Foreword

Lawrence W. Sherman

Published online: 21 August 2010
© Springer Science+Business Media B.V. 2010

In the late 1960s, City University of New York psychologist Morton Bard invented a new strategy for policing domestic violence. Applying the “crisis intervention framework of clinical psychology, he convinced New York City Police Commissioner Howard Leary to allow him to train a team of some 10 police officers from the “Spanish Harlem” area around upper Broadway in Manhattan. The training consisted of methods for calming people down, especially the technique of two officers each taking one party into a separate room to hear their side of the story. The goal of the program was to achieve on-the-spot reconciliation, a peaceful basis for leaving a family intact, while returning police to patrol duties as soon as the goal could be achieved in each case (Bard 1970).

By 1980, Bard’s approach had fallen out of fashion. One reason was that punishment, not peacemaking, had become the more politically prominent objective for police to achieve. A less obvious reason was that Bard’s experiment had never been properly evaluated. His invention was a case of what a famous surgeon once enjoined a young medical inventor: “Don’t think, try” (Sherman 2010b). Bard did not take the next step required by the motto of the Cambridge University Police Executive Programme: “Don’t just try. Test.” The work of Bard’s Family Crisis Intervention Unit (FCIU) was not subjected to any randomized controlled trial or even a matched comparison. Bard did not even track future violence or calls for service for FCIU cases. While he implied that FCIU officers were less likely to be assaulted in domestic calls than officers without FCIU training, he published no direct evidence on that hypothesis. Thus, when the political fashion switched from peacemaking to punishment, there was little that could be said about the possible merits of not arresting.

L. W. Sherman (✉)
Cambridge University and University of Maryland, Cambridge, MD, USA
e-mail: ls434@cam.ac.uk

This volume shows how far the field of experimental criminology has come since then, at least in testing responses to domestic violence. As the Davis and Auchter article reports, there has been a steady growth in the use of RCT designs to test responses to domestic violence since the first randomized trial of policing domestic violence (Sherman and Berk 1984). These responses are, for the most part, new inventions that were not even conceived when Bard pioneered the entire idea of innovation in domestic violence response. Even arrest for misdemeanors, which Bard discouraged, was legally impossible in the 1960s unless police had witnessed an offense. From the standpoint of the injunction to “First, do no harm,” the good news is that new inventions for domestic violence are now much more likely to be tested rigorously than they were at the dawn of this inventive era.

The fact that many of these tests show negative results is also good news. Unlike Stalin’s Russia, where statisticians were executed for reporting unwelcome news—or even the UK government under a recent Cabinet member who try to suppress unwelcome research results—the United States has been able to handle the truth, at least at its National Institute of Justice. Davis and Auchter’s article about NIJ reflects some uncertainty about just how good it is to prove that point with bad news, understandably so for a government agency created to find better solutions to crime problems. But the lesson of the Minneapolis Domestic Violence Experiment (Sherman and Berk 1984) and its replications is that one size is not likely to fit all cases, regardless of the public appetite for simple solutions.

A yearning to feed that appetite is unfortunately evident in the way many people read the evidence from those six experiments, which constitute about one-third of the NIJ-funded RCTs on domestic violence. The evidence from three of those experiments (Milwaukee, Omaha and Miami) clearly shows that arrest causes *more* violence against women among unemployed suspects, measured by the only indicator that covered 100% of the suspects—arrest or other official records. Victim survey data in all six experiments suffered woefully low response rates, making that evidence largely uninterpretable due to potential selection bias in the responses. Yet many advocates of arrest for misdemeanor domestic violence, for example, cite only the Minneapolis experiment, and assert that arrest “works.” NIJ’s failure to publish its own report on this multi-million dollar research program has helped to cloud this bad news, if not to bury it. Perhaps it is time for a new NIJ Director to help consolidate what NIJ has learned, overall, from its experiments on domestic violence.

This volume would be an excellent place to start. Perhaps its most important contribution is the Davis, Weisburd and Hamilton paper on the Redlands experiment, a definitive blow to the lingering hope that second responder programs may be effective. It is a welcome antidote to the sugar-coated analysis in the current Campbell Collaboration review on the same subject (Davis et al. 2008), which leaves practitioners with the impression that an *increase* in police calls about domestic violence is an indicator of success! The much clearer (and more negative than the Campbell) interpretation of the Redlands experiment published in this volume should be added as soon as possible to an updated Campbell review that is much more sceptical about the value, and more explicit about the danger, of second-responder programs.

The Maxwell, Davis and Taylor article in this volume makes a major contribution to deterrence theory, with its evidence that ongoing participation in monitoring is associated with less violence. That evidence is unfortunately not definitive, since the key evidence is correlational within treatment groups rather than experimental across treatment groups. But it suggests a new way to develop further refinements to court supervision of batterers, as part of the ongoing process of developing and testing (Sherman and Strang 2004; Sherman 2006).

The Taylor, Stein and Burden paper on their randomly assigned Cleveland classrooms is a fine example of controlled experiments with aggregate populations, which is especially important for universal prevention strategies. Their results are both encouraging and challenging: encouraging because the interaction-based curriculum reduced teen violence relative to control classes, challenging because of some evidence that the program *increased* dating violence. Their paper shows further support for the idea of experiments as inventions, rather than verdicts (Sherman and Strang 2004).

The Labriola, Rempel and Cissner paper in this volume is a welcome contribution to what should be a growing literature on the implementation of RCTs. While many of the lessons learned have been published before, they have not deterred the uninitiated from diving into RCTs without any technical preparation for implementation. Some of the authors whose problems are documented by Davis and Auchter, for example, ignored counsel from experienced experimenters about implementation, only to report failures of implementation. Some of that literature has recently been assembled in planning guidance explicitly aimed at experimental criminology (Sherman 2010a), but an ongoing distillation of those lessons is required for such a complex field. Clearly the literature on experiments needs more reading, as well as publishing, by anyone who would undertake experiments in domestic violence.

Further evidence for better quality is marshalled in this volume by Davis and Auchter, who catalog the many methodological issues presented by poorly-executed experiments. That list of issues adds force to their argument that CONSORT standards of reporting experiments (www.consort-statement.org) should be required by NIJ, and indeed by this Journal. Until we can raise the technical standards that grantors and journals require of experimental criminologists, the field will waste money on studies failing to reach the high level of quality that the articles in this volume generally represent. That is clearly not good news, either for research or for the victims of domestic violence.

References

- Bard, M. (1970). *Training police as specialists in family crisis intervention*. Washington, D.C.: U.S. Department of Justice.
- Davis, R., Weisburd, D., & Taylor, B. (2008). Effects of second responder programs on repeat incidents of family abuse. *Campbell Systematic Reviews* 15 doi:[10.4073/csr.2008.15](https://doi.org/10.4073/csr.2008.15)
- Sherman, L. W. (2006). To develop and test: The inventive difference between evaluation and experimentation. *Journal of Experimental Criminology*, 2(3), 393–406.
- Sherman, L. W. (2010a). An introduction to experimental criminology. In A. Piquero & D. Weisburd (Eds.), *Handbook of quantitative criminology* (pp. 399–436). New York: Springer.

- Sherman, L. W. (2010b). Criminology as invention. In M. Bosworth and C. Hoyle (Eds.), *What is criminology?* Oxford: Oxford University Press. (in press)
- Sherman, L. W., & Berk, R. A. (1984). The specific deterrent effects of arrest for domestic assault. *American Sociological Review*, 49(2), 261–271.
- Sherman, L. W., & Strang, H. (2004). Verdicts or inventions? Interpreting randomized controlled trials in criminology. *The American Behavioral Scientist*, 47(5), 575–607.

Lawrence W. Sherman is the Wolfson Professor of Criminology and Director of the Jerry Lee Centre of Experimental Criminology at the Institute of Criminology, Cambridge University. He is also Distinguished University Professor at the University of Maryland, College Park, Department of Criminology and Criminal Justice. The founding President of the Academy of Experimental Criminology and current Chair of the ASC Division of Experimental Criminology, he has designed or directed over 30 randomized controlled field experiments, including the Minneapolis and Milwaukee domestic violence arrest experiments.