

The Criminologist

The Official Newsletter of the American Society of Criminology

Vol. 34, #1

January/February 2009

Editors Note: ASC President Todd Clear has initiated a discussion and reconsideration of ASC's role in advocacy and policy. To help us all think through this important subject, I have solicited a set of lead articles for *The Criminologist* that express a variety of views on relevant issues. Many thanks to Akiva Liberman for providing the first of the series.

Wayne Osgood, ASC Vice President

ADVOCATING EVIDENCE-GENERATING POLICIES: A ROLE FOR THE ASC

Akiva M. Liberman, National Institute on Drug Abuse

Note: Views expressed here are solely those of the author, and do not represent official positions or policies of the National Institute on Drug Abuse.

Criminologists often lament their lack of influence on policy, and ASC members continually debate how the Society can exert greater influence. This discussion generally presumes that the route to policy impact is the collective advocacy and/or endorsement of particular policies. If so, the key impediment would seem criminologists' inability to agree on which policies to support. The discussion about policy impact thus often reduces to debate about which particular policies to support, based on what types and standards of evidence, and on how to achieve consensus. Such discussions are at the heart of criminology, but are rarely resolved. To date, ASC has taken only two policy positions, the first in 1989 against the death penalty, and the second in 2007 opposing the use of UCR data to rank the safety of American cities.

I believe that an alternative conception of appropriate advocacy would lead to a more constructive approach to policy. I refer to this approach as advocating *evidence-generating policies*, by which I mean advocating that policy-makers treat policy changes as experimental and expected outcomes as hypotheses, and implement those policies in such a way as to generate reasonable evidence about their effects. This approach was broadly laid out by Donald Campbell (1969, 1988/1971) in *Reforms as Experiments* and *The Experimenting Society*. It is not at all trivial to enact, but has the potential to generate much greater consensus within the Society and enable the Society to advocate for the importance of research in improving policy. In contrast to advocating *evidence-based* policies, advocating *evidence-generating* policies would draw on criminologists' methodological and substantive expertise to advocate that policy changes be used to generate evidence, while often frankly acknowledging the lack of an unambiguous evidence base.

The key here is to stress the distinction between the decision to try a policy from an *a priori* commitment to its effects. As Campbell (1988/1971, p. 293) describes in *The Experimenting Society*: "Faced with a choice between innovating a new program or commissioning a thorough study of the problem as a prelude to action, the bias would be toward innovating." The basic difference between an experimenting versus non-experimenting society is not whether policy is made in advance of unambiguous evidence; the basic difference is in whether policy changes themselves are used to generate evidence.

(Continued on page 3)

IN THIS ISSUE...

Teaching Tips.....	9-11	2008 ASC Annual Meeting Pictures.....	33-36
2009 Election Slate.....	11	2009 ASC Award Nominations.....	37-41
Featured Articles.....	13-23	Position Announcements.....	45-49
2009 ASC Annual Meeting CFP.....	25-32		

(Continued from page 1)

The Problem of Premature Evidence-Based Policy

While evidence-based policy is a popular goal among researchers and policy-makers, the evidence is insufficient to inform many policy questions. What role can our professional societies play when definitive evidence is not available, and "evidence-based policy" is premature? We seem left on the horns of a dilemma: Absent definitive evidence, we might either advocate that policy-makers wait for a definitive study, or we might provide policy-makers with our best judgment, however speculative. The former is naive; the latter seems risky for scientific organizations, although potentially quite reasonable for individual criminologists.

If clear and unambiguous evidence is a prerequisite for establishing scientific consensus leading to collective policy recommendations, then strengthening the evidence base would obviously help, especially if we can set *a priori* evidentiary standards for recommendations. Researchers pursuing this strategy have mobilized to advocate stronger methods both in primary research and evaluation (e.g., randomized control trials) and in research synthesis (e.g., meta-analysis), and to clearly articulate standards of evidence. The promising results of this strategy are evidenced in numerous reviews including those of the University of Colorado's *Blueprints for Violence Prevention*, metaanalyses conducted by the Campbell Collaboration's Crime and Justice Group, Mark Lipsey and colleagues, and Steve Aos and colleagues, as well as government efforts such as reviews of violence prevention interventions by the CDC Task Force on Community Preventive Services, and SAMHSA's National Registry of Evidence-Based Programs and Practices. We seem to have an accumulating body of knowledge about effective programs and practices. While skeptics remain, we seem to have made considerable progress since Martinson's (1974) famous conclusion that "nothing works."

Unfortunately, however, the promise and progress in identifying evidence-based *programs* is considerably greater than for identifying evidence-based *policies*. For programs, the promise of the evidence-based movement is to identify programs that can be taken "off the shelf" and implemented as originally designed in new settings or contexts to reliably produce desired effects.

Although growing, the set of effective programs is still small relative to need. Even for identified programs considerable challenges remain. What are the necessary conditions for implementing a given effective program? How can effective local programs be brought to national scale, while maintaining their integrity and fidelity? Can effective programs be adapted to local conditions while retaining sufficient program fidelity to maintain their effects? A growing literature finds implementation fidelity to be key for achieving promised program effects. Yet local adaptability often seems key to program adoption, creating a difficult tension much discussed by program evaluators. What, then, are the critical elements of effective programs that must be maintained, while allowing adaptation of other elements? The difficult work to unpack the "black box" of an effective program to identify its critical elements must generally rely on weaker methods than those used to identify the effective program itself, because strong experimental designs are used less often to experiment with program elements than to test programs as a whole. An alternative approach to resolving this tension is to identify *principles* of effective programs, but I will not belabor here its considerable theoretical and practical difficulties, let alone whether these principles can generate the consensus of evidence-based programs.

Even with all of these challenges, the search for evidence-based *programs* is considerably farther along than the search for evidence-based *policies* in criminal justice. The evidence-base is generally weaker for criminal justice policies than for programs, and even less often based on strong research designs. The circumstances in which new policy is made also often seem to differ considerably from the circumstances under which prior evidence was generated. Furthermore, the notion of fidelity that seems key to the effective implementation of evidence-based programs is not well worked through in many realms of policy. Finally, criminal justice policy makers are even less likely to simply adopt prior policies without their own adaptations than are administrators searching for effective programs. (We may sometimes have clearer evidence of what policies do *not* work, but policy-makers seem unwilling to merely follow recommendations of what not to do.)

Opportunities for Evidence-Generating Policy

Framed differently, the absence of sufficient evidence regarding a proposed policy offers a research opportunity. When a policy is changed, we have the possibility of a natural experiment. But will the policy be implemented in such a way that anything can be learned? Usually not.

(Continued on page 4)

(Continued from page 3)

As social scientists, we can collectively advocate that policy changes and their implementation be used to *generate* much better evidence. Many of these natural policy experiments will be of limited methodological quality, but they can certainly be much improved. A particular policy can often be implemented in ways that make it more conducive to comparisons with alternatives, using techniques such as staggered policy implementation over time and place, sunset provisions, and the gathering of pre-implementation data. The goal of evidence-generating policy is to improve the "counter-factual" comparisons involved in thinking about policies' effects. That is, we often know what happens after a policy change, but that is insufficient information. To discuss the effect of a policy, we must make assumptions about what would have happened without the policy change.

Methodological rigor serves to reduce the range of alternative explanations for observed outcomes, and methodological improvements lead to stronger inferences by eliminating alternative explanations and inferential threats. (This is essentially the approach of Campbell & Stanley, 1963, in their classic work on quasi-experiments; they use the term "sources of invalidity" for what I refer to here as alternative explanations and inferential threats.) In these terms, double-blind randomized trials are designed to eliminate all alternative explanations but chance, but are essentially unavailable for criminal justice policy. Non-blinded randomized trials eliminate most alternative explanations, and are sometimes available in criminal justice.

Weaker methods, too, can eliminate important alternative explanations and inferential threats. As a simple example, absent random assignment, pre-post designs with comparison groups -- even non-equivalent comparison groups -- are much stronger than either post-only designs with comparison groups (threatened by unmeasured pre-experimental differences) or than pre-post designs without a comparison group (threatened by history effects). This is understood by most researchers, but by few policy-makers.

For many new policies, a modest investment to identify reasonable comparison groups in advance, and to check that comparable pre-measures are available for both groups, would greatly strengthen later inferences about the policy's effects. In some cases, a slight delay in program implementation to allow collection of some pre-measures might have a large inferential payoff. Identifying comparison groups, populations, or places in advance can also eliminate the threat or suspicion of evaluators cherry-picking amongst plausible comparisons *post-hoc*, in order to support their preferred conclusions.

The goal advocated here is modest; it is to push for policy implementation that generates better evidence. Criminologists will continue to evaluate the effects of implemented policies regardless, devising *ad hoc* inferential strategies to make use of available data. Many of those studies will be published, some in ASC's own *Criminology and Public Policy*, many will be disputed there as well. Modest steps to proactively eliminate easily identifiable inferential threats in evaluating criminal justice policies could considerably improve the strength of that research. If we can incrementally improve the evidentiary possibilities that result from policy implementation, we may also incrementally increase the receptiveness of policy audiences to stronger methods.

Planning for Evidence-Generating Policy

Policy domains vary in at least four key respects that determine which evidence-generating strategies are useful. Advance work to think these through would enhance our ability to advocate evidence-generating policy at the moments when such advocacy is needed.

First, policies differ in their key outcomes. Trivial as this may seem to academics, policy-makers rarely articulate the key outcomes they anticipate from their policies. Program evaluators often find that a key early step is working with a program's organization and staff to articulate the program's key outcomes and the logic by which the program is hypothesized to influence those outcomes. The same holds for policy. Criminologists can help to articulate relevant outcomes in advance. For example, are the intended outcomes of prison policies only in-prison outcomes, or also successful reentry and reduced reoffending?

Second, different policy domains are threatened by different key inferential threats. For example, a key inferential threat for evaluating local burglary prevention policies is unmeasured geographic displacement, in which local crime appears to have declined but has merely moved to another locale. Displacement can also threaten evaluations of policies to disrupt local drug markets. In contrast, for evaluations of drug courts, a key threat is net-widening, where enthusiasm for a program leads to expanding its reach to include less serious offenders than previously, which may generate spurious declines in reoffending.

(Continued on page 5)

(Continued from page 4)

Third, policies differ in feasible methodological approaches to addressing key inferential threats. Some policy changes can more easily be used to generate policy-relevant evidence than others. For example, police practices can sometimes be experimented with in some precincts but not others, in experimental designs. In contrast, policies that involve legislative change seem much less amenable to such experimentation. One would be hard-pressed to implement state sentencing policies in a staggered way or in only some of the effected jurisdictions. Nonetheless, even in such cases, consideration of counter-factual comparisons before policy implementation can enable much better inferences later, by identifying reasonable comparison jurisdictions, non-affected but similar offenders to those affected, or similar strategies.

Fourth, policy domains differ in the administrative data available to assess relevant outcomes, both before and after a policy change. For cross-jurisdictional comparisons, an important question may be whether comparable data exists in comparison jurisdictions. How would the various administrative data systems at federal, state, or local levels need to be supplemented in order to be useful in evaluating policy changes in different domains? When reasonable data do not exist for a policy domain where we anticipate policy changes, criminologists can advocate that the existing data systems be supplemented or that relevant data systems be created. As a case in point, when the U. S. Sentencing Guidelines were established, preexisting administrative data did not allow strong projection of the prison populations that would be expected, counterfactually, without the sentencing guidelines. To allow later estimates of the effects on the federal prison population, the U. S. Sentencing Commission persuaded the U. S. Pretrial and Probation Service to collect supplemental data before the guidelines were implemented and to establish systems for collecting later data (Gaes, Simon, & Rhodes, 1993).

Collectively, we should engage in advance work to examine how different policy domains relate to these four issues. In each policy domain, what are the relevant key outcomes, the key inferential threats, and the possible inferential yield of various methodological approaches to address those threats? What administrative data are available, what are their strengths and weaknesses, and how would that data need to be supplemented before policy changes to improve later inferences? ASC members have the substantive and methodological expertise necessary to elucidate these key aspects of implementing policy in an evidence-generating way. We are often reminded that when policy changes are being considered, there is usually little time for extensive deliberations by academics. Advance work as suggested here, perhaps resulting in white papers, would position us to more effectively advocate using policy change as a vehicle for generating evidence, and to promote such efforts quickly in the brief moments of opportunity that arise.

Conclusion

Advocating particular policies has been problematic for ASC. The perennial debate over whether to advocate specific policies reflects the continual tension between our desire for more-informed criminal justice policy and our high standards for evidence. Given that tension, we may ultimately take very few policy positions. Many criminologists believe that ASC would risk its scientific and non-partisan credibility by advocating policies supported by anything less than unambiguous evidence and virtual unanimity. Many social scientists also feel that issue-advocacy pushes for categorical statements at odds with the social scientist's respect for the limitations of current knowledge.

In contrast, advocating evidence-generating policy draws on the skills and predilections of our research community. It involves an appreciation of both the benefits and limitations of research, along with a constructive agenda to incrementally improve the formulation, implementation, and ultimately the content of criminal justice policy. The key is to stress the distinction between the decision to try a policy versus an *a priori* commitment to its effects.

Of course, such an approach runs counter to currents of political decision-making which can push fast, relatively uninformed, and even faddish policy-making, often in response to crises. Yet that is precisely why ASC should advocate the role of research, and think hard about what would be involved in evidence-generating policy implementation. Promoting evidence-generating policy should be part of our general advocacy of informed policy, and part of our attempt to educate policymakers and practitioners about the usefulness and relevance of research.

References

- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Campbell, D. T. (1969). Reforms as Experiments. *American Psychologist*, 24, 409-29.
- Campbell, D. T. (1988/1971). The Experimenting Society. In S. Overman (Ed.), *Methodology and Epistemology for Social Science: Selected Papers* (pp. 290-314). Chicago: University of Chicago Press.
- Gaes, G. G., Simon, E. S., & Rhodes, W. M. (1993). 20/20 hindsight: Effectiveness of mimulating the impact of federal sentencing legislation on the future prison population. *Prison Journal*, 73, 5-29.
- Martinson, R. (1974). What works? Questions and answers about prison reform. *Public Interest*, 10, 22-54.