

The Ten-Step Guide for Conducting In-House Experimental Evaluations

A guide for law enforcement practitioners

Renée J. Mitchell, Ph.D.

Sacramento Police Department

Cody W. Telep, Ph.D.

Arizona State University, School of Criminology and Criminal Justice

Cynthia Lum, Ph.D.

George Mason University, Center for Evidence-Based Crime Policy

Department of Criminology, Law and Society



The Center for Evidence-Based Crime Policy
George Mason University

CITATION FOR THIS REPORT: Mitchell, R.J., Telep, C.W., and Lum, C. (2017). *The Ten-Step Guide for Conducting In-House Experimental Evaluations*. Fairfax, VA: Center for Evidence-Based Crime Policy, George Mason University.

*This guide was created as part of the **Matrix Demonstration Project**. For more information on the project and the Center for Evidence-Based Crime Policy at George Mason University, go to <http://cebcp.org/evidence-based-policing/the-matrix/matrix-demonstration-project/>.*

This project was supported by Grant No. 2011-DB-BX-K012 awarded by the Bureau of Justice Assistance. The Bureau of Justice Assistance is a component of the Office of Justice Programs, which also includes the Bureau of Justice Statistics; the National Institute of Justice; the Office of Juvenile Justice and Delinquency Prevention; the Office for Victims of Crime; the Community Capacity Development Office; and the Office of Sex Offender Sentencing, Monitoring, Apprehending, Registering, and Tracking. Points of view or opinions in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

Table of Contents

| | | |
|---|--|----|
| 1 | The Purpose of this Guide | 1 |
| 2 | What is Evidence-Based Policing?..... | 2 |
| 3 | What Role do Experiments Play in Evidence-Based Policing? | 5 |
| 4 | The Sacramento Hot Spots Experiment..... | 9 |
| 5 | The Ten-Step Guide for Conducting an In-House Experiment..... | 11 |
| | Step 1: Deciding to evaluate..... | 11 |
| | Step 2: Deciding what to evaluate | 12 |
| | Step 3: Creating an initial working group with key stakeholders..... | 15 |
| | Step 4: Anticipating challenges | 17 |
| | Step 5: Garnering support for the experiment specifically and for research generally | 18 |
| | Step 6: Planning and strategizing for the experiment..... | 22 |
| | Step 7: Designing the experiment..... | 26 |
| | Step 8: Ensuring successful experiment implementation | 29 |
| | Step 9: Ensuring successful data collection and analysis | 30 |
| | Step 10: Completing the experiment..... | 31 |
| 6 | Challenges and Frequently Asked Questions | 34 |

1 The Purpose of this Guide

Law enforcement agencies are becoming more and more interested in how to implement evidence-based policing and be more attuned to the benefits of research and evaluation. As discussed in Section 2 below, this effort involves both the use of high-quality research as well as developing a culture of self-evaluation.

While there are many forms of scientific evaluation that police agencies can use to evaluate the impact of a particular program (whether that impact is on crime prevention, community relations, or internal improvements), experiments are recognized as among the most rigorous approaches to address the question of program effectiveness. This guide, written as part of the Matrix Demonstration Project (funded by the Bureau of Justice Assistance [BJA]), leads agencies through one practical example of an experiment used to examine the impact of directed patrol using the Koper Curve principle in one agency in California (Sacramento).

We encourage agencies to use this guide *in addition to* partnering with researchers and to supplement the guide with structured evaluation textbooks such as those by Boruch (1996); Weiss (1997); Shadish, Cook, and Campbell (2001); or Rossi, Lipsey, and Freeman (2004). The guide is not all-inclusive, but it does point out a number of possible rewards and challenges of agencies conducting their own experimental evaluations and discusses a number of practical issues for agencies interested in conducting a randomized trial. We have also included a number of resources on evidence-based policing and experimentation at the end of each section that follows.¹

More information on the Matrix Demonstration Project (Principle Investigators: Cynthia Lum and Christopher Koper) can be found at <http://cebcp.org/evidence-based-policing/the-matrix/matrix-demonstration-project/>.

For more information on textbooks on evaluation and experimentation:

- Boruch, R. F. (1996). *Randomized experiments for planning and evaluation: A practical guide*. Applied Social Research Methods Series vol. 44. Thousand Oaks, CA: Sage Publications.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach*. 7th ed. Thousand Oaks, CA: Sage Publications.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.
- Weiss, C. H. (1997). *Evaluation: Methods for studying programs and policies*. 2nd ed. Upper River Saddle, NJ: Prentice Hall.

¹ We have included many links in the sections that follow that we recognize may change or be removed after the publication of this report. All links were live as of April 2017.

2 What is Evidence-Based Policing?

Evidence-based policing is a law enforcement perspective and philosophy that asserts that police can benefit from using the products, processes, and knowledge from high-quality research, evaluation, and crime analysis in various aspects of law enforcement decision-making (Lum and Koper, 2017; Sherman, 1998). Evidence-based policing can include paying attention to and using the results of studies that examine what types of police approaches are effective in reducing crime, improving job satisfaction, reducing stress, increasing workforce efficiency, or managing fleet services. Evidence-based policing also supports the use of rigorous crime analysis and survey methods to better understand crime, fear of crime, police legitimacy, or internal affairs. Evidence-based policing incorporates all sorts of research evidence, from evaluations of interventions to criminological research, to descriptive information. Further, it does not presume that such knowledge can influence every issue in law enforcement. However, evidence-based policing does suggest that when research is used, it should be the best research available (Sherman, 1998) and it should be translated into useable forms that are meaningful to the police (Lum, 2009; Lum et al. 2012; Lum and Koper, 2017).

The idea of using objective scientific information and criteria to inform public policy and agency decision-making reflects the important value in modern democracies of accountability for governmental actions and spending. Although this principle is mirrored in many social arenas, it has become especially significant in the fields of public health and medicine. Indeed, there are many requirements, laws, and liabilities stipulating that believable and rigorous scientific testing and replication must support medical treatments and remedies sold to the public and that harmful side effects or null effects also be reported.

Evidence-based crime policy and evidence-based policing support similar expectations. Evidence-based policing emphasizes that law enforcement should at least use the knowledge already known about policing interventions and organizational practices, as well as crime analysis generated in their own jurisdictions, to make strategic, managerial, and operational decisions. In other words, rather than rely on tradition, past practice, or even hunches and experience alone, police should deploy resources and personnel in ways that achieve sought-after outcomes, whatever those outcomes might be (e.g., crime or fear reduction; legitimate, fair, and respectful treatment; police responsiveness). While traditional, procedural, or political decision-making models could continue to be valuable, evidence-based policing suggests that decisions should also be influenced by high-quality information, science, and analysis.

Evidence-based policing is as much about generating high-quality research, knowledge, and analysis about law enforcement tactics, crime, or internal affairs as it is about *using* and *institutionalizing* research into daily police practices. There are a number of organizational, cultural, and procedural challenges in institutionalizing the use of research in practice in policing. However, three points should be remembered about this approach to policing, as Lum and colleagues point out: (1) evidence-based policing is a decision-making *perspective*, not a panacea; (2) it is grounded in the idea that policies and practices should be *supported by* research evidence and analytics, not blindly determined by them; and (3) it suggests that research is *not ignored*, and that it at least becomes a *part of the conversation* on what to do about reducing crime, increasing legitimacy, or addressing internal problems (Lum and Koper, 2017: p. 20).

For more information about evidence-based policing and evidence-based crime policy:

Evidence-Based Policing training, resources, and videos. Center for Evidence-Based Crime Policy, George Mason University. <http://cebc.org/evidence-based-policing/resources-tools/>

- “Evidence-based policing, the basics” (4-chapter, police academy video training module): <https://www.youtube.com/playlist?list=PLA5F09BA4A10764D6>.
- “2012 CEBCP Evidence-Based Policing Leadership Workshop” (4 videos): <https://www.youtube.com/playlist?list=PL4E509820FD3010E9>.
 - Workshop workbook: <http://cebc.org/wp-content/evidence-based-policing/PolicingWorkshopWorkbook2012.pdf>.
- “2011 CEBCP Evidence-Based Policing Workshop” (13 videos): <https://www.youtube.com/playlist?list=PL85DBF4AC525688DC>.
 - Workshop workbook: <http://cebc.org/wp-content/evidence-based-policing/Policing2011Workbook>.
- “2014 Evidence-Based Policing Training for First Line Supervisors” (8 videos) <https://www.youtube.com/playlist?list=PLoagclHgvlin4vK1bM7DMXPBmeWX69IT>.
- Evidence-Based Policing Matrix website: <http://policingmatrix.org>.
- Matrix Demonstration Project website: <http://cebc.org/evidence-based-policing/the-matrix/matrix-demonstration-project/>.
- What Works in Policing? website: <http://cebc.org/evidence-based-policing/what-works-in-policing/>.
- Evidence-Based Crime Policy Resources (Lum, 2011): http://cebc.org/wp-content/onepagers/EBCrimePolicySources_Lum.pdf.

American Society of Evidence-Based Policing <http://americansebp.com/home/>.

Bueermann, J. (2012). Being smart on crime with evidence-based policing. *NIJ Journal*, No. 269 (March). Available at: <http://nij.gov/journals/269/Pages/evidence.aspx>.

Coalition for Evidence-Based Crime Policy. (2012). How to read research findings to distinguish evidence-based programs from everything else. Five part video series on evidence-based policy and how to read a study. <http://coalition4evidence.org/help-desk/workshop/videos/>.

Why evidence-based policy video: <http://coalition4evidence.org/rationale-for-evidence-based-policy/>.

Lum, C. (2009). Translating police research into practice. *Ideas in American Policing*. Washington, DC: Police Foundation. Available at: <https://www.policefoundation.org/publication/translating-police-research-into-practice/>.

Lum, C. & Koper, C. (2017). *Evidence-Based Policing: Translating Research into Practice*. Oxford, UK: Oxford University Press.

Lum, C., Koper, C., & Telep, C. W. (2011). The evidence-based policing matrix. *Journal of Experimental Criminology*, 7, 3–26. Available at: <http://cebc.org/wp-content/evidence-based-policing/the-matrix/MatrixPaperJEC.pdf>.

- Lum, C., Telep, C., Koper, C., & Grieco, J. (2012). Receptivity to research in policing. *Justice Research and Policy*, 14, 61–95. Available at: <http://cebcp.org/wp-content/evidence-based-policing/matrix-demonstration-project/Lum-et-al-2012.pdf>.
- Palmer, I. (2011). *Is the United Kingdom Police Service receptive to evidence-based policing? Testing attitudes towards experimentation*. Master's thesis submitted to University of Cambridge, Cambridge, UK. Available by request.
- Puddy, R. W., & Wilkins, N. (2011). *Understanding evidence part 1: Best available research evidence. A guide to the continuum of evidence of effectiveness*. Atlanta, GA: Centers for Disease Control and Prevention. Available at: http://www.cdc.gov/ViolencePrevention/pdf/Understanding_Evidence-a.pdf.
“Understanding Evidence” learning modules (CDC),
<http://vetoviolence.cdc.gov/apps/evidence/>.
- Sherman, L. W. (1998). Evidence-based policing. *Ideas in American Policing*. Washington, DC: Police Foundation. Available at: <https://www.policefoundation.org/publication/evidence-based-policing/>.
- Sherman, L. W. (2013). The rise of evidence-based policing: Targeting, testing, and tracking. In M. Tonry (ed.), *Crime and Justice: A Review of Research*, vol. 42 (pp. 377–451). Chicago: University of Chicago Press. Available at: <http://cebcp.org/wp-content/evidence-based-policing/Sherman-TripleT.pdf>.
- Telep, C. W., & Weisburd, D. (2012). What is known about the effectiveness of police practices in reducing crime and disorder? *Police Quarterly*, 15, 331–357. Abstract available at: <http://pqx.sagepub.com/content/15/4/331>.
Earlier version of this paper available at: http://johnjay.jjay.cuny.edu/files/Telep_Weisburd.pdf.
- Weisburd, D., & Eck, J. E. (2004). What can the police do to reduce crime, disorder, and fear? *The ANNALS of the American Academy of Political and Social Science*, 593, 42–65.
Available at: <http://cebcp.org/wp-content/publications/WeisburdEck04.pdf>.
- Weisburd, D. & Neyroud, P. (2011). Police science: Toward a new paradigm. *New Perspectives in Policing*. Harvard Executive Session on Policing and Public Safety. Cambridge, MA: Harvard University. Available at: <http://www.ncjrs.gov/pdffiles1/nij/228922.pdf>.

3 What Role do Experiments Play in Evidence-Based Policing?

One anchoring philosophy of evidence-based policing is that law enforcement should not only use and translate existing research into daily practice, but it should also generate the evidence base through self-evaluation. Agencies face questions daily that could benefit from evaluation. For example, the use of a new investigative model could be evaluated to see if it improves time-to-closure or arrest rates. A targeted deterrence approach in a city block that has high amounts of shootings and gun crime could be examined to see if the approach did in fact reduce crime. Even internal practices need evaluation. A temporary adjustment of shift schedules could be examined for its effects on officer satisfaction and fatigue. Or, a new approach to monitor overtime use or sick leave might be examined for whether abuses are mitigated. A new policy for vehicular pursuits might be evaluated in terms of reduced accidents and injuries. These evaluations would add to the general research knowledge in policing and, more importantly for practitioners, benefit a city by reducing policing costs.

A major part of generating this evidence base for decision-making involves researchers and practitioners (officers, analysts, civilians) working together to evaluate an agency's practices. Almost all of the evaluations of police practices in the Evidence-Based Policing Matrix were generated in this way.

Agencies that seek partnerships with researchers and include evaluation in their strategic planning reflect one evidence-based policing notion that law enforcement agencies should take ownership of research and evaluation (Weisburd and Neyroud, 2011). Such ownership helps to institutionalize evaluation as a regular organizational practice (Lum and Koper, 2017). While researchers may bring resources and expertise to the table that agencies do not have, sometimes research partners might not be available or funded to evaluate a specific agency program. They often do not have the authority, institutional knowledge, leverage, or data access to implement changes in operational or managerial practices to conduct evaluations quickly or smoothly. Evaluations that begin in-house (of course, that could involve research partners) could generate a more dynamic approach to agency strategizing and assessment.

Conducting in-house evaluations has also become important during a time of fiscal austerity and reduced resources. In the current economic and political environments, agencies can no longer sustain enforcement efforts that do not lead to reductions of crime and fear, improvements in their interactions with citizens, or enhancements to officer work environment and satisfaction (Veigas and Lum, 2013). Agencies are becoming more attuned to accurately understanding the benefits of certain technologies before investing more money into them. Agencies have also seen the benefits of developing the capacity to carry out their own evaluations and have increased their crime analytic capacity. Federal agency support for building this capacity has also grown. The Bureau of Justice Assistance's "Smart Policing Initiative" and its interest in building crime analysis capacity in local agencies are two examples of a federal funding project that supports in-house evaluations of interventions, encouraging law enforcement agencies to not only use existing research, but also evaluate the effectiveness (including cost-effectiveness) of their practices. Non-profit foundations have also been supporting the use of in-house evaluations. The

Laura and John Arnold Foundation, for example, released a funding solicitation to support low-cost randomized controlled trials.²

There are many ways law enforcement agencies can evaluate whether a program, practice, or intervention in policing leads to desired outcome—many of which are not covered in this single guide. However, not all of these approaches are equally reliable. Implementing a crime control crackdown and simply measuring the before and after crime statistics at the location in which the crackdown was implemented might make one believe that the crackdown "worked" if crime was reduced. But crime might have gone down even without the program. Or, perhaps crime was declining in the county or city already, and the decline in the area measured was actually *less* than the overall jurisdiction's decline. Or, perhaps crime went down not because of the crackdown but some other program that had been in place during the crackdown. Weak evaluations in which alternative explanations are not considered (and controlled for) could lead to deceiving results about the effectiveness of law enforcement activities. Indeed, in police research more generally, we now know that evaluations with lower internal validity—a notion pointing to the believability of the results based on the methods used—are more likely (and perhaps misleadingly) to show optimistic results (Weisburd et al., 2001).

Stronger evaluations use control groups or statistical modeling to see whether a policing tactic really led to the effects experienced. Control groups allow tactics, strategies, and deployments (or internal adjustments) to be made on some units and not others. In this way, intervention outcomes can be measured and compared among places or individuals who did and did not receive the intervention. Even within this type of evaluation, there are good and not-so-good practices. Implementing an intervention in Neighborhood A and not in Neighborhood B and then comparing the outcomes to determine the effectiveness of the intervention is indeed an experimental evaluation. However, the two neighborhoods might be vastly different. For example, Neighborhood A might have had less crime, more neighborhood cohesion, less poverty, or may have been more amenable to the intervention than Neighborhood B. Thus, a very positive outcome in Neighborhood A might result from this evaluation not because the intervention actually works, but because it works only in neighborhoods that are very amenable to such programs or that are already on the mend.

Much like the medical field, social scientists are beginning to use randomized controlled experiments to increase the believability of the findings of a study. For those interested in extra reading on experimental methods, citations are provided below. In short, randomized experiments are considered the "gold standard" of research evaluation because they provide the best estimate of whether or not a treatment or intervention works in reducing crime and disorder. These experiments achieve this by randomly assigning recipients of an intervention to a control or experimental group and then comparing what happens when treatment is given to one group and not the other.

Randomized experiments are actually simpler than some might believe, because they use randomization to determine if a person (or place) will end up in the treatment or the control group. When done correctly, experiments minimize the chance that there might be systematic

² See <http://www.arnoldfoundation.org/laura-and-john-arnold-foundation-announces-expanded-funding-for-low-cost-randomized-controlled-trials-to-drive-effective-social-spending/>.

differences between the treatment and control groups that may lead to a particular result.³ Randomized experiments are not always possible in evaluating police interventions, and evidence-based policing is not simply about conducting randomized experiments. However, when they can be implemented, experiments can provide us with highly believable findings about the impact of police interventions or organizational practices.

For more information on experiments in policing:

Chalmers, I. (2003). Trying to do more good than harm in policy and practice: The role of rigorous, transparent, and up-to-date evaluations. *The ANNALS of the American Academy of Political and Social Science*, 589, 22–40. Abstract available at: <http://journals.sagepub.com/doi/abs/10.1177/0002716203254762>.

Coalition for Evidence-Based Policy. (2010). *Checklist for reviewing a randomized controlled trial of a social program or project, to assess whether it produced valid evidence*. Washington, DC: Coalition for Evidence-Based Policy. Available at: <http://coalition4evidence.org/wp-content/uploads/2010/02/Checklist-For-Reviewing-a-RCT-Jan10.pdf> ; Key elements of a well-conducted randomized controlled trial video: <http://coalition4evidence.org/video-key-elements-rct/>.

Justice Research and Statistics Association. (2011). *Is this a good quality outcome evaluation report? A guide for practitioners*. Washington, DC: Bureau of Justice Assistance Center for Program Evaluation and Performance Measurement. Available at: https://www.bja.gov/evaluation/reference/Quality_Outcome_Eval.pdf.

Morgan, A., & Homel, P. (2013). *Evaluating crime prevention: Lessons from large-scale community crime prevention programs*. Trends & issues in crime and justice No. 458. Canberra: Australian Institute of Criminology.

Sherman, L. (2003). Misleading evidence and evidence-led policy: Making social science more experimental. *The ANNALS of the American Academy of Political and Social Science*. 589, 6–19. Abstract available at: <http://ann.sagepub.com/content/589/1/6.short>.

Ward, K., Chibnall, S., & Harris, R. (2007). *Measuring excellence: Planning and managing evaluations of law enforcement initiatives*. Washington, DC: Office of Community Oriented Policing Services, U.S. Department of Justice. Available at: <https://ric-zai-inc.com/Publications/cops-p129-pub.pdf>.

Weisburd, D. (2000). Randomized experiments in criminal justice policy: Prospects and problems. *Crime and Delinquency*, 46, 181–193. Abstract available at: <http://cad.sagepub.com/content/46/2/181.short>.

Weisburd, D. (2003). Ethical practice and evaluation of interventions in crime and justice: The moral imperative for randomized trials. *Evaluation Review* 27, 336–354. Abstract available at: <http://erx.sagepub.com/content/27/3/336.short>.

³ See David Weisburd’s presentation on experiments and science in policing at the 2011 CEBCP Evidence-Based Policing Workshop: <http://youtu.be/P8DY-W2ymJ8> View his slides from the workshop: <http://cebcp.org/wp-content/evidence-based-policing/Policing2011Workbook.pdf#page=119>.

- Weisburd, D. (2005). Hot spots policing and criminal justice research: Lessons from the field. *Annals of the American Academy of Political and Social Science*, 599, 220–245. Abstract available at: <http://ann.sagepub.com/content/599/1/220.short>.
- Weisburd, D. (2011). Agency-led evaluations of effectiveness of deployment strategies. Presented at the 2011 CEBCP Evidence-Based Policing Workshop. Video available at: <http://youtu.be/P8DY-W2ymJ8>.
- Weisburd, D., Lum, C. and Petrosino, A. (2001). Does research design affect study outcomes in criminal justice? *The ANNALS of the American Academy of Political and Social Science*, 578, 50–70. Available at: <http://cebcp.org/wp-content/publications/Does%20Research%20Design%20Affect%20Study%20Outcomes.pdf>.

4 The Sacramento Hot Spots Experiment

This guide provides a ten-step process for law enforcement agencies to carry out their own experimental evaluations based on an actual experiment that was conducted and led in-house by a police sergeant in the Sacramento, California, Police Department (the first author). This activity was unique; although researchers from the Center for Evidence-Based Crime Policy (CEBCP) at George Mason University (GMU) provided guidance, Sgt. Mitchell and the officers and supervisors of the Sacramento Police Department (SPD) were responsible for carrying out the project and collecting the data needed to evaluate the hot spots intervention.

Sacramento is a fairly large urban environment with a variety of crime problems. Mitchell and the SPD were interested in determining whether targeted deterrent patrols in short time bursts (15 minutes) could significantly reduce crime in those targeted places. This study was propelled by a straightforward and commonly shared interest among police agencies across the United States: to reduce violent crime problems in high crime places or “hot spots.” While hot spots policing had been generally shown to be an effective strategy for reducing crime, Mitchell and the SPD wanted to determine what type of patrol strategy would be useful with limited resources.

The SPD experiment began with Mitchell reaching out to Cynthia Lum and David Weisburd at GMU about designing a study in-house.⁴ Lum drew the idea for the SPD experimental design from a recent experiment she had conducted in Northern Virginia using license plate readers in hot spots. She suggested using the “Koper Curve” principle as a type of hot spots policing based on Christopher Koper’s 1995 study. This approach suggests that officers can optimize their impact on crime by visiting hot spots in short periods of time (15 minutes or so) on an unpredictable basis, in contrast to placing officers in hot spots for long periods of time.

Through more conversations, the research partners assisted Mitchell with the design of the experiment. The goal was to maximize the deterrent ability of police to address crime on high crime street blocks by both increasing presence in medium-length visits and randomly changing the timing and order of these stops to keep offenders uncertain about when officers would return. Forty-two hot spots were paired for the experiment, and half were randomly selected to receive hot spots policing using the Koper Curve principle, while the other half continued to receive standard policing deployment as before during the 90-day experimental period. The detailed description of this study and its results are reported in Telep, Mitchell, and Weisburd (2014).

The Sacramento experiment showed significant overall declines in both total calls for service and Part I crime incidents in the treatment hot spots compared to the control hot spots both in the period of the evaluation as well as compared with the previous year. There was some variation in results across the hot spot pairs; not every treatment hot spot showed a crime reduction, but overall, the intervention was associated with a decline in calls and serious incidents.

Most importantly, the SPD carried out this study without any outside funding. In an era of limited economic resources for policing, this experiment suggests a model by which police agencies can take ownership of science and oversee the implementation and evaluation of evidence-based interventions. Along the way, Mitchell and the SPD garnered a number of lessons about doing such in-house experiments. In the next chapter, we describe these lessons

⁴ Academic centers and researchers are often willing partners for police in these informal ways (see the “e-Consortium” at <http://gmuconsortium.org> for a list of researchers near you).

learned by outlining ten important steps to help agencies design and implement their own experimental evaluations.

For more information on hot spots policing, police-researcher partnerships, and the Sacramento experiment:

- Alpert, G. P., Rojek, J., & Hansen, A. (2013). *Building bridges between police researchers and practitioners: Agents of change in a complex world*. Washington, DC: National Institute of Justice, U.S. Department of Justice. <https://www.ncjrs.gov/pdffiles1/nij/grants/244345.pdf>.
- Braga, A. A., Papachristos, A. V., & Hureau, D. M. (2012). Hot spots policing effects on crime. *Campbell Systematic Reviews*, 8(8). <https://www.campbellcollaboration.org/library.html>.
- Koper, C. (1995). Just enough police presence: Reducing crime and disorderly behavior by optimizing patrol time in crime hotspots. *Justice Quarterly*, 12, 649–672. One page summary available at: <http://cebc.org/wp-content/onepaggers/KoperHotSpots.pdf>.
- Koper, C. (2013). Putting hot spots research into practice. Presented at the 6th International Evidence-Based Policing Conference, Institute of Criminology, University of Cambridge. Video available at: <http://youtu.be/y4bB8JJcrmw>.
- Mitchell, R.J. (2017). Frequency versus duration of police patrol visits for reducing crime in hot spots: Non-Experimental Findings from the Sacramento Hot Spots Experiment. *Cambridge Journal of Evidence-Based Policing*, 1-16. DOI 10.1007/s41887-017-0002-2.
- Mitchell, R. J. (2011). Hot spots experiment in Sacramento. Presented at the 2011 Center for Evidence-Based Crime Policy Evidence-Based Policing Workshop. Video available at: http://youtu.be/HeY2_3x4D80.
- Mitchell, R. J. (2013). Research in brief: Hot-spot randomized control works for Sacramento. *Police Chief*, 80(2), 12. Available online at: <http://www.policechiefmagazine.org/research-in-brief-hot-spot-randomized-control-works-for-sacramento/>.
- Mitchell, R. J. (2012). Adapting to challenges and change. *Translational Criminology*, Fall. <http://cebc.org/wp-content/TCmagazine/TC3-Fall2012>.
- Office of Community Oriented Policing Services. (2012). A hot spots experiment: Sacramento Police Department. *Community Policing Dispatch*, 5(6).
- Rojek, J., Smith, H., & Alpert, G. (2012). The prevalence and characteristics of police practitioner-researcher partnerships. *Police Quarterly*, 15, 241–261. Abstract available at: <http://pqx.sagepub.com/content/15/3/241>.
- Telep, C. W., Mitchell, R. J., & Weisburd, D. (2014). How much time should the police spend at crime hot spots?: Answers from a police agency directed randomized field trial in Sacramento, California. *Justice Quarterly*, 31, 905–933. Abstract available at: <http://www.tandfonline.com/doi/abs/10.1080/07418825.2012.710645>.
- Weisburd, D. (2008). Place-based policing. *Ideas in American Policing*. Washington, DC: Police Foundation. Available at: <https://www.policefoundation.org/publication/place-based-policing/>.

5 The Ten-Step Guide for Conducting an In-House Experiment

Below are ten steps to help you and your organization conduct in-house experiments. We strongly encourage agencies to seek out both researchers and other law enforcement officers who have been involved in experimental evaluation for advice, guidance, and ideas.

Step 1: Deciding to evaluate

Police organizations should understand why they want to carry out an experiment, and chief executives should clearly convey this knowledge, motivation, and information to their commanders and officers to increase personnel receptivity to the task. Experiments are often carried out to answer specific questions. These questions might include: Does this strategy used by a specialized unit reduce gang violence? Does this patrol strategy reduce calls for service? Does this social media approach increase our communication and responsiveness to the community? Should we adopt this new strategy that we heard about at a recent conference? Is a particular technology useful in preventing crime?

While experiments can be used to answer many questions of interest to law enforcement organizations, they commonly have been used in policing to evaluate whether particular police programs and strategies are effective in reducing calls for service and crime incidents. Approximately 150 examples of such evaluations are available in the Evidence-Based Policing Matrix (<http://policingmatrix.org>), a database of rigorous crime control police interventions. Thirty-nine of the Matrix evaluations are randomized experiments. Research questions addressed in experiments within the Matrix include, for example: Is Drug Abuse Resistance Education (DARE) effective in reducing adolescent drug and alcohol use?⁵ Is a problem-oriented hot spots policing approach effective in reducing crime and disorder in drug markets?⁶ Do monthly newsletters with crime data provided to neighborhood residents help reduce victimization rates?⁷

Agencies also use experiments to push the envelope of police performance measures by examining *outcomes* instead of looking only at *outputs*. Traditionally, success in policing has often been measured by process-based *outputs*, such as the number of arrests made, the number of hours spent on a particular deployment, the number of guns and drugs seized, or the number of positive media reports. Experiments, however, can help to assess outcome-based performance measures. *Outcomes* can include crime reduction and prevention, improved community relations, reduced fear of crime, or improved case clearance. These performance measures are central to police agency success and cost-effectiveness. Experimental research can benefit a police organization by increasing engagement in practices that are shown to be effective in reducing and preventing crime; adjusting practices that need work; and discontinuing practices that are inefficient, ineffective, and unsustainable.

Taking the time to discuss and decide upon the reasons why an experiment will be carried out, what will be evaluated (see Step 2 below), what support is needed, as well as determining what outcomes and outputs an agency would like to measure are important considerations in deciding whether or not to carry out an evaluation. After deciding to engage in research, departments must

⁵ See <http://cebcp.org/evidence-based-policing/the-matrix/individuals/individuals-perry-et-al-2003/>.

⁶ See <http://cebcp.org/evidence-based-policing/the-matrix/micro-places/micro-places-weisburd-and-green-1995/>.

⁷ See <http://cebcp.org/evidence-based-policing/the-matrix/neighborhood/neighborhood-pate-et-al-1985-houston/>.

carefully consider if they have the time, resources, and personnel to embark on a scientific study (see Step 6). Agencies must be prepared to expend the means to continue a study, even when unforeseen circumstances (e.g. urgent community, crime, or personnel issues) consume departmental resources. When an agency is planning a project, it should include a strategy for the design, implementation, and analysis of the project. This discussion should encompass both the vision for the project, and, in a broader sense, the vision for the police department. Knowing that research is an important part of a greater plan can help command staff and line employees stay the course, even when other priorities crop up.

Step 2: Deciding what to evaluate

Many different problems, tactics, and strategies could be evaluated within any law enforcement organization. Choosing which of these to evaluate requires carefully considering the overall priorities of the organization, the potential impacts of the specific research project on the organization (both positive and negative), as well as the reasons for evaluation as described above. Being strategic about a research and evaluation agenda is just as important with regard to departmental resources, morale, and support of personnel as being strategic about other changes, innovations, or technologies being implemented in an agency. How do agencies decide what its most pressing issues are?

Taking the time to determine what to evaluate, and then to properly embed and align the research activity into the greater vision within the police department is an important step in conducting in-house experimental evaluations. Decisions about what to evaluate also mean making decisions about how the findings will be used once an evaluation has been completed. For example, will the results of this evaluation just be considered “interesting,” or will they be used to make changes in agency policies? Without a vision and leadership for institutionalizing the research findings and processes into organizational policies, the resources used to carry out the experiment may be wasted. One might consider treating study findings the way that special commission findings are often treated in agencies—with a specific commitment from leadership up front to apply the findings to current policies and practices.

We suggest actively engaging in discussions during managerial meetings and with a working group (see Step 3 below) to understand what areas of policing need to be evaluated and for what purposes. Here are a few possible options:

- The agency may wish to evaluate an existing and long-standing deployment approach (such as foot patrol in a downtown area). The research on foot patrol in terms of its crime control and also fear reduction effectiveness is mixed, and the effectiveness of the tactic seems to depend on how it is implemented (see Ratcliffe et al., 2011; Sorg et al., 2013)⁸. Perhaps the agency would like to know if changing the deployment approach of foot patrol could be more effective than its current approach.
- The agency may wish to evaluate a new deployment approach such as problem-solving in places with high levels of property crime (theft from autos, auto thefts, burglaries,

⁸ Ratcliffe, J. H., Taniguchi, T., Groff, E. R., & Wood, J. D. (2011). The Philadelphia Foot Patrol Experiment: A randomized controlled trial of police patrol effectiveness in violent crime hotspots. *Criminology*, 49, 795–831. Sorg, E. T., Haberman, C. P., Ratcliffe, J. H., & Groff, E. R. (2013). Foot patrol in violent crime hot spots: The longitudinal impact of deterrence and posttreatment effects of displacement. *Criminology*, 51, 65–101.

property damage, etc.) Experimental and quasi-experimental evaluations might then be used to evaluate the effectiveness and feasibility of a newly-implemented program.

- The agency might have an internal issue it wants to evaluate—for example, abuses in medical leave. Perhaps a new proactive policy or approach might discourage such abuses, and an agency wants use a rigorous methodology to evaluate whether this new policy works.
- The agency might want to conduct an evaluation to increase its formal relationship with a researcher or another community group. In those cases, the research topic may be chosen to maximize the benefit to both groups.
- The agency may be interested in whether adding new training on police-citizen relationships and procedural justice improves citizen perceptions of the police. An experiment could be designed in which some officers receive the training and some do not, and the views of citizens these groups of officers have contact with could be compared.

An agency's priorities and interests will likely guide the overall area of deployment or internal management and affairs chosen to be evaluated; however, agencies should also look to existing knowledge and research to sharpen their choice of what to evaluate. This can help agencies learn whether a similar intervention has been evaluated already, or what the general literature says about either the intervention or the internal policy. Consulting prior research is also useful in learning about the theories (e.g. deterrence) guiding successful interventions. Successful interventions typically have a strong theoretical backing that links the program with intended outcomes and do not rely simply on common sense. Sometimes there will be no research on the topic. Including an outside researcher in the group could be beneficial in accessing this type of knowledge quickly. Further, there are a number of freely available resources that can be useful. These include:

- *Fairness and effectiveness in policing: The evidence* (National Research Council, 2004) – This publication reviews research on police behavior, the effectiveness of police activities, and police legitimacy. A PDF of this book can be downloaded at: http://www.nap.edu/catalog.php?record_id=10419.
- The Evidence-Based Policing Matrix (<http://www.policingmatrix.org>) – The Matrix is a free translation tool developed by the CEBCP at GMU designed to help organize rigorous policing studies visually to make it easier for agencies to view not only the results of particular studies but also generalizations that can be drawn from the field of research. All randomized experiments in policing that focus on crime control outcomes are included in the Matrix.
- The Campbell Collaboration Crime and Justice systematic reviews (http://www.campbellcollaboration.org/crime_and_justice/index.php) – These are systematic reviews of large amounts of literature distilled into reports by type of intervention. A number of policing reviews include randomized experiments including those on hot spots policing (Braga et al., 2012), problem-oriented policing (Weisburd et al., 2008), and second responder programs for domestic violence (Davis et al., 2008). Search for these and other policing-related reviews at: <https://www.campbellcollaboration.org/library.html>.

- What Works Centre for Crime Reduction (<http://whatworks.college.police.uk/Pages/default.aspx/>) – The What Works Centre in the United Kingdom focuses on integrating evidence into policing practice and includes a crime reduction toolkit (<http://whatworks.college.police.uk/toolkit/Pages/Toolkit.aspx>) summarizing the research evidence on a number of crime prevention initiatives.
- National Institute of Justice (NIJ <http://www.nij.gov>) – NIJ is the main federal funding agency for criminal justice research and has a long history of funding randomized experiments in policing. Learn more about their funded research on a number of law enforcement topics at: <http://nij.gov/topics/law-enforcement/Pages/welcome.aspx>.
- National Police Research Platform (<http://www.nationalpoliceresearch.org/>) – The Platform involves researchers collaborating with a number of agencies to develop tools that can help agencies measure and monitor internal and external indicators of effectiveness. The Platform has also been involved in a randomized experiment on police training in Chicago.
- Office of Justice Programs CrimeSolutions.gov – This website reviews research in policing and other areas of criminal justice and rates programs and practices based on their level of effectiveness. Evidence from experimental studies is given greater weight over non-experimental studies in establishing effectiveness. More on policing strategies is available at: <https://www.crimesolutions.gov/TopicDetails.aspx?ID=84>.
- Bureau of Justice Assistance (BJA; <http://www.bja.gov>) – BJA provides grants for a number of policing programs and has recently become increasingly interested in funding police-researcher partnerships that include a rigorous evaluation. The Smart Policing Initiative (see <http://www.smartpolicinginitiative.com/>), for example, has funded more than 40 police agencies in efforts to implement and evaluate data-driven and evidence-based approaches.
- Groups like the Police Foundation (<http://policefoundation.org>), the International Association of Chiefs of Police (<http://www.theiacp.org>), and the Police Executive Research Forum (www.policeforum.org) also have numerous resources related to research and have conducted a number of important studies (for example, see <https://www.policefoundation.org/projects/randomized-experiments/>).
- The National Criminal Justice Reference Service (<https://ncjrs.gov/>) of the Office of Justice Programs is also a large resource for reports of many evaluations across criminal justice institutions.

When finding evaluations on topics similar to your agency’s interest, it may be useful to ask a few more questions, including: Is the agency or city similar? What was the experiment testing? Was the experiment testing the problem you are trying to solve? How were the results collected and analyzed? Can we contact the researchers for more information?

It is also important to be confident that the agency is able to carry out a study to address a specific question. For example, one might want to know whether a particular anti-gang unit intervention is effective in reducing gang activity in a city. This may require you to be able to discern what crimes and disorders are “gang activity,” know the universe of gangs or gang membership in your jurisdiction, and be able to measure “gang activity.” Some evaluations

require data and knowledge that an agency might not yet have, and a priority of the agency might then be to build the data to do an evaluation in the future. A good example of this problem occurs with “juvenile crime.” Whether an intervention reduces juvenile crime may be difficult to discern, since most crime does not result in arrest, and most suspects’ ages are unknown. Measuring the “effectiveness” in dealing with juvenile crime by examining numbers of arrest may simply measure police activity and increases in juvenile arrests, not necessarily decreases in juvenile-committed crime.

Step 3: Creating an initial working group with key stakeholders

Running a research project potentially affects the organization as a whole; administrative, training, patrol, and investigative units may all be involved, which means the project will need strong support from command staff. It is important for everyone to know that the chief and command staff support the project and will expect everyone in the organization to also demonstrate their support. Chief and command staff support are imperative, as often the project lead is a sergeant or lieutenant who may be working with multiple units throughout the project, dealing with peers, and attempting to get compliance from supervisors, line-level officers, analysts, and civilian staff. This is difficult for any project lead, but with the backing of the chief and command staff, the process will be smoother.

A chief executive should initially establish a working group consisting of key stakeholders, which will be crucial for implementing an experimental evaluation. An experiment is more likely to succeed when representatives from multiple units within the department come together with external stakeholders to discuss and plan the experiment. Bringing various perspectives to the table helps ensure that the experiment will focus on something of concern to multiple units in the agency and will be designed in a way that will make implementation feasible, especially if an experiment requires significant operational change. We discuss below some of the key stakeholders that should be part of this working group. In Step 5, we discuss in more detail external stakeholders who can join the working group as the project plan develops.

In addition to executive command and those who actually initiated the project idea (who sometimes might not be part of the command staff), first line (i.e. patrol) officers and supervisors should also be part of the working group, particularly when the intervention will be implemented at the patrol level. By incorporating key agency personnel, you can prevent creating a research design that does not have practical application. If the design calls for a strategy or practice that officers cannot assimilate into their daily activities due to other demands, then no matter how solid the research design is, it will fail. The group that will be most responsible in implementing the intervention should be present.

Another key group that should be in the working group is the agency’s crime analysts, who will be needed to assist with data analysis and the evaluation before, during, and after the experiment. These individuals have in-depth knowledge of how to access data, analyze crime trends, and display patterns, and they will be key in developing outcome and performance measures to be used in the experiment. It can also be useful to include dispatchers or other civilians who interact with the public (and thus have an understanding of community concerns) and Information Technology (IT) staff who will have an understanding of the department’s software and computing capabilities.

Including a stakeholder who has expertise in research design in the working group will also be critical. Several issues will arise when creating a study. An expert will keep an agency from spending time and resources generating uninformative data. But whom should the agency reach out to? One option might be to contact those who have researched the topic of interest before (for example, Mitchell contacted Weisburd, Lum, and Koper, who had conducted studies on implementing the Koper Curve or hot spots policing using experiments). However, some experts may not be available or geographically nearby. Location of a research partner could be very important; a university partner who is nearby means easier and more frequent interaction that may be needed during all phases of the study. As noted above, check the e-Consortium for University Centers and Researchers for Partnership with Justice Practitioner (<http://gmuconsortium.org/>) to search for universities and research centers interested in partnering with criminal justice agencies.

It is also important to consider non-police or research stakeholders in this initial working group. Having stakeholders from external agencies who might be involved somehow in either the implementation or the outcomes of an experimental evaluation can help develop a more targeted intervention to evaluate and also improve communication about the study to citizens. Initial working groups may include a few key individuals from relevant groups such as:

- The mayor's office or city council.
- City administrative and information technology units (ITU). Many law enforcement agencies do not have their own ITU staff but have access to their city department ITU or Geographic Information System (GIS) staff. If this is the case, then these groups will have to be aware of the requirements for the study and have the ability to incorporate the workload into their daily activity.
- City park staff; park rangers; or bus, light rail, or train employees, especially if you are working on a study that involves these or other related groups.
- Public works agencies who might be needed to implement certain interventions like Crime Prevention Through Environmental Design (CPTED) or nuisance abatement.
- State criminal justice agencies, which might be able to fund small projects.
- For agencies under a consent decree, the Department of Justice will be an integral stakeholder.
- Neighborhood activists and leaders, citizen groups, and other community organizations.

As an example of how the working group model can work in practice, when the Portland Police Bureau (PPB) decided to test a hot spots program, its first meeting consisted of their chief, the researchers from Portland State University, the head of the strategic services division, the crime analysis sergeant, crime analysts, and an external consultant (Mitchell) from an agency who had implemented a hot spots experiment. These stakeholders were involved in the initial component of creating and designing the study to ensure that everyone was on the same page as to what question the research was trying to answer, what resources would be needed, and who would be involved in the project. The project lead was the crime analysis sergeant.

After working through some kinks in the data systems at PPB, the sergeant's next step was to reach out to the stakeholders in patrol who would be implementing the experiment. These were the lieutenants, sergeants, administration staff, etc. that would be responsible for following research protocols. These stakeholders were pivotal in the success of the research. If research protocols are not followed, everyone's time, energy, and money could be wasted. Once PPB determined the type of study design, the study question, and where the study was going to be carried out (see more on this below), the crime analysis sergeant identified who would be impacted by the research. They created two working groups—one at the chief level, which would meet every month to make sure the project stayed on track, and one at the line level, which would meet weekly. The line-level working group would consist of the patrol sergeants, lieutenants, and commanders whose districts were affected, along with the Portland State University researchers, the Sacramento sergeant, and crime analysts. This working group would involve the stakeholders who were in the trenches operationalizing the field research. The crime analysis sergeant also conducted three different meetings during the study to obtain feedback from the line-level officers.

Although not used in policing as often as in other arenas, the working group model can increase the amount of perspectives about agency priorities and can increase buy-in to research more generally.⁹ The working group should try to find common themes throughout the various units and hierarchies in the organization that can help lead to a problem or group of problems the agency can address. The working group could also focus on how the group might address or alleviate the challenges of implementing a study and could discuss how results might be used.

Step 4: Anticipating challenges

Conducting any type of research can be fraught with challenges. When thinking about how a study will work, the working group should try to anticipate some of the challenges ahead of time and proactively address them.

Many of these potential challenges are discussed in two publications on implementing evidence-based practices. The National Registry of Evidence-Based Programs and Practices has put together a useful list of questions to ask as you explore the possible implementation of an intervention in your agency. These are available at:

http://www.nrepp.samhsa.gov/docs/Questions_To_Ask_Developers.pdf. We revisit some of these questions in Step 6. The Coalition for Evidence-Based Policy has also developed a guide for successfully implementing evidence-based programs with a series of recommendations available at: <http://coalition4evidence.org/wp-content/uploads/2012/12/PublicationHowToSuccessfullyImplement06.pdf>.

Some questions to consider include:

- Have we identified the reason why we are carrying out this study, and how will we clearly convey this to the organization in a positive fashion?

⁹ Useful examples of the working group model can be found in some of the Project Safe Neighborhood interventions. See the Project Safe Neighborhoods toolkit at: https://www.bja.gov/programs/psn/psn_toolkit.pdf.

- If the study is led by lower-ranking officers or units, is leadership (i.e., the chief/top commanders) supportive of the study, and what commitments are given as to using the results of the study?
- If the study is led by high-ranking officers or the chief, are line officers and first-line supervisors on board?
- Are there enough time, resources, and personnel to complete the study? You should decide on a length of time to run the study (e.g. 90 days, 120 days, 6 months, etc.) that is academically sound but also practical for the agency (see more on this in Step 7). It is also best if a project lead is assigned to the project full-time.
- Can the agency devote officers to the experiment (i.e. should officers be assigned full-time to implement a strategy or program)? If the experiment is not evaluating existing deployments but an innovation, can that innovation be incorporated into existing deployment approaches?
- Is the crime analysis unit suitably staffed and trained to deal with the demands of an ongoing research project? Will extra staff be needed for crime analysis to keep up with its regular duties and the study?¹⁰
- Does the agency have support from a researcher who can provide guidance on experiments?
- What outcomes do you want to measure in the research? How will these outcomes be measured?
- Will there be resistance to the study from community groups, politicians, or groups of officers? How might such concerns be addressed and alleviated before the study begins?

Step 5: Garnering support for the experiment specifically and for research generally

One significant challenge of doing anything new or innovative in a police agency is the possibility of low support, cynicism, and resistance to implementation by groups inside and external to the agency. While many challenges may be anticipated, this particular challenge can determine if an experimental evaluation can be successfully implemented. Inconsistent levels of support or conflicting levels of support across units, commanders, and supervisors can sabotage any evaluation (or innovation in policing, for that matter). Even if a strong working group is established, garnering support for the experiment, specifically, and for the idea of research generally, should be a priority.

In this step, we discuss a variety of areas where support is needed and possible ways to garner this support.

¹⁰ The CEBCP has produced two training modules on crime analysis that provide a more extensive introduction to the importance of crime analysis for effective policing. See “Crime Analysis for Operations” (<http://www.youtube.com/playlist?list=PLoagclcHgvIj9SJAHObyWzvtDcf3B1fLi>) and “Crime Analysis for Commanders” (<http://www.youtube.com/playlist?list=PLoagclcHgvIi4zkWDi0ZJ829bQL4fbTvP>).

Leaders

Good organizational leadership is the driving force behind any successful research project. Thus, the chief executive's support is absolutely necessary for experimental evaluations, and he or she should communicate this support clearly and often in positive, enthusiastic ways. Expressing this support is especially important when the interest in evaluation does not come from chief executives, but from analysts or lower-ranking supervisors who lack the similar authority and leadership as the chief executive. If the chief executive leaves, the study should be a part of the regular briefing to the incoming chief executive about the goings-on in the department. Most importantly, if the chief executive supports the research, so should his or her executive team. A number of projects have failed because some of the members of a management team openly opposed research that a chief initiated or supported. Unlike ad-hoc projects or short-term solutions, evidence-based policing demands rigorous leadership as it often challenges long-term cultures and traditions and uses unfamiliar ideas of accountability and evaluation in a highly traditional institution.

Part of this leadership requires bringing every level of the organization on board to understand and support the project. Just one officer who refuses to carry out the deployment or a supervisor who grudgingly conveys the assignment to his or her squad can thwart an evaluation. Analysts who refuse to help collect data or commanders who actively criticize or put down the project will also contribute to an evaluation's failure, both during implementation and when results are being translated into daily practice. Without a unified approach, weaknesses in the research design could be exploited, and officers, supervisors, and other staff could undermine the research project.

Avoiding resistance to evaluation, research, and any organizational change cannot be successfully done with orders to comply. However, in hierarchal organizations, line employees and first-line supervisors are often told the "what" and the "how" but are rarely inspired by the "why." Chief executives play a pivotal role in working with officers, supervisors, staff, and citizens to understand why an intervention is being carried out as well as their role in its success. Further, it is important that these explanations and direct support come from the Office of the Chief Executive/Sheriff rather than a particular unit to avoid the targeting of a particular group or unit within an organization for resistance.

Some suggestions for a chief executive in garnering support for research include:

- Inspire the officers by letting them know why the research is important for the organization and why the chief is personally interested. Why is it important to the policing profession? Why is the project important to them? Will it increase officer safety, reduce calls for service, etc.? Why do police try to improve their performance at all? Why does the way we police affect our community?
- Talk to supervisors and command staff about the way you want to convey the research. Be specific about the type of language that should be avoided. For example, a statement such as "Hey guys, I don't really know why the chief wants us to do this, and I don't want to do it, but if he says to do it, then that's what we need to do" could significantly reduce support of any innovation or evaluation by an implementing unit.
- Set clear expectations for your organization about who will carry out the evaluation and who the results of the evaluation will affect. Develop a timeline as to when the

project will begin and end. Develop plans ahead of time as to how to use the evaluation results.

- Set clear goals for the project. Be clear as to which units or staff members are responsible for specific aspects of the project. This might include reducing the workload for certain units that are also responsible for delivering products for the evaluation. Crime analysts who are already overloaded by generating statistics for Compstat meetings, for example, may not be able to also generate data and analysis for a new study.
- Train officers and prepare them for the project itself (see Step 6).
- Chief executives should also work with their community and public relations units to notify the media of research partnerships and projects.

While the chief's support is essential for an experiment to be successful, other top department leaders can also play an important role. All supervisors and commanders should convey a clear and consistent message about the reasons for the evaluation as well as express consistent support for the evaluation. Total command and supervisory support is especially needed for unusual, innovative, progressive, or non-traditional activities in policing.

Support can be communicated through an announcement at a Compstat meeting, a department-wide email, or an article in the department's newsletter. Commanders should reinforce the importance of the study every week at Compstat. Additionally, they should ask for updates from the project leader as to the progress of the study and incorporate evaluation into the regular conversation of command meetings. Maintaining scrutiny for lack of implementation or problems with following experimental protocols is important. Commanders can also reward officers and supervisors for their efforts in implementing an evaluation. Additionally, such efforts should be noted as important to promotions and performance measures.

Commanders also play an important role in overseeing the experiment's implementation with the project lead. For example, commanders can ensure the sergeants are following through with their officers and verifying that the officers are being deployed as planned. Commanders can update patrol staff as to the progress of the study in roll call and identify innovation champions (see below). If the experiment is being tested at the line-level, first-line supervisors need to also be involved and supportive, and they must understand the "what", "why", and "how" of the evaluation.

Line officers, detectives, supervisors, and civilian staff

It is important to gain buy-in from the officers from the start of the project. To achieve this, leaders should start by conveying with specificity and transparency exactly why the evaluation is being done. Officers will want to know how evaluations can benefit their everyday work and profession. For example, officers have an incentive to cooperate in a study that evaluates the benefits of different shift lengths, because changes instituted as a result of the study could increase personal time and/or improve health and stress-related outcomes. Evaluations that save line-level employees from an ineffective, expensive, resource-intensive program can increase employee morale. Interventions that are being evaluated for their crime control potential should also be carefully explained with regard to the specific connection between the intervention and crime reduction/prevention that is being implicated in the deployment.

Buy-in also requires listening to the officers' concerns about the study, their workload, whether or not the study will benefit them in the long run, etc. A department should take the time to listen and honestly address these concerns, working with officers to develop solutions. Officers and supervisors might also be included in determining how to implement the evaluation or perhaps the geographic boundaries of a study. Inclusion can increase buy-in and reduce resistance to implementation.

Opinion leaders and innovation champions

Whenever someone is introducing an innovation or experimental evaluation into an organization, there will always be resistance. One important group that can facilitate change is innovation champions, also called opinion leaders. These individuals can be informal leaders who influence the behaviors of other officers toward leaders within the organization. Identifying these individuals beforehand and involving them in the working group would go a long way towards influencing the whole organization. Research has shown that opinion leaders are capable of influencing others to adopt an innovation more so than statistics or "experts" in the field.

Innovation champions can help get other officers on board with the experiment. To do this, champions need to be trained in evidence-based policing generally, as well as the evaluation more specifically. In turn, these champions can serve to train rank-and-file officers about the research project and help develop ways to implement evaluations successfully, as well as troubleshoot problems as they arise.

Building support for research generally

It is also easier to build support for any given experiment when your agency generally supports and has experience with research and has an infrastructure in place for conducting research. Building this infrastructure can occur in a number of ways:

- Assessing agency receptivity to research. A first step can be assessing officers' current views about research and evidence-based policing. Lum and Telep designed a survey instrument that can be administered to examine the extent to which officers comprehend, use, and are open to research evidence (see <http://cebcp.org/evidence-based-policing/the-matrix/matrix-demonstration-project/receptivity-to-research/>).
- Establishing formal and long-term partnerships with researchers. In the past, these long-term partnerships have led to a series of experiments in a number of agencies including Jersey City, NJ and Minneapolis, MN.¹¹ A number of resources for building research-practitioner partnerships are available at: <http://cebcp.org/evidence-based-policing/the-matrix/matrix-demonstration-project/partnerships/>.
- Embedding a criminologist in the department. Hiring an academically trained criminologist or evaluation researcher to work in the agency part- or full-time is one way to help create a long-term partnership and shows a clear agency commitment to

¹¹ Experiments in Jersey City included the first problem-oriented policing experiment (see <http://cebcp.org/evidence-based-policing/the-matrix/micro-places/micro-places-braga-et-al-1999/>) and a drug hot spots intervention (see <http://cebcp.org/evidence-based-policing/the-matrix/micro-places/micro-places-weisburd-and-green-1995/>). Experiments in Minneapolis included the original domestic violence arrest study (see <http://cebcp.org/evidence-based-policing/the-matrix/individuals/individuals-sherman-and-berk-1984/>) and the original hot spots policing study (see <http://cebcp.org/evidence-based-policing/the-matrix/micro-places/micro-places-sherman-and-weisburd-1995/>).

consistently engaging in research. Read more about this model in Braga's (2014) Police Foundation paper on this topic where he discusses working as an embedded researcher in the Boston Police Department: <https://www.policefoundation.org/publication/embedded-criminologists-in-police-departments/>.

- Making research findings and processes a regular part of the conversation during Compstat or managerial meetings. See more on this at: <http://cebcp.org/evidence-based-policing/the-matrix/matrix-demonstration-project/rethinking-compstat/>.
- Institutionalizing research and analysis into training and everyday practice. The Matrix Demonstration Project (<http://cebcp.org/evidence-based-policing/the-matrix/matrix-demonstration-project/>) has a number of tools and suggestions that can be used to help institutionalize support for research. These include academy training modules on evidence-based policing (see <http://cebcp.org/evidence-based-policing/the-matrix/matrix-demonstration-project/academy-curriculum/>) and suggestions for incorporating research knowledge into field training (see <http://cebcp.org/evidence-based-policing/the-matrix/matrix-demonstration-project/transforming-field-training/>). For more information, see <http://cebcp.org/wp-content/evidence-based-policing/matrix-demonstration-project/Lum-et-al-2012.pdf>.

An expanded discussion of these issues can be found in Lum and Koper's *Evidence-Based Policing: Translating Research into Practice* (Oxford).

External Support

As already mentioned in Step 3, involving external stakeholders can also help garner support for research in one's agency. Research can bring legitimacy and justification to specific strategies and tactics as well as to the agency more generally. Reaching out to the mayor's office, city council, city administrative groups, neighborhood groups, community leaders, schools, other criminal justice institutions, public works, and the media is important.

Step 6: Planning and strategizing for the experiment

Once a working group has been established, support has been garnered, and the subject of evaluation has been chosen, planning for the intervention is the next crucial step for the successful implementation of any evaluation. This involves not only designing the experiment itself (Step 7), but also determining timelines for the study, personnel involved in the study, as well as the costs and resources needed.

Estimating time

New to research design, the SPD gave itself a tight deadline to test a research hypothesis and attempted to implement the study within a month of its conception. However, this timeline was unrealistic. For example, it took longer than expected to figure out where the hot spots were located in Sacramento. This delay arose out of software difficulties, staffing issues, and working through the best approach to identify hot spots. At the project's onset, crime analysts had to divide the entire city by street segments, use GIS to map crime to those segments, and also adjust GIS for common problems during mapping, including inaccurate mapping or missing data. Once the calls for service were mapped to the street segments, SPD personnel looked at all of the types of calls received over a two-year period. Analysts went through each call type and decided

whether high visibility intermittent random policing could potentially have an effect on that type of crime. If the crime was indoors and out of view (e.g., domestic violence), then that crime type was removed from the data set. Further, deciding on the exact boundaries of the hot spots was also a challenge. These issues are just some examples of why the initial timeline SPD personnel set for the project was not realistic.

The processes described in Steps 1-5 also take time, especially in terms of choosing the project team and deciding what to study. Initial information and institutionalized knowledge may be needed from various stakeholders, and garnering support from stakeholders can be challenging. Designing the study; choosing how to operationally roll out the study; and establishing what data to collect, how to collect it, and who will collect it (discussed in Steps 7-9) are also all time-consuming.

Other time-consuming efforts involve data collection and recording performance and process measures. Data needs to be continuously collected, checked, and backed-up during the research project, including the taking of meticulous notes about all aspects of the project. This process is critical not only for other agencies to replicate a successful program, but also to record why decisions were made and to ensure internal validity and reliability in the data collection and overall project. Training may be needed to help personnel with this data collection and recording to avoid mistakes or improper data collection. Further, as data is collected, it should be reviewed, and may often need to be “cleaned” for common errors such as misspellings of street names (which will reduce the ability of the crime to be geocoded) or missing data due to lack of recording by officers involved in the experiment.

Doing experiments also means apportioning time for adjustments, modifications, and other challenges. Even though an agency can meticulously map out a timeline, create procedures, and thoroughly train staff, mistakes or unintended mishaps can occur that cause staff to refocus their energy and take time away from the study itself. For example, during the SPD hot spots study, the project team was told that some officers were not initially following the protocol of implementing the hot spots tactic. Because correct implementation was central to the study, the project team had to take time to make sure that officers were implementing the study in the way intended. Data collection was immediately halted while the crime analysis sergeant notified the watch commander and area captain, emails were sent out, and officers were sent to roll calls to reinforce the importance of following the specified hot spot order. Interestingly, after examining automated vehicle locator (AVL) data, it was discovered that the belief that officers were not following the project protocol was simply a rumor. However, this is an example of how unforeseeable events can divert resources and increase the amount of time, energy, and money a department expends on a study.

Further, the initial research design may not be compatible with an agency’s existing operational strategy, software systems, or organizational culture. For the SPD project, street officers were not familiar with hot spots policing and were deployed in a traditional, area-based, reactive patrol mode. Conducting research was also relatively new to the agency, and there was natural resistance to evidence-based policing. Training officers and supervisors on something new will also take time, as will the efforts needed by supervisors to monitor the implementation of a new intervention or deployment style. Project staff will need to regularly follow-up with officers and consider ahead of time how resistance to the study or a new intervention will be handled. The chief executive plays a major role in ensuring an agency is ready for such changes.

Finally, any agency conducting an experiment or other evaluation should conduct some pre-testing before implementing the actual intervention and evaluation. Pre-testing might include trying out the intervention in small area of the jurisdiction with amenable officers and supervisors to anticipate problems with the intervention and research ahead of time. Focus groups with officers involved in pre-testing can also provide a qualitative understanding of these issues.

Choosing personnel to conduct the actual evaluation

By this stage, there will likely be a group of individuals already identified who will help to conduct the actual evaluation of the intervention, and they may be able to conduct the statistical analysis of the results of the study as well. However, you may need to identify those with specific expertise in data collection and analysis who may not have been part of the working group. Depending on the personnel and resources available, the same or different individuals may play the roles of project lead, crime analyst, information technology expert, subject matter expert, or report writer/editor.

Ideally, the project lead will have knowledge of experimental processes and will have capabilities to adequately design, implement, and analyze research. Project leads should also have some leadership role in the organization and the ability to make some decisions and assert some authority over the project when necessary. It is not necessary for the project lead to be an expert, but he or she should have an idea of what good research design looks like. A background in statistics would be useful. As noted earlier, CEBCP has created a series of training modules on evidence-based policing and crime analysis that can be useful in providing additional background for the project lead (and any personnel involved in the study) on the importance of rigorous research to improve policing.¹² A project partner, such as a university researcher, can also help the project lead in evaluating the intervention. It should be noted that outside researchers might request to attempt to publish the study in a peer-reviewed journal as an incentive for assisting an agency in their research.

Involving research and planning partners or an agency's crime analysis unit is an important ingredient to a successful evaluation project. In the case of SPD, there was a crime analyst sergeant, a patrol sergeant assigned to the unit, an officer, a community service officer, an administrative analyst, and an administrative technician working on the study. While there was enough staff to monitor the project, every employee still had their normal workload to complete on a weekly basis. This led to the crime analysis unit becoming overworked during the design, implementation, and analysis of the study, especially the administrative technician who was directly responsible for all of the data being generated. Over-working analysts for purposes of special evaluation projects is a common challenge in evidence-based policing, and we suggest that agencies consider investing in increasing their analytic units (preferably with embedded criminologists) if they intend to pursue an evidence-based approach to their strategic portfolio.

If there are not enough personnel, police agencies might consider extending the time allotted for an evaluation or reducing the scope of the experiment. Another idea is to consider partnership agreements with universities to let advanced graduate students work on the evaluation as part of their dissertations or theses. But if research and evaluation is a long-term priority (which it must

¹² See footnote 9 and <http://cebc.org/evidence-based-policing/the-matrix/matrix-demonstration-project/academy-curriculum/>.

be if an agency chooses to become more evidence-based), then investing time and resources into these goals (possibly at the expense of adding new police officers, for example) will be a difficult, yet worthwhile decision.

Resources

In addition to the planning resources discussed above as well as resources regarding the evaluation team, the largest resource for experimental evaluations will come from the officers implementing the evaluation. Experiments like the one in Sacramento can and have been implemented entirely without the addition of new officers or the use of overtime, as they work in existing operational frameworks and shift schedules. Officers may be tweaking the way they respond to an incident, or they might change what they do during their non-committed time, for example. The belief that evaluations in policing can only be done using overtime officers is simply not true.

However, officers may need to be trained on the new intervention, which can cost labor hours. Some training can be done quickly or with a simple directive or order. Other training may be more extensive and require consultants and subject matter experts to be brought to the agency. For example, hot spots experiments may only entail telling officers to go to specific locations for 12-15 minutes in between calls for service, using their discretion, experience, and existing toolkits to try and prevent and control crime in those locations. However, officers in a community legitimacy experiment may need hours of training on alternative ways to speak to community members, learn problem-solving techniques, or be trained in restorative or procedural justice techniques and concepts. Training resources needed to carry out the experiment should be considered when deciding what to evaluate.

Project managers might also consider allowing officers to suggest ways of incorporating the evaluation strategy into their daily work, which can help increase their feelings of connection and ownership of the project. It is also useful if the training provides officers at least a basic understanding of randomized controlled trials and research design. Supervisors also require training, especially on how to ensure that officers are following research protocols. These leaders are essential in ensuring that the daily implementation of an intervention being evaluated is being carried out. In particular, the project team may bring in a well-respected leader who can work with first- and second-line supervisors on leadership skills that are useful for especially difficult projects that might be met with high levels of resistance.

Further, equipment might be required for the study. Are there enough patrol or unmarked vehicles, computer terminals, and/or other equipment available for everyone involved in the study? If the study focuses on license plate readers, are those available for the evaluation? Agencies will need to ascertain whether existing software systems can download, search, and sort necessary data in an efficient manner, or whether software licenses (e.g., for crime mapping or statistical analysis) are available for all personnel involved in the study. It is essential to know your data systems. Verify that you can extract the data you want without using extensive resources to get it. Different personnel (i.e. command staff, dispatchers, patrol officers, detectives, etc.) will require different access for different use of the data. At every level of involvement throughout the organization, determine what software tools will be needed, how those tools will be used, and that stakeholders have the ability to retrieve the data needed for their level of involvement.

Step 7: Designing the experiment

In this step, we outline the most important aspects of designing an experimental evaluation. A well-designed experiment is important for ensuring that the findings of an evaluation are believable and allows a police agency to draw conclusions about the impact of an intervention or program on intended outcomes. For example, agencies often engage in large operations to control or reduce open-air drug markets, monitor newly released parolees, quell gang activity, or deal with a rash of burglaries or thefts from automobiles. Countywide or regional operations can involve multiple agencies, including the surrounding local police agencies, the Drug Enforcement Agency (DEA), the Bureau of Alcohol, Tobacco, Firearms and Explosives (ATF), probation, and parole. Often, such interventions require intensive resources and time to plan, frequently involving dozens of officers working across multiple days. These operations are common in law enforcement agencies and account for great cost. The lead agency will spend time locating suitable targets, writing an operational game plan, coordinating meetings with the involved agencies, and preparing team leads for the operation. At the end of the operation, the lead agencies may coordinate a press conference to show off the number of guns and the amount of drugs and illegal paraphernalia they confiscated. They may display the number of arrests and how many offenders were locked up, demonstrating the outputs of the operation rather than looking at the outcomes.

Do these operations help reduce crime? For the most part, such operations usually don't have an evaluation component to them, despite their costs. Agencies do not follow up to see if the operation correlates to lowered drug use, reduced recidivism of parolees, decreased gang activity, or the prevention of thefts from automobiles. Take an anti-gang operation, for example. Agencies might target identified gang members and places with high levels of gang activity using an "Operation Ceasefire" or "pulling levers" approach.¹³ Often, agencies might examine crime rates before and after the operation in their city, which is a poor way to determine whether the intervention had an effect (even if a crime reduction was noted). What if crime was already declining before the operation began? What if a blizzard immediately after the start of the operation affected gang crime citywide? What if crime was abnormally high in the month prior to the operation, so examining any other post-operation period would make it appear that crime was dropping? Because the before/after design has no comparison or control group, it is difficult to conclusively attribute changes in crime to the agency operation. As another example, the Sacramento study could have simply examined whether or not crime went down in the 21 hot spots where police focused their attention. However, crime in Sacramento was declining citywide at the time, so any intervention evaluated in this way would appear to be effective. The control group and randomized design in an experiment takes into account these trends, while just looking at data before and after does not. In Sacramento, we can be more confident that the policing strategy caused a decline in crime because we examined the randomly chosen treatment hot spots in relation to the randomly chosen control hot spots that did not receive extra police attention.

A well-designed evaluation can help better pinpoint whether the intervention caused a drop in crime, gang activity, or whatever outcome is being measured. Returning to our gang example, agencies might consider setting up a randomized controlled trial by targeting a random selection of half of the gang members on their list for the operation and letting the other half receive

¹³ See <http://www.nij.gov/topics/crime/gun-violence/prevention/Pages/ceasefire.aspx>.

typical police attention. This approach is especially useful if the agency does not have the resources to target all the gang members on their list. One outcome measured may be recidivism by these individuals in one, three, and six months following the intervention, comparing them to those who did not receive the intervention. Or, if there are specific places with high levels of gang activity, a random selection of those places might be selected for police intervention, again comparing crime rates in those specific (or surrounding) places with areas that did not receive the intervention. While there may be different ways to evaluate the impact of this anti-gang intervention, randomized experiments provide the most believable answer to questions about the effectiveness of a program or strategy. That is because experiments allow you to say with greater certainty that a program or practice *caused* a change in crime (i.e. they have a higher level of internal validity). Other research designs suffer from threats to internal validity, which make those designs less able to make statements about causality. In Step 8, we discuss some of these threats in more detail and how they can even affect an experiment if there are problems with implementation.

Randomized experiments are not always appropriate. You cannot randomly allocate the death penalty, for example. Nevertheless, in many instances in policing, randomized experiments are possible without any ethical concerns. In Sacramento, for example, the department only had available the resources to focus on about 21 hot spots. In that case, randomization is a fair way to decide how the police will use limited resources. The treatment was about giving some hot spots in the city extra attention and extra patrol. None of the hot spots was losing any police resources though. Hot spots in the control group continued to receive their normal levels of patrol.

There are several important considerations when designing an experiment as described by Sherman (2010)¹⁴:

1. *Having a suitable research question*: As described above, the first crucial step is the formulation of a research question that can be answered using a randomized trial. Randomized experiments rely on the random allocation of an intervention to at least one treatment and at least one control group (see more below) and so the research question must be designed in a way that an experiment can be used to answer it. For example, the SPD wanted to know if a new hot spots policing intervention would reduce crime. A randomized design could be used to address this question because the hot spots intervention could be randomly implemented on some hot spot street segments and not others.
2. *Deciding how/what to randomize*: The process of randomization involves assigning people or places to either a treatment (i.e. intervention) or control group using a coin, a random generator, or some unbiased way of group assignment. Usually, all units should be randomly allocated at the beginning of the study. For example, in Sacramento, the department initially identified 42 high crime hot spots. These hot spots were then paired so that each pair had similar hot spots and then in each pair, one hot spot was randomly assigned to treatment (the hot spots intervention) and one hot spot was randomly assigned to a control group (receiving standard levels of policing) using a random number generator. It is also possible to run an experiment in which cases (e.g. arrestees) are randomly allocated as they come into the system. However, this may not be feasible or desired, and so agencies new to experimentation

¹⁴ This section draws on Sherman's (2010) "An Introduction to Experimental Criminology."

- may want to focus on randomly allocating all cases at once. Places are typically easier units to work with than people are, in part because crime data is often readily available for specific streets or neighborhoods and because places are easier to keep track of than people are.
3. *Carrying out the randomization:* There are a number of different ways to randomize people or places into treatment and control groups. Perhaps the simplest way is just flipping a coin for each place and assigning the group based what side the coin lands on. Ideally, however, the treatment and control groups will be of equal or at least similar size. Just by chance, flipping a coin could lead to groups of different sizes (e.g. you could get six heads in a row). Random.org is a freely available online random number generator that can be used to generate a series of random numbers. Randomizer.org can also be used for random assignment. Computer programs with random number generators (e.g. Microsoft Excel or IBM SPSS Statistics) can also be used to assign a person or place to treatment or control and these programs can ensure that groups are of equal size. Places or people chosen for the treatment group receive the intervention or program, while control group people or places do not receive any special treatment.
 4. *Choosing appropriate outcome measures:* For most policing experiments, the focus will be on whether the intervention reduced crime. Fortunately, police agencies already routinely collect crime-related data that can be used to examine changes in crime. The SPD used three outcomes measures to assess the impact of the hot spots treatment: all calls for service, Part I crime incidents, and soft crime incidents (less serious crimes with domestic violence removed). Depending on the intervention, the agency may want to focus on a specific subset of calls or incidents (e.g. gun-related calls and incidents for a gun crime intervention or just burglary calls and incidents for a burglary reduction initiative). Other outcomes measures such as resident or offender surveys or interviews or observations of disorderly activity can also be collected but tend to be more expensive than using official crime data. If an intervention is focused on the behavior of individual offenders, frequency measures (i.e. how many times an offender reoffends) are better than simply recording if an offender recidivates. Outcome measures and what measures will be considered primary outcomes should also be determined before the study begins.
 5. *Ensuring you have a statistically powerful design:* A statistically powerful design is one that is likely to show an effect on crime and disorder (or whatever outcome you're measuring) from the intervention if one exists. Statistics use samples from the total population studied to draw inferences about the population. If a sample size is too small, the study will not show an effect, even if one exists (i.e. there will not be enough power). For example, taking 10 officers for a study on the effects of a training from an organization that has 1,000 officers might not be a large enough sample to accurately reflect the potential effects of the training on the organization as a whole. A sample size must be large enough to demonstrate a natural average of the population. To assist with this, an outside research partner can be very useful in helping to ensure you have designed an experiment that has sufficient power to show an effect. Power can often be increased by increasing the sample size or the length of

the study, so a longer experimental period or a greater number of people or places involved in the study may be useful in increasing power.

6. *Implementing the experiment and ensuring treatment is delivered appropriately and consistently* (see Step 8): After randomization, the experiment can begin. The project team should consider conducting a small pre-test prior to full implementation (see above). It is also important to monitor adherence to randomization and experimental protocols throughout the study. In Sacramento, for example, this meant confirming that officers were visiting the treatment hot spots in the specified order each day and spending about 15 minutes in each hot spot. Additionally, officer location was monitored through AVL to ensure that officers were not spending extra time in control hot spots.

Step 8: Ensuring successful experiment implementation

Several implementation issues can arise during an experiment, all of which can threaten the integrity of the study and the ability to reach reliable and valid conclusions about effectiveness when the study is complete. We discuss some of these implementation challenges below and considerations on how to address these issues if they arise.

Program fidelity and adherence to experimental protocols

Program fidelity refers to the extent to which the intervention has been implemented as expected based on the planning processes described above. Ensure compliance with experimental protocols throughout the study. A number of different approaches can be used. AVL or Global Positioning System (GPS) data can be useful for this purpose, especially if the experiment involves officers visiting specific locations, like in the SPD study. Are the officers in the location they are documenting? Personnel should use departmental technology for verification whenever possible rather than relying only on verbal confirmation. Data logs can also be used to have personnel document their activities. Volunteers (ideally from university or research organization) can be used to engage in social observation or ride-alongs. Surveys, interviews, and focus groups with officers involved in the intervention can also be used to assess adherence to experimental protocols and potential reasons for any deviations from these protocols.

To verify that the study protocols are being carried out properly, sergeants or other supervisors should be required to randomly evaluate employees' work products. During the SPD study, 10% of the hot spot calls for service were selected by a computer-generated randomizer every week. Then, a sergeant checked the officers' hot spot calls with the automatic vehicle locator to determine whether the officers were going to the correct hot spots in a randomized order. This practice ensured compliance with the study parameters.

Contamination

Contamination occurs when both the treatment group and the control group are receiving the treatment (and therefore, when you compare the groups, significant differences between them might not be apparent). This could occur, for example, if officers were conducting extra patrols in the control hot spots as well as the treatment hot spots. As noted above, monitoring the location of officers is one way to try to limit contamination. In a place-based study, it is also advisable to limit officer knowledge of the control locations. Withholding this information can

limit intentional contamination by officers who may try to provide resources to control locations. Only individuals directly involved in the evaluation need to know the location of control sites.

Attrition and staffing issues

Attrition occurs in a person-based study when individuals drop out of the experiment before the experimental period ends. This is less of an issue in a place-based study, but staff attrition can still occur during the period. Officers may be reassigned during the experiment, for example, which can create complications in ensuring the treatment is successfully delivered throughout the study. Any pre-experiment training should be filmed, if possible, so that officers brought onto the project team during the experiment can receive the same training. Agencies should also ensure that sufficient personnel will be available to carry out the experiment throughout the study period. If, for example, the agency is concerned about sufficient personnel being available in the summer because of vacations or rising crime, then the experimental period should not include summer months.

Changes in instrumentation and data systems

If possible, any forms used during the experiment should be kept consistent throughout the study. Changes in instruments, forms, or logs can create complications in data analysis. Similar problems can occur if the agency makes changes to dispatch or records management systems. If the agency is planning a change to a new dispatch or records management system, any experiments should begin after this change has been made to ensure consistency in how data is collected during the study.

Selection bias

Selection bias can occur when individuals or places that receive treatment are chosen because they are the easiest to treat or most likely to show successful results. An experiment ideally reduces this bias, because individuals or places are randomly assigned to a treatment or control group. Selection bias can still occur in an experiment, however, particularly if officers are involved in the randomization process. In the Minneapolis domestic violence experiment, for example, misdemeanor domestic violence offenders randomly received one of three treatments (arrest, separation, or counseling). For eligible cases, officers were supposed to open an envelope that specified the assigned treatment for that case. Selection bias could occur here if, for example, officers ignored the assigned treatment and made an arrest because they believed it was the “right” treatment for an offender. Ideally, researchers should oversee the randomization process when possible. Researchers at the University of Cambridge have developed the “Cambridge Randomizer” (Ariel et al., 2012) which is useful for person-based experiments and allows researcher oversight of the randomization process through an online portal.¹⁵

Step 9: Ensuring successful data collection and analysis

Ensuring that data is successfully collected and analyzed is essential for accurately reporting the results of an experimental study. The data used and the appropriate analysis strategy will vary

¹⁵ Learn more about the Cambridge Randomizer at: <http://link.springer.com/article/10.1007/s11292-012-9141-4/fulltext.html> and <http://www.crim.cam.ac.uk/events/documents/EBPCIVtcr.pptx>.

based on the study, but we provide some general tips here on data management for an experiment.

Observe how the data are being collected, inputted, and processed. Make sure research personnel review each of the data. Organize the data into easily understood formats, and label the data meticulously so there is no confusion as to where the data came from and what the data represents. As a study progresses, the amount of data exponentially increases, and this can potentially become overwhelming. Labeling and filing will help to organize the data. Be sure to back up all data files so if one becomes corrupt there is not the chance of losing the data forever.

As noted above, conducting a pre-test before the experiment begins is useful to ensure data systems are operating properly. It is also important to examine all data within the first few days of the experiment to verify data are accurate, accessible, and clean. This will help verify whether personnel are complying with the research protocol. It is imperative to discover early in the research process if personnel do not understand the parameters they are supposed to be following, or if they are purposefully not following research protocols. Data also might not be accurately captured. For example, systems may not collect data in the intended manner and examining data early in the process can help identify any problems. Information technology staff are an incredibly useful resource for data collection and analysis. These staff can ensure forms, databases, and analytic interfaces are available and working.

Allow time for personnel to gather, search, sort, and analyze the data and create GIS maps, charts, tables, etc. This will take an extensive amount of time. You should allocate extra time for innovative approaches to the data. The study may generate new questions that the data may be able to answer. Make sure the data collected answers the hypothesis. Remember that analyzing the number of arrests made does not give an indication of crime going down. It is just an indication of work productivity. When analyzing data, choose personnel who have a strong background in statistical analysis or ask a local university for a graduate student for assistance. Crime analysts will be particularly useful for data analysis. The crime analysts' expertise in GIS and data systems will be incredibly valuable for analyzing the data. Analysts can determine what type of data can be mined from the software systems and how efficiently these data can be accessed.

Be sure to double check that the data are accurate. Having more than one set of eyes on every part of the study is very useful. An outsider may be able to see something you missed.

Step 10: Completing the experiment

Once you have completed the study and data analysis, share its results with internal and external groups. These results can be shared in many ways including in-house reports and presentations, media releases, academic articles and presentations, and practitioner-oriented articles and presentations.

A brief written report and/or PowerPoint presentation should be provided to departmental leadership, so they have a clear understanding of the results of the experiment and can act accordingly to adjust agency policies and practices as necessary. If results suggest an intervention implemented in select districts reduced crime, for example, then the agency ideally

will use that information to expand the intervention citywide. The results then provide a clear justification to officers for any shifts in departmental strategies.

All agency personnel should also receive information on study results. Results of the study should be made easily available, whether they are posted on an internal or public website, referenced in the agency's newsletter, or described in Compstat or at roll call presentations. Personnel should be allowed to question results to enhance their understanding of a scientific approach to policing. All the results of the study should be released—not just an outcome that makes the intervention look particularly favorable. Even if a study shows that an intervention did not reduce crime, for example, this is still important information for the department to consider in decisions about whether to continue using that intervention in the future. The use of experimental designs provides the best answers to the question of whether a policy or practice is effective, and so the need for using good science to evaluate departmental strategies should be continually emphasized, even if the results of an experiment are unexpected or disappointing.

The results should also be presented to external groups. If the study suggests beneficial results (e.g. a reduction in crime from a new intervention), a press release can help inform the public about the study and can lead to positive print and broadcast media attention for the agency. In the SPD case, Sgt. Mitchell conducted several interviews with Sacramento newspapers and television stations following a press release touting the successful results of the hot spots study.

A series of publications and presentation followed the press release. The results of the SPD experiment were published in a scholarly journal in collaboration with researchers from the CEBCP (*Justice Quarterly*, Telep et al., 2014), a practitioner-oriented magazine (*The Police Chief*, Mitchell, 2013) and online through the e-newsletter of the Office of Community Oriented Policing Services. Mitchell also gave presentations about the experiment at a number of academic and practitioner-oriented events including the Jerry Lee Crime Prevention Symposium, the Academy of Criminal Justice Sciences Annual Meeting, and the CEBCP Evidence-Based Policing Workshop.

Publishing the study allows other agencies to benefit from the information generated from the study. Information flow between agencies in the United States is sluggish; publishing allows for a faster flow of information. Ideally, the study should be published in both academic and professional publications to cater to both audiences.

It is also important to follow-up with other agencies interested in the study. Once presentation, published articles, or media releases are available, other agencies will likely request more information on the study. Depending on the innovation or success of the study, the number of agencies asking for information can be overwhelming. The Los Angeles Police Department and Santa Cruz Police Department had hundreds of agencies call about the predictive analytics software they were using. This amount of intense attention can tie up scarce resources.

To prepare for this possibility, create a team to answer questions about the study. Make sure the team members who will be answering questions understand all aspects of the study. Law enforcement agencies will ask how the idea came to mind, how it was designed, how it was implemented, and what the results were. Other agencies will want to know what technologies were used and exactly how they were used. They will want to know the officer's reaction to the study and best practices for achievement officer buy-in. Decide beforehand who will oversee media interviews, give presentations, or return phone calls. In the end, some 30-35 agencies

called the SPD asking for assistance in implementing a hot spots strategy within their own agencies. SPD personnel spent time looking over agency proposals and giving suggestions on implementation and strategy. Some agencies requested site visits from SPD personnel. Individual officers, sergeants, and lieutenants even called the SPD looking for information on the hot spots study to use for promotional interviews.

Although planning, designing, and implementing research might seem like a daunting task, the return on the investment is worth it. Both SPD and PPB command staff and crime analysis were invigorated by the idea of challenging conventional thinking in policing, and then creating a rigorous evaluation to test that new thinking. Challenging an organization to reach beyond its abilities stimulates creativity, tenacity, critical thinking, and a renewed sense of purpose—excellent traits for the evidence-based law enforcement organization.

6 Challenges and Frequently Asked Questions

Below are responses to frequently asked questions about issues that may arise in designing and implementing an experiment.

Why should my department engage in research?

Sir Dennis O’Conner, formerly Her Majesty’s Chief Inspector of Constabulary in the United Kingdom, once said that ignoring police research or evidence constituted professional negligence. In other words, research helps the police profession truly become professional, and an evidence-based approach builds mechanisms and infrastructure for self-evaluation. Using the scientific method to understand whether police strategies, tactics, and practices are effective allows the police profession to move away from being a trade embedded in a tradition of apprenticeship toward a professional model of crime prevention and internal managerial strategies that have been demonstrated as effective. Imagine if every single police agency in the world implemented just one study; we would have thousands of studies in policing to build a body of knowledge that all police professionals could benefit from. Currently, there are many unknowns in policing, and if these unknowns were better researched, we could potentially generate more efficient or effective policing approaches.

My personnel don’t want to participate in the study. How do I overcome this?

There are many ways to overcome resistance. At the most basic level, officers need to be trained, not only in carrying out the experiment, but in principles of evidence-based policing more generally, and in research findings more specifically. Institutionalizing a culture of research and evaluation requires training, just as officers are trained in how to use their firearms, how to arrest individuals, or how to submit evidence.

Often in police organizations, resistance comes in the form of reduced morale or concerns raised by police unions. If this is the case, an agency may want to rethink the timeline for implementation. If the resistance is strong, then more time may be needed up front to find the opinion leaders who are resisting the study and pull them into the project group. Finding out why the officers do not want to participate is important, and those opinion leaders can assist in sorting out that issue out. Of course, continuity and strong leadership applied to strategic approaches like evidence-based policing are essential ingredients.

In the case of the SPD, two studies were run in succession. The hot spots study was run in the first part of 2011, and a prolific offender study was run in 2013. The studies demanded different resource allocations from the officers. In the hot spots study, officers had to commit to getting to the hot spots multiple times in one day. In the prolific offender study, specialized teams made contact with prolific offenders over a period of time. In both studies, officers did what they were asked to do, but the hot spots study required more time from them. Not surprisingly, there was more pushback from the officers in the hot spots study because it was an additional job duty in an already demanding shift. However, having been exposed to the hot spots study first made the second one less overwhelming to the officers as they were becoming comfortable working within the parameters of field research.

This demonstrates that when employees do not want to participate in a study, it may be due to their fear of doing something outside their comfort zone. Once an organization adapts to the experimentation process, employees may feel less resistant to participate.

There are complaints from the community about the study. How do I handle this?

In the Sacramento hot spots study, community members—mostly employees of local businesses—complained about the officers parking in front of their businesses causing customers to perceive there was an increased crime problem in the area. SPD had a sergeant reach out to the businesses that were making the complaints to explain the study and why the department felt research was important for the business community, and to ask if there was anything we could adjust that would benefit the company without jeopardizing the experiment.

If it would not damage the validity of the study, outreach before the study begins would be the best way to ward off potential complaints. If this were not possible, then having employees available to discuss the study with individual complainants would be useful. Just as in the private sector, customer service goes a long way.

The results suggest one of our tactics is not effective. Now what?

It depends. Is this a tactic that other studies have found to be effective, or was this study testing a tactic for the very first time? Either way, the project lead should research all the previous literature about that type of study or studies from similar tactics. Using research to drive police strategy is a new step in the Scanning, Analysis, Response, Assessment (SARA) model.¹⁶ Thus, an organization would use this information in the assessment portion of SARA to determine whether to keep using the tactic or form a new strategy.

By using SARA, you may decide to look at the strategy that was evaluated to determine if there is a way to tweak the strategy to make it effective. For example, the Sacramento study tested Koper's suggestion to use 15-minute police patrols in hot spots. If a study had determined that this tactic was not effective, then a follow-up study with 10-minute, 20-minute, or 30-minute patrols could be done to identify a better alternative.

Police agencies might view a study result showing ineffectiveness as a failure, but remember that an experiment just gives you information about your organization. It is up to you how you interpret that information and use it in the best interest of your organization.

We discovered that personnel were not following the research protocol. Where do I go from here?

Determine what part of the protocol they were not following and why. This would be a good time to get opinion leaders involved in a solution. If they were involved from the beginning of the research project, they may be able to tell you what went wrong in research implementation. For example, did officers have difficulty understanding the protocol? Or did officers have difficulty with time management during the study? Once the underlying reason has been uncovered, then the project team should create a new approach. The team must decide what approach will best address the issue. This could be, for example, undertaking new training, having supervisors apply closer scrutiny of officers, or developing a creative reward system. If implementation is slow or nonexistent at the outset, it may be necessary to extend the experimental period for an additional month or two to ensure a sufficiently long treatment period. Whatever the approach, all data should be collected, and any implementation issues should be well-documented. This way when the study resumes, the interruption can be documented in the analysis. In the original hot spots study in Minneapolis, Sherman and

¹⁶ Learn more about the SARA model at: <http://www.popcenter.org/about/?p=sara>.

Weisburd (1995) had difficulty maintaining officer dosage during the summer months, and then other issues with a new computer system that was installed during the study created data problems. Sherman and Weisburd (1995) addressed both issues in their article, demonstrating that even though difficulties arise during implementation, studies can still be salvaged, but proper documentation is imperative.

My old supervisor supported this study; my new supervisor does not. What can I do?

Every police supervisor has different concerns. If the new supervisor does not support the study, find out what his/her concerns are and engage in open communication with your new supervisor. Is the supervisor concerned about the budget, the amount of resources being used, the time it takes to document the study, or the public's view of the study? Alternatively, perhaps the supervisor or commander wants to make his or her own mark in the unit and not engage in previous activities. Individuals who find themselves in this position might find out from the supervisor's old team what was important to him or her. What projects he did like to work on? What projects did she vocalize support for? What type of supervisor was he? Does she like to be innovative? Does he like to keep things status quo?

Once you have a general idea of why your supervisor opposes the study, you will be in a better position to respond. If the concern is the budget, make a spreadsheet demonstrating where the money is coming from and how the study will benefit the department financially. If the concern is about the public, demonstrate how the unit is using social media or local media outlets to advertise the positive aspects of the study. When the SPD began its hot spots study, every media outlet from the newspapers to radio stations covered the launch with a positive slant. Whatever the supervisor's issues are, finding ways to address those concerns proactively may help.

If the study is already underway, it should already have the support of your chief executive or sheriff, which will also increase the likelihood that your supervisor will be supportive. Find a way to professionally alert your supervisor about your chief's support. Take the time to explain the study and its process to your new supervisor and provide him or her with as much documentation as possible. Even if your supervisor is not completely supportive, obtaining at least his or her cooperation is important to any study. And don't give up. Not every supervisor will support every study or project. There is evidence that shows within agencies a minority of officers and supervisors might be opposed to research. However, many supervisors do support research and innovation as well.

Are there examples of successful experiments conducted in police agencies that I can read about?

Yes, there are a number of experimental evaluations in policing. Many of these focus on police efforts to reduce crime, but some experiments have also covered non-crime related issues such as use of force and the quality of interactions with citizens.

Examples of crime control experiments:

The Evidence-Based Policing Matrix (<http://policingmatrix.org>) includes all randomized experiments on crime control topics. Some of the most famous crime control experiments in policing include:

- Minneapolis domestic violence arrest experiment (Sherman & Berk, 1984) –

Matrix page: <http://cebcp.org/evidence-based-policing/the-matrix/individuals/individuals-sherman-and-berk-1984/>.

Police Foundation report: <https://www.policefoundation.org/publication/the-minneapolis-domestic-violence-experiment/>.

- Minneapolis hot spots experiment (Sherman & Weisburd, 1995) –

Matrix page: <http://cebcp.org/evidence-based-policing/the-matrix/micro-places/micro-places-sherman-and-weisburd-1995/>.

One-page summary: <http://cebcp.org/wp-content/onepagers/ShermanWeisburd1995.pdf>.

- Jersey City problem-oriented policing in violent crime places experiment (Braga et al., 1999) –

Matrix page: <http://cebcp.org/evidence-based-policing/the-matrix/micro-places/micro-places-braga-et-al-1999/>.

One-page summary: http://cebcp.org/wp-content/onepagers/POPInViolentCrimePlacesRCT_BragaEtAl.pdf.

- Philadelphia foot patrol experiment (Ratcliffe et al., 2011) –

Matrix page: <http://cebcp.org/evidence-based-policing/the-matrix/micro-places/micro-places-ratcliffe-et-al-2011/>.

One-page summary: http://cebcp.org/wp-content/onepagers/PhiladelphiaFootPatrol_Ratcliffe.pdf.

Examples of departmental policy experiments:

- Shift length experiment comparing 8-, 10-, and 12-hour shifts (Amendola et al., 2011) –

Police Foundation website on the experiment: <https://www.policefoundation.org/projects/the-shift-length-experiment/>.

Police Foundation practitioner guide: <https://www.policefoundation.org/publication/the-impact-of-shift-length-in-policing-on-performance-health-quality-of-life-sleep-fatigue-and-extra-duty-employment-full-report/>.

- Chicago Quality Interaction Program experiment on training for procedural justice (Rosenbaum & Lawrence, 2012) –

National Institute of Justice report: <http://www.nationalpolicersearch.org/storage/research-papers/Rosenbaum%20%20LawrenceTeaching%20Respectful%20Police-Cit%20%20Encounters.pdf>.

- Experiment on the impact of TASERS on police use of force decisions in field training (Sousa et al., 2010) –

Abstract available at: <http://link.springer.com/article/10.1007/s11292-010-9089-1>.

- Impact of officer-worn cameras on officer behavior, use of force, and complaints (experiments completed in Rialto, CA and Mesa, AZ) –

Summary of the Rialto experiment: <https://www.policefoundation.org/publication/self-awareness-to-being-watched-and-socially-desirable-behavior-a-field-experiment-on-the-effect-of-body-worn-cameras-on-police-use-of-force/>.