneous collection of other sophisticated program evaluation techniques. Social scientists have one other favorite tool — regression analysis — used not so much to determine causality (as it mostly establishes correlations rather than causal linkages) but used at an earlier stage of inquiry to identify *factors* that might exert significant influence on specific *outcomes.* Identifying such factors, of course, could lead eventually to clues about potential interventions and policy effects. However, there would normally be a lot of ad hoc probing, prodding, and messy experimentation before a regression finding (establishing the significance of one factor or set of factors) could be translated into an intervention design. Nevertheless — and perhaps because of the sophistication and apparent ubiquitous applicability of the tool — regression analysis also seems to have earned a place in the social science elite's preferred toolkit.

Other Ways of Knowing

Perhaps it is worth bearing in mind that Sir Isaac Newton established the laws of motion and elasticity without using any of these preferred

methods. Using his trademark combination of scientific curiosity and creativity, he first estimated the speed of sound in air by clapping his hands at one end of a walkway in Neville's Court (Trinity College, Cambridge) and measuring the interval between the clap and the echo returning from a wall at the far end of the courtyard. Having no stopwatch, he synchronized the swing of an adjustable-length pendulum to match the delay and later computed the period of the pendulum. He surely conducted experiments. He did so to test the theories that he developed to explain the observations that he so carefully made. Observation begat theories, and theories begat further observation. His experiments were not randomized, nor controlled, and involved no meta-analyses nor regressions.

Perhaps it is also worth bearing in mind that the vast majority of modern medical knowledge has accumulated without the use of this elite toolkit. Yes, specific remedies are now tested through randomized clinical trials, but medical students first learn anatomy and are required to dissect a cadaver as part of their training so that they can see how the human body is put together. They learn how the musculo-skeletal system works, then the cardio-pulmonary system, the endocrine system, the nervous system, and so on. Next, they learn about the myriad ways in which physiological failures can occur. During their training, they talk to hundreds or thousands of patients with various symptoms and conditions. They do most of this learning by using their own eyes and ears, aided by microscopes, stethoscopes, scanners of one kind or another, patient interviews and examinations, and lab tests galore.

Only at a very late stage, when the medical community wants to check the efficacy of one treatment protocol compared with another, in relation to a specific condition or diagnosis, does it turn to controlled experiments. When it does, medicine has many advantages over policing. Throughout the world, the human body works basically the same way and is subject to common modes of failure or dysfunction. (The same is not true for societies, communities, neighborhoods or crime problems).[57] These medical failure modes are finite in number and have already been codified as a list of diagnoses (not true for policing problems). For any one diagnosis, there are at least thousands of cases, if not millions (not true for policing problems). For clinical trials in medicine, hundreds or thousands of patients can generally be identified who not only share the same underlying diagnosis but also satisfy any additional demographic filters that experimenters may choose to apply.

Modern medicine generates numerous clinical trials, in part, because of the interests of corporations. Manufacturers of drugs and medical devices have powerful incentives to overstate the effectiveness of their products and to press those claims on doctors and patients alike. Regulators (such as the U.S. Food and Drug Administration) require manufacturers to supply evidence from clinical trials before granting approval for new products or certifying new uses for them. Stringency in testing seems natural and appropriate in such circumstances, given the commercial incentives in play.[58]

Randomized studies turn out to be easier to run, as a practical matter, for drugs than for other types of medical intervention. The administration of

drugs is relatively easy to standardize. As medical researchers have pointed out,

> [T]here is a lack of generalisability once we move away from drugs to manual interventions. For example, difficulty in devising practice policies in surgery arises because decisions depend on the features of a particular patient (obesity, anatomy, quality of tissue), the particular surgeon, and various external factors (equipment available, competence of assistants).[59]

What is true for surgery is most certainly true for policing, with little prospect of precisely replicating interventions across jurisdictions. The good news, in medicine, is that for pharmaceuticals — an area where commercial propositions deserve the most careful scrutiny — the treatments happen to be relatively generalizable, which makes clinical trials feasible.

It may be good news for policing that there are relatively few commercial interests at stake in advancing one crime prevention strategy over another. We should certainly beware those cases where specific commercial products are closely associated with specific policing strategies or tactics (as may be the case with the recent emergence of *predictive analytics,* the adoption of technical products such as Tasers and particular types of firearms, body armor or vehicles). Such circumstances demand heightened skepticism, closer scrutiny and stricter evaluative standards. There do not appear to be any particular commercial interests behind problem- or

community-oriented policing, so expensive research to safeguard against commercially motivated and overblown claims of effectiveness probably are not needed in these areas. Lower levels of evidentiary support for these strategies might serve the profession perfectly well.

Natural Science Inquiry Methods in Policing

Does the police profession use the equivalent of natural science inquiry methods? Absolutely. I would suggest that crime analysis, intelligence analysis, intelligence-gathering, investigations, investigative field-craft and general surveillance techniques all fall squarely into this category. These are the ordinary processes of discovery, structured and unstructured, through which police find out what is happening, and why, and begin to explore how best to intervene. Such methods can be more or less sophisticated, of course, and they can be very sophisticated indeed without involving any of the tools from EBP's elite toolkit. Moreover, police and scholars can collaborate closely and productively around such methods.

The Boston Gun Project provides an obvious example of such a collaboration. Three Harvard scholars worked closely with Boston police and other agencies to find the causes of escalating juvenile homicide rates in Boston and figure out what might be done. They were given some hypotheses, developed more of their own, and tested these hypotheses by talking to street workers, gang members, and anyone else who could provide useful insights. They inherited one particular theory — that the violence was fuelled

by an uncontrolled supply of guns from southern states. They checked out that theory by tracing guns used in past homicides back to their point of first sale. What they found out (most of the guns used in homicides were sold first in Massachusetts and were relatively new when used) demolished that theory, and the team quickly abandoned it. Next, they searched for new ideas, listening carefully to a broad range of players. Like natural scientists running back and forth between the lab and the field, these researchers moved back and forth constantly between data analysis and street-level inquiry, each form of investigation informing, enriching and redirecting the other. Eventually, "the structure of the knot" came into focus, and its internal dynamics became clear. The researchers and their police partners saw clearly the structure of the 61 gangs, the patterns of established gang "beefs," and the role played by peer pressure within the gangs when it came to violence. Finally, once they understood the structure of the problem, the project team devised a tailor-made strategy to reverse the effects of peer pressure within the gangs.

What did these researchers *not* do? For this project, they did not conduct any randomized experiments, perform any meta-analyses, nor did they use regression analysis. The entire project was set up and funded (by NIJ) as a problem-solving demonstration project, not as a program evaluation or criminological research project. In fact, there was no formal experimental structure for the project, which may leave the EBP community wondering whether or not Operation Ceasefire really worked, or whether the 63 percent reduction in the youth homicide rate,[60] which quickly followed

implementation of the Ceasefire strategy, was merely some kind of fluke.[61] Maybe, several years later, EBP scientists will come up with some method to confirm (subject to their own standards of evidence) that Operation Ceasefire actually saved lives. Even if they do, we should be grateful for all the lives that will have been saved in the meantime.

What a shame it would be if this type of cooperation between police and scholars were not valued, just because this partnership used nothing from the toolbox of elite science. What a shame it would be if the many forms of analysis this team (and others like them) employed along the way, when unraveling a serious crime problem, were deemed unsatisfactory. What a tragedy if operational policing ever had to wait for social science to catch up.

Because some social scientists use natural science methods, and natural scientists occasionally run controlled experiments, drawing a sharp line separating the two sets of research methods is somewhat problematic. However, distinguishing the much smaller set of social science methods approved by the EBP elite from all other scientific methods is actually much easier, simply because the preferred toolbox is so small and its contents quite easily enumerated.

Data Analysis and Pattern Recognition in the Natural Sciences

Some may make the mistake of assuming that natural science methods look only locally, through the microscope or by way of lab tests, at one object at a time; and that any methods involving analysis of large data sets (such as crime analysis) must obviously belong to the social sciences. This is plainly

wrong. The entire field of *pattern recognition* techniques, for example, aligns better with the instincts of natural scientists than with those of social scientists. Fraud detection algorithms (which operate across massive databases of financial and transactional data) have nothing to do with program evaluation or controlled experimentation, and everything to do with searching for anything strange that might be there and exploring the nature of anything that appears.

Nicholas Christakis (Professor of Medicine and Medical Sociology at Harvard Medical School) explores the mechanisms through which disease or health effects are transmitted through social networks. Through the application of network analysis and other analytic methods, he has shown, for example, how obesity can be transmitted through social ties as individuals influence the attitudes and behaviors of family and friends around them. Christakis reports that the advent of social networking sites such as Facebook have presented researchers in this area enormous repositories of data, electronically available and ripe for analysis. His use of them is highly sophisticated, deeply scientific and analytical in nature. Nevertheless, his instincts align more with the mindset and methods of investigation and exploration rather than program evaluation, hence more with the habits of natural scientists than those of social scientists.[62] In a recent interview with *Harvard Magazine,* Christakis explained the significance of natural curiosity and open-mindedness, coupled with a broad range of analytic instruments, in finding out how things work. He applies the same mindset, he implied,

when exploring terabytes of social network data as Galileo employed when he peered through his telescope to fathom the structure of the heavens:

> In some ways the availability of these new kinds of data is like what the microscope was to Van Leeuwenhoek or the telescope to Galileo. When the telescope was invented, Galileo just started looking at stuff. He looked at the moon and he saw mountains. He looked at Jupiter and found moons encircling it. He looked at the sun and found sun spots. There's this huge part of science which is just about careful observation and curiosity about the world.[63]

This "huge part of science" routinely dwarfs social science in making contributions to knowledge. It would be strange indeed if Galileo and Newton, who have taught us so much about the way the universe works, were deemed not to have engaged in "high science" simply because their methods did not rely on randomized experiments or program evaluation techniques.

There is no prima facie reason why the ratio of natural science methods to social science methods applicable to policing should differ markedly from this ratio in other areas. One can obtain a rough sense of where that ratio lies, in general, by comparing the rate at which new articles are abstracted into various academic citation indices. For the United States, the rate at which articles are being added to the general science citation indices runs at roughly five times the rate at which articles are being added to equivalent

social science citation indices.[64] Across a range of industrialized nations, this ratio varies between 5:1 and 10:1. In other words, social science may account for no more than 10 to 20 percent of new science.[65] Given that the elite toolbox and preferred methods of EBP represent a relatively small subset of the overall social science toolkit — certainly less than half — then it might be reasonable to guess that EBP should represent no more than 5 to 10 percent of the investments the police profession could usefully make in scientific inquiry. From this perspective, the notion of EBP playing a central or dominant role in the relationship between police and scholars begins to look woefully unbalanced.

Weisburd and Neyroud do mention the natural sciences and engineering, in passing, but they lump these together under the general rubric of *devices* or *technologies,* which they say the police are much too eager to adopt. They virtually ignore natural science inquiry mechanisms, normally the larger piece of the scientific pie. W&N do briefly mention crime analysis, commenting positively on some recent advances in its sophistication and versatility. However, they do not seem to seize on crime analysis (as I believe we should) as an example of a different type of science that is more directly relevant to operations. W&N observe little "involvement between scientific work in universities and the work of crime analysis in policing."[66] Specifically, they complain:

> Police departments do not … encourage their scientific staff to publish in scientific journals in criminology; indeed, they generally discourage them … . Science in this sense is not a part of large policing centers.

> The implication of this is that the scientific quality of crime analysis units is often relatively low.[67]

In other words, W&N suggest that crime analysis should involve the same type of analytic sophistication as criminological research, and any crime analyst worth his or her salt should be publishing studies in scientific journals. On this point, as on so many others, W&N seem to equate science with criminological research and ignore the significance of inquiry and analytic methods that are perfectly valuable for diagnosing crime problems and guiding operations but lie well outside the realm of evidence-based policing and criminological research.

It is quite a different thing to make the police profession "an arena of evidence-based policies"[68] rather than a sophisticated user of scientific methods. Conflating these purposes may well serve to elevate the status and interests of social scientists but would be disastrous for police. To set things more properly in balance, one might surmise that evidence-based policing, because it is unlikely to meet more than 5 percent of the police profession's overall scientific needs, should probably receive no more than 5 percent of the funding for police science and a commensurate level of attention.

If such a rule seems remotely reasonable, then the police, along with their scholarly supporters, will need to make a serious commitment to figuring out what mix of investments should constitute the remaining 95 percent of the science agenda because, so far, we have heard less about this part. It is not too hard to identify some of the priorities

in this space. The police profession, aided by the scholarly community, should:

- Aim to broaden the range of crime analysis techniques available, beyond the narrow traditions of spatial analysis and CompStat. We should help police understand that problems come in a daunting array of shapes and sizes, and help them to develop the broader analytic versatility required to reveal a broader range of problems and bring them into clearer focus.

- Learn more about the interplay between data-mining and investigative field-craft, so that macro-level analysis and micro-level examination can each inform, refocus and complement the other in a continuous cycle, as police seek to identify and comprehend the complex phenomena they confront.

- Continue to develop intelligence analysis techniques versatile enough to assess local, regional, national and international crime problems (because the security threats that confront police continue to diversify and vary considerably in scale).

- Develop a clearer vision of what might constitute *analytic vigilance* for the profession, learning to avoid "failures of imagination," knowing how much time and resources to spend on looking, and knowing how to look, even when there might be nothing to find.

- Explore and import a much broader array of pattern recognition techniques to help police spot emerging, invisible and unfamiliar problems earlier and more reliably.

- Define and refine the (several) supporting roles for data analysis, measurement and monitoring during the different phases of the problem-solving process.

- Invest in the quality of analytic support available to operational policing and dramatically increase the availability of analytic services throughout departments.

- Continue the drive to elevate crime analysis and intelligence analysis to the status of a profession,[69] taking care to prevent this emerging discipline from being confused with (or captured by) criminology or the social sciences.

All of these investments would be deeply analytical and could draw on diverse streams of scientific knowledge and scholarship.

Toward a More Stable and Sustainable Relationship

The relationship between academia and the police profession remains tenuous and vulnerable, but significant progress has been made in developing fruitful collaborations of many types. Scholars have worked with police on political management, organizational design, organizational change, police culture, training, enhancing educational standards within the ranks, and developing analytic methods as well as helping to develop operational strategies and tactics. Scholars have participated in problem-solving projects, chaired inquiries and commissions and have served extensively as consultants to police executives.

All of this is too valuable to jeopardize. Giving evidence-based policing a central position or allowing it to dominate interactions between police and academia may stifle the relationship.

The form of the relationship proposed by proponents of evidence-based policing offers virtually no benefits for police. The best they can hope for is that the scientists they have invited in, after months or years of research work, will finally confirm what police thought they knew already: that an intervention or program the department had previously deployed did actually work. The downside risk for police is much greater. Maybe the research findings will prove to the world that police actions were irrelevant or ineffective and that apparent successes turn out to be bogus or mere luck. For police managers, what joy! No wonder many executives scratch their heads and wonder why they would want to enter into such a partnership. Meanwhile, the scholars offer police no real help with pressing operational needs because they have such a short list of approved methods. The scholars bear no responsibility for the consequences of action or inaction and feel no obligation to invent anything useful. They mostly want to *evaluate.*

While the benefits for police seem minimal, the costs loom large. Police must proceed more slowly, even in the presence of urgency, in order to satisfy the demands of experimental design. Police agencies must accommodate scholars, providing them access to staff and data, and confronting the legal issues that arise when outsiders are allowed in. Police end up driving the scholars around, keeping them safe, and generally looking after them. Police executives voluntarily subject their own actions and their officers' actions to scrutiny, dealing with the associated press inquiries and reputational risks. Managers have to persuade their own officers to cooperate with researchers despite their workloads, beliefs and worries about outside scrutiny — a task made no easier if the scholars use condescending phrases such as "high science" and "elite scientists." In addition to all of these costs, W&N now propose earmarking "a significant percentage of [a police department's] budget" for research and evaluation,[70] which would exacerbate tensions over resources even further.

Evidence-based policing does have a place in policing, but it needs to be kept in its rightful place. EBP employs expensive and complex methodologies that need to be strategically deployed. There are many areas of policing where these methods are not, and will never be, relevant or useful. Problem-oriented policing may well be one such area. EBP should recognize that and simply leave it alone.

There are other areas where EBP's rigorous evaluative techniques seem more appropriate. Where police adopt programs or methods that are expensive, long term, potentially permanent — and which are deployed in a sufficiently standardized way across many departments — evaluating these programs with a reasonable degree of rigor may well be important.[71] With respect to a small number of major programs, EBP may deliver some value. Then again — given the substantial difficulties involved in conducting controlled experiments within a policing context — EBP might extend its disappointing track record, offering valuable insights few and far between.

The profession should not overlook the many other useful contributions that scholars can make and that science can offer. There are many other forms of scientific inquiry, more akin to natural science methods, that need more urgent development within policing. These are more relevant to the bulk of operational policing challenges and should take priority among science investments at this time.

In closing, consider W&N's key question, "How can we move police science to a central place in the policing industry?"[72] Preferably by understanding the particular and limited contributions that social science research methods can make to operational policing, and by embracing a substantially broader range of investigative, analytic, inquiry and intelligence techniques more generally suited to the operational demands of the profession.

Endnotes

1. Weisburd, David and Peter Neyroud, "Police Science: Toward a New Paradigm," paper of the Second Harvard Executive Session on Policing and Public Safety, Washington, DC: National Institute of Justice, 2010. NCJ 228922.

2. Ibid., p. 2.

3. Ibid., p. 3.

4. Ibid., p. 3.

5. Ibid., p. 3.

6. Ibid., p. 4.

7. Ibid., p. 12.

8. Ibid., pp. 6-7.

9. Ibid., p. 7.

10. Ibid., p. 1.

11. Ibid., p. 1.

12. Ibid., p. 16, table.

13. Ibid., p. 5.

14. Weisburd and Neyroud have claimed, in response to initial drafts of this paper, that their paper presents a broader view of *police science* and a more nuanced view of the proposed relationship between scholars and police than we normally hear from the champions of EBP. Indeed, they do mention in passing, and label as *scientific methods,* a considerable range of innovations in technology and analysis, including problem-oriented policing, CompStat, intelligence-led policing, and crime analysis more generally. Nevertheless, they use the phrase *evidence-based* no less than 23 times (followed by *policing, practice(s), policy(ies), model, approach, profession, research,* and *science*) and state rather plainly in their opening sentence that the goal is to make policing "an arena of evidence-based policies." W&N refer to the type of science that the police profession needs as an "elite" science (p. 13), a "blue chip" science (p. 14), and a "high level" science (p. 15). These characterizations echo EBP's emphasis on professional social-scientific evaluation techniques and scholarly criminological research. Thus, the W&N paper

seems closely aligned with the core themes of the evidence-based policing movement.

15. Sherman, Lawrence W., "Ideas in American Policing: Evidence-Based Policing," Washington, DC: Police Foundation, July 1998, p. 3.

16. Ibid., pp. 3-4.

17. Lindblom, Charles E., *Inquiry and Change: The Troubled Attempt to Understand and Shape Society,* New Haven, CT: Yale University Press, 1990: preface, vii.

18. Ibid., p. 11.

19. Moore, Mark H., "Improving Police Through Expertise, Experience, and Experiments," in *Police Innovation: Contrasting Perspectives,* ed. David Weisburd and Anthony Braga, Cambridge, England: Cambridge University Press, 2006: 325.

20. Moore, Mark H. "Learning While Doing: Linking Knowledge to Policy in the Development of Community Policing and Violence Prevention in the United States," in *Integrating Crime Prevention Strategies: Propensity and Opportunity,* ed. Per-Olof H. Wikstrom, Ronald V. Clarke and J. McCord, Stockholm, Sweden: National Council of Crime Prevention, 1995: 301-331.

21. Moore, "Learning While Doing," pp. 302-303.

22. Ibid., p. 310.

23. Sparrow, Malcolm K., *The Character of Harms: Operational Challenges in Control,* Cambridge, England: Cambridge University Press, 2008: 8-9.

24. Moore, "Learning While Doing," p. 307.

25. For this purpose, single group pretest/posttest designs are perfectly adequate, whereas these are regarded as "inadequate and uninterpretable" by the experimentalists. See "Standards of Evaluations in Problem-Oriented Policing Projects: Good Enough?" in *Evaluating Crime Reduction Initiatives,* ed. Johannes Knutsson and Nick Tilley, Crime Prevention Studies Book Series, vol. 24, Monsey, NY, 2009: 21, 23.

26. Eck, John E., "Learning From Experience in Problem-Oriented Policing and Situational Prevention: The Positive Functions of Weak Evaluations and the Negative Functions of Strong Ones," in *Evaluation of Crime Prevention,* ed. Nick Tilley, Crime Prevention Studies Book Series, vol. 14, Monsey, NY: Criminal Justice Press, 2002: 109.

27. For a detailed exploration of the differences between functions, processes and problems and the implications for agency operations, see chapter 2, "A Different Type of Work," in Sparrow, *The Character of Harms,* pp. 47-72.

28. I refer to this elsewhere as the *whack-a-mole* model for risk control operations; see Sparrow, *The Character of Harms,* pp. 143-146.

29. Nick Tilley points out that, in asking "what works," the *what* can refer to particular interventions, classes of interventions, mechanisms, strategies, or other more complex combinations of the four. He stresses the need to be clear about which level of object one is evaluating. See Tilley, Nick, "What's the 'What' in 'What Works?' Health,

Policing and Crime Prevention," in *Evaluating Crime Reduction Initiatives,* ed. Johannes Knutsson and Nick Tilley, Crime Prevention Book Series, vol. 24, Monsey, NY, 2009: 121–145.

30. Braga, Anthony A. and Brenda J. Bond, "Policing Crime and Disorder Hot Spots: A Randomized Controlled Trial," *Criminology* 46 (3) (2008): 585.

31. Weisburd, David, Cody W. Telep, Joshua C. Hinkle and John E. Eck, "Is Problem-Oriented Policing Effective in Reducing Crime and Disorder? Findings From a Campbell Systematic Review," *Criminology & Public Policy* 9 (1) (2010): 139-172.

32. Such reviews follow the protocols of the Campbell Collaboration and focus on experimental and quasi-experimental studies. See http://www.campbellcollaboration.org, accessed on October 6, 2010.

33. Weisburd et al., "Is Problem-Oriented Policing Effective?," p. 153.

34. Vote-counting essentially grants one vote to each study incorporated into a meta-analysis. Tallying the votes provides an overall score indicating whether a specific intervention produces positive outcomes more often than not, according to the compiled evidence. It is generally regarded as an unsophisticated approach because it makes no corrections for the relative sizes and quality of the different experiments. However, when researchers aim to combine the results from several studies involving *different* interventions (as is

the case with Weisburd's Campbell Review), any of the more sophisticated *statistical* techniques for combining outcomes might be regarded as mathematically inappropriate, and vote-counting might seem more reasonable in these circumstances.

35. Weisburd et al., "Is Problem-Oriented Policing Effective?," p. 153.

36. Ibid., p. 159.

37. Ibid., p. 164.

38. Pawson, Ray and Nick Tilley, *Realistic Evaluation,* London: Sage Publications, 1997: 50. For a discussion of growing dissatisfaction among scholars with prevailing approaches to evaluation and evidence — spanning medicine, education, early childhood programs, and international development, see Smyth, Katya Fels and Lisbeth B. Schorr, "A Lot to Lose: A Call to Rethink What Constitutes 'Evidence' in Finding Social Interventions That Work," Working Paper Series, Cambridge, MA: Malcolm Weiner Center for Social Policy, Harvard University Kennedy School of Government, January 2009. Accessed September 29, 2010, at lisbethschorr.org/doc/Alottolosejan.2009.pdf.

39. Weisburd et al., "Is Problem-Oriented Policing Effective?," p. 140.

40. Ibid., p. 164. An earlier study by Weisburd and Eck, designed to test the efficacy of various policing strategies, drew the same essential conclusion: "[T]he authors find that many policing practices applied broadly throughout the United

States either have not been the subject of systematic research or have been examined in the context of research designs that do not allow practitioners or policy makers to draw very strong conclusions" (42). See Weisburd, David and John E. Eck, "What Can Police Do to Reduce Crime, Disorder, and Fear?" *Annals of the American Academy of Political and Social Sciences* 593 (May 2004): 42-65.

41. Paquet, Gilles, *Crippling Epistemologies and Governance Failures: A Plea for Experimentalism,* Ottawa, Ontario, Canada: University of Ottawa Press, 2009: xv.

42. Although nearby residents, if they knew what was happening, supposedly might.

43. Braga and Bond, "Policing Crime," pp. 577-607.

44. Ibid., p. 598.

45. Ibid., p. 599.

46. For a detailed discussion of the dimensionality of problems and the implications for control strategies, see chapter 4, "Defining Problems: Picking the Dimensions," in Sparrow, *The Character of Harms,* pp. 93-100.

47. Weisburd et al., "Is Problem-Oriented Policing Effective?," p. 147. The study also located six quasi-experimental designs, and four of these were also place-based. The remaining two quasi-experiments focused on probationers and parolees, respectively, who are presumably more easily subjected to experimental manipulation than other segments of the public.

48. Paquet, *Crippling Epistemologies,* p. xvii.

49. Pawson and Tilley, *Realistic Evaluation,* p. 57.

50. Discussion at Harvard Law School, moderated by Professor Philip Heymann and Mathea Falco, January 30, 2009.

51. Ernest Nagel, *The Structure of Science,* New York: Harcourt, Brace & World, 1961:5. Quoted in Lindblom, *Inquiry and Change,* pp. 161-162.

52. Lindblom, *Inquiry and Change,* p. 137.

53. The term *randomistas* appeared first in the field of international development economics, where it applies to those who consider randomized controlled trials (RCT) to be the gold standard when it comes to determining intervention effects.

54. This approach stems from the work of Cook and Campbell in 1979. See Cook, T.D. and D.T. Campbell, *Quasi-Experimentation: Design and Analysis Issues for Field Settings,* Chicago: Rand McNally, 1979. In 1998, a formal scale, now known as the *Maryland Scientific Methods Scale,* was laid out in Sherman, L.W., D.C. Gottfredson, D.L. MacKenzie, J. Eck, P. Reuter and S.D. Bushway, *Preventing Crime: What Works, What Doesn't, What's Promising,* Research in Brief, Washington, DC: National Institute of Justice, July 1998, NCJ 17176. A more recent version of the scale is laid out in Welsh, Brandon C., "Evidence-Based Crime Prevention: Scientific Basis, Trends, Results and Implications for Canada," Research Report 2007-1, Ottawa, Ontario, Canada: National Crime Prevention Centre, Public Safety, June 2007, pp. 13-14.

55. Welsh, "Evidence-Based Crime Prevention," p. 15.

56. Ibid., p. 5.

57. Jim Manzi observes that therapeutic biology "could often rely on the assumption of uniform biological response to generalize findings from randomized trials" but that "higher causal densities in social sciences" make such generalizations "hazardous." See Manzi, Jim, "What Social Science Does — and Doesn't — Know," *City Journal,* Summer 2010. Accessed September 29, 2010, at http://www.city-journal.org/printable.php?id=6330.

58. In fact, the commercial sector seems to have figured out some ways to prevent the FDA from properly monitoring the trials. Most of the trials are now conducted offshore, where the FDA cannot and does not supervise them. Roughly 80 percent of approved marketing applications for drugs and biologics contain data from foreign clinical trials, and more than half of clinical trial subjects were located overseas. The FDA inspects only 0.7 percent of foreign clinical trial sites. See Levinson, Daniel, "Challenge to FDA's Ability to Monitor and Inspect Foreign Clinical Trials," Report OEI-01-08-00510, Washington, DC: Office of Inspector General, Department of Health and Human Services, June 2010: p. ii.

59. Black, Nick, "Evidence Based Policy: Proceed With Care," *British Medical Journal* 323 (August 4, 2001): 275.

60. The youth homicide rate involves victims ages 24 and younger, and the juvenile homicide rate, involves victims ages 17 and younger.

61. For a full account of the project, see Kennedy, David M., Anne M. Piehl and Anthony A. Braga, "Youth Violence in Boston: Gun Markets, Serious Youth Offenders, and a Use-Reduction Strategy," *Law and Contemporary Problems* 59 (1) (Winter 1996): pp. 147-196. For a subsequent analysis of pre/post data and comparisons with the trajectory of youth homicide rates in other cities, see Braga, Anthony A., David M. Kennedy, Elin J. Waring and Anne Morrison Piehl, "Problem-Oriented Policing, Deterrence, and Youth Violence: An Evaluation of Boston's Operation Ceasefire," *Journal of Research in Crime and Delinquency* 38 (2001): 195-225.

62. Christakis does use techniques from social network analysis, which arose first as a subdiscipline of social science. However, social science has no monopoly on the uses and applications of the core ideas from social network analysis. In fact, mathematicians had been studying networks, which they called "graphs," for several hundred years before social science began to realize their significance for the study of social phenomena.

63. Interview with Professor Christakis, reported in Gudrais, Elizabeth, "Networked: Exploring the Weblike Structures That Underlie Everything From Friendship to Cellular Behavior," *Harvard Magazine,* May-June 2010, p. 50.

64. See, for example, "Articles Abstracted to the Thomson-Reuters and Scopus Databases, 2007," *World Social Science Report 2010,* Annex 1, which are basic statistics on production of the social sciences.

65. This is a crude proxy, of course, for the rate of use of various inquiry methods because the categorization of articles is based heavily on the field of study as well as the research methods used.

66. Weisburd and Neyroud, "Police Science," p. 7.

67. Ibid., pp. 7-8.

68. Ibid., p. 1.

69. This cause was championed for many years by the International Association of Law Enforcement Intelligence Analysts (IALEIA).

70. Weisburd and Neyroud, "Police Science," p. 16, and table at p. 16.

71. For a discussion of the circumstances affecting the costs and benefits of strong and weak evaluations, see Eck, "Learning From Experience," pp. 93-117.

72. Weisburd and Neyroud, "Police Science," p. 11.

About the Author

Malcolm K. Sparrow is professor of the Practice of Public Management at the John F. Kennedy School of Government at Harvard University. This paper was written in support of the Executive Session on Policing and Public Safety at the Harvard Kennedy School of Government. The author acknowledges valuable feedback on early drafts from Anthony Braga, Christine Cole, Thomas Feucht, Peter Neyroud, Kris Rose, Ellen Scrivner, Darrel Stephens, Chris Stone and David Weisburd.

Members of the Executive Session on Policing and Public Safety

Chief Anthony Batts, Oakland Police Department

Professor David Bayley, Distinguished Professor, School of Criminal Justice, State University of New York at Albany

Dr. Anthony Braga, Senior Research Associate, Lecturer in Public Policy, Program in Criminal Justice Policy and Management, Kennedy School of Government, Harvard University

Chief William J. Bratton, Los Angeles Police Department

Chief Ella Bully-Cummings, Detroit Police Department (retired)

Ms. Christine Cole (Facilitator), Executive Director, Program in Criminal Justice Policy and Management, Kennedy School of Government, Harvard University

Commissioner Edward Davis, Boston Police Department

Chief Ronald Davis, East Palo Alto Police Department

Chief Edward Flynn, Milwaukee Police Department

Colonel Rick Fuentes, Superintendent, New Jersey State Police

Chief George Gascón, San Francisco Police Department

Mr. Gil Kerlikowske, Director, Office of National Drug Control Policy

Chief Cathy Lanier, Washington D.C. Metropolitan Police Department

Dr. John H. Laub, Director, National Institute of Justice

Ms. Adrian Nicole LeBlanc, Visiting Scholar, New York University

Professor Tracey Meares, Walton Hale Hamilton Professor of Law, Yale Law School

Chief Constable Peter Neyroud, Chief Executive, National Policing Improvement Agency (U.K.)

Christine Nixon, Chair, Victorian Bushfire Reconstruction and Recovery Authority (Australia)

Chief Richard Pennington, Atlanta Police Department

Mayor Jerry Sanders, City of San Diego

Professor David Sklansky, Professor of Law, Faculty Co-Chair of the Berkeley Center for Criminal Justice, University of California, Berkeley, School of Law

Mr. Sean Smoot, Director and Chief Legal Counsel, Police Benevolent and Protective Association of Illinois

Professor Malcolm Sparrow, Professor of Practice of Public Management, Kennedy School of Government, Harvard University

Chief Darrel Stephens, Charlotte-Mecklenburg Police Department (retired)

Professor Christopher Stone, Guggenheim Professor of the Practice of Criminal Justice, Kennedy School of Government, Harvard University

Mr. Jeremy Travis, President, John Jay College of Criminal Justice

Mr. Rick VanHouten, President, Fort Worth Police Association

Professor David Weisburd, Walter E. Meyer Professor of Law and Criminal Justice; Director, Institute of Criminology, Faculty of Law, The Hebrew University and Distinguished Professor, Department of Criminology, Law, and Society, George Mason University

Dr. Chuck Wexler, Executive Director, Police Executive Research Forum

Learn more about the Executive Session at:

NIJ's website: http://www.ojp.usdoj.gov/nij/topics/law-enforcement/executive-sessions/welcome.htm
Harvard's website: http://www.hks.harvard.edu/criminaljustice/executive_sessions/policing.htm