On being right, but unhappy

Elliott Currie
University of California, Irvine

An occasion like this one offers a rare opportunity to look back over what you have been trying to say for some time and to reflect on how well it has held up in a way that we do not often get to. I'm grateful, but this is also a perilous enterprise. The potential for a downside is obvious, particularly for those of us who have chosen, for whatever reason, to presume to voice opinions about big issues of social policy. You run the risk of discovering that you were basically wrong from the beginning and of being embarrassed—big time—by how much of your early thinking needs to be thrown out.

Another less anticipated danger is the discovery that you were basically *right* but wish you were not. I think that is where I am today as I look back through arguments I made about the roots of violent crime in America and about what would and would not work and to address it seriously. The good news is that I think my analysis was largely correct. The bad news is that I think my analysis was largely correct.

Let me go back 25 years, to the mid 1980s, when I was drafting the manuscript of the book that later became titled *Confronting Crime* (which was not the original title—I wanted to call it *The American Affliction*, which I thought was kind of sexy and also substantive, but my editor nixed that choice, for better or for worse).

This was in the first flush of the Reagan era. It is difficult to exaggerate the impact of conservative policy ideals at that time—not just regarding crime, but with respect to every aspect of social policy. It was said widely that more optimistic or traditionally "liberal" solutions to social ills, including crime, had been tried and had failed. Or worse, they had helped cause (or at least exacerbate) the problems. The absence of credible evidence for that belief did not faze increasing numbers of people across a surprisingly large swath of the political spectrum. So, in addition to the forceful and vocal case being made by newly resurgent conservatives, there was a pervasive sense of doubt, a failure of nerve even among people who thought of themselves as traditional liberals. The resulting vacuum created an opening for the resurfacing of ideas about crime and punishment (as well as much else) that, for some time, had been mostly confined to the fringes of public debate. It was a time that (at least when it came to our ideas about crime)

Direct correspondence to Elliott Currie, University of California, Irvine, 300 Social Ecology I, Irvine, CA 92697 (e-mail: ecurrie@uci.edu).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 1 12/21/09 7:13:51 PM

recalled the great line from Yeats's poem "The Second Coming": "the best lacked all conviction, the worst were full of passionate intensity."

The persistence of a very real state of crisis in many American cities contributed to the loss of faith in more progressive solutions to the problems of crime, violence, and drugs. Violent crime was bad, the inner-city drug trade was exploding, and entrenched problems of poverty and joblessness were barely budging, and in many places were worsening. As a result, some were throwing up their hands, and others were saying, in effect, that it was time to take the gloves off.

It seemed to me that both of these responses were wrong-headed and oddly timed because they appeared to ignore what we knew from an increasingly persuasive body of research and experience about the roots of the problems besetting us. It seemed timely to try to confront those ideas head on.

The resulting argument, as it began to form in *Confronting Crime* (Currie, 1985), took a substantial number of pages to lay out (278, to be exact), but its basic principles were simple. At its core was a reassertion of the crucial importance of systemic or structural issues in understanding why the United States had such an unusually severe level of violent crime, and in helping us to understand what it would take to change that tragic prominence. It pinpointed the extremes of inequality and the erosion of stable work, communities, and families as the key to the American affliction (along with an accompanying culture of predatory competition). It also suggested that these social ills were related—not separate and isolable problems, but part of a larger pattern of what I called "disruptive" social and economic development in which market forces had been allowed (or encouraged) to drive our social and economic policies to an unusual degree among advanced societies. And this pattern was not a matter of fate, human nature, or economic necessity; to a vital extent, it was the product of choices we had made and failed to make: "We have the level of criminal violence we do because we have arranged our social and economic life in certain ways rather than others" (Currie, 1985). Let me indulge in a quote from the concluding chapter (Currie, 1985: 225):

> The result is an unusual degree of erosion of the institutions that bear much of the responsibility for achieving socialization and social cohesion … the natural tendency of the marketplace is to put considerable strain on the institutions of family, work, and community … the growth of urban industrial consumer society has increased those strains throughout the world. In most developed nations, this process has been ameliorated to some extent by countervailing mechanisms of social obligation and support. In the United States, by contrast, this growth has been marked by the relatively unbuffered play of market forces and has been exceptionally disruptive of the institutions essential for the healthy development of character and community.

And some fundamental implications for policy flowed directly from this analysis (Currie, 1985: 225-226):

09012-CrimJournal-Guts.indd 2 12/21/09 7:13:51 PM

> If we are serious about attacking the roots of this American affliction, we must build a society that is less unequal, less depriving, less insecure, less disruptive of family and community ties, less corrosive of cooperative values. In short, we must begin to take on the enormous task of creating the conditions of community life in which individuals can live together in compassionate and cooperative ways.

The concern with the underlying structural forces responsible for American violence, however, emphatically did not mean that nothing short of a radical restructuring of the social order would make any difference. Contrary to the views of many people on both the right and the left at that time, I argued that many things could be done to enhance children's chances of growing up healthy and free of violence, to reintegrate offenders into productive lives, and to increase the safety of communities and the fairness and efficacy of the criminal justice system. All of these things were worth doing and deserved our attention and investment. But I was equally emphatic that no *solution* could be found for the problem on its American scale through piecemeal measures alone—inside or outside of the criminal justice system. Acting as if such a solution existed at best would bring us disappointment and at worst would create a society and a criminal justice system that we might not recognize and certainly should not want.

The argument was hardest on the conservative social policies and ideas that were in the ascendancy at the time—particularly the seductive idea that putting ever increasing numbers of people in prison would do the trick. But it also chastised liberal policy makers for too often failing to take the social context of American crime seriously enough. I pointed to what I thought was a curious disjunction between the liberal analysis of the roots of the crime problem and the remedies they were willing to put on the table. Liberals generally understood the importance of the structural forces I was pinpointing, but in practice, they often responded with narrow (and typically under-funded) attempts to change the behavior of individuals or to beef up the capacity of local communities to control themselves—in the face of structural economic forces that were ripping those communities apart and systematically eroding opportunities for productive and dignified lives. So just as there could be no prison solution to the crisis of American violence, neither would there be a neighborhood watch solution, or a group therapy solution, or a "defensible space" solution. That contradiction between analysis and action, I argued, had set liberal policy up for the charge of failure and boosted the case of those who argued that nothing short of massive increases in incarceration would work.

The implications of this analysis for the future were not spelled out thoroughly in *Confronting Crime* (Currie, 1985), but I did point to some fairly bleak possible scenarios. We could follow the logic of the conservative strategy to even greater extremes than we already had—and I worried that we might; although at the time, I did not believe it, because I thought the costs of doing so relative to any credible benefits would be obvious to nearly everybody.

I did a little thought experiment: Suppose we were to triple the existing prison population, which then stood at around 454,000 inmates. Based on a generous reading of the most credible incapacitation research, I thought we might achieve reductions of 20% to 25% for some

09012-CrimJournal-Guts.indd 3 12/21/09 7:13:51 PM

crimes. But that would happen only at great cost, if it happened at all—not just in dollars but in ultimately more important things as well (Currie, 1985: 90-91):

> Remember, too, that although these expenditures [on vastly increasing incarceration] would give us an incarceration rate more than 20 times that of some European countries, they would (even assuming that they did accomplish a 25% reduction in crimes) still leave us with far and away the highest rates of criminal violence in the developed world. Nor of course are these the only costs involved or even the most important … the United States already presents the demoralizing spectacle of one of the world's most highly advanced societies confining an unparalleled proportion of its citizens in order to achieve even the most minimal level of public safety. Much greater increases in incarceration would turn the American penal system, already swollen out of all proportion, into a homegrown gulag of dreadful proportions. This is bad enough in itself: but it would also profoundly disrupt community and family life…. The prison already looms as an ugly but pervasive presence in many American communities, especially among disadvantaged minorities…. Further drastic increases in incarceration would decimate these communities beyond recognition and would amount, in practice, to writing off a substantial part of entire generations of minority men.

Again, I did not really think we would go that far. What I *most* likely thought was that we mainly would follow a policy of muddling through. We would continue to skirt, for the most part, the measures we really would need to take if we finally wanted to address the economic and racial inequalities and the shrinking of opportunities that had given us the highest violent crime rate in the industrial world. We would probably make some incremental improvements to the justice system, and, at least sporadically, fund some genuinely helpful programs for children and families—although rarely on a sustained basis or on a scale that would match the magnitude of the need. In my heart of hearts, I thought that fiscal conservatism alone probably would forestall much greater increases in incarceration (now *that* was way wrong). But that muddling approach, of course, would not work either. The forces ripping apart the institutions of care, support, social control, and opportunity in the United States were entrenched deeply; in many ways, they were accelerating in the 1980s and seemed likely to continue to do so. If we did not make a concerted effort to take those forces on directly, then violence would remain *endemic* in many of our cities no matter how many prisons we built, how many neighborhood watch groups we started, how many community policing programs we established.

That is not to say that we would not have fluctuations in the crime rate—perhaps large ones. In fact, as I pointed out at the time, *Confronting Crime* was being written in the midst of a notable drop in crime from the scary peak we had reached at the beginning of the 1980s. But, barring a more structural approach to the toxic mix of ills that underlay it, violence would continue to be a source of fear, injury, and death for many Americans, especially the most vulnerable and disadvantaged Americans.

09012-CrimJournal-Guts.indd 4 12/21/09 7:13:51 PM

Okay, that was 25 years ago, roughly. After the intervening quarter of a century, what have we got?

We have got pretty much what I thought we would get if we continued to make the choices we seemed bent on making in the early 1980s. Oddly enough, policy makers turned out to be mysteriously immune to what I took to be compelling and careful arguments on my part (on the inside cover of *Confronting Crime*, a sympathetic but overly optimistic copywriter predicted that the book was "sure to change the way we see our number-one national problem—and the way we solve it." As my teenaged daughter used to say, "butohwell…"). So we did, in fact (with numerous exceptions), *mainly* do what I most feared in my least optimistic moments and then some. We tripled the prison population, in fact, we more than tripled the population, after having already doubled it by the time I wrote (in the early 1980s, I was outraged about a prison incarceration rate of 217 per 100,000). My hypothetical example of what would happen if we tripled the population was intended as just that, a hypothetical designed to show how nutty it would be to imagine that we could imprison our way out of the crime problem. Again, I did not really think we would do it. But we did, and the impacts (both physical and human) that now are evident tell me that I was disturbingly on target in predicting the real-world consequences of that choice.

I said back then that taking this road would turn out to be so costly that we would cut into every other type of public spending, including those that might really help to reduce violence. At that time in California, we spent a few hundred million dollars on our state prisons. Now, we spend $11 billion. This year in California we are cutting social programs of every size, shape, and description to cope with a budget crisis that is now the worst in recent history, but has been a persistent, nagging problem for years—a crisis that is not wholly caused by that $11 billion in prison spending, but most certainly is exacerbated by it. This recurrent drain on our public budget would be even worse (far worse) except for the fact that we actually have done this massive prison expansion on the cheap. We would need billions more, as the federal courts have forcefully insisted, to provide minimally adequate housing or health care to the prisoners we have packed into the system. Thus, we spend that $11 billion on operating costs alone while also stacking prisoners in triple-deck bunks in prison gyms, and we routinely kill several inmates every month from a lack of the most basic health care.

And the impact on communities and families? Again, I had said in the mid-1980s that the prison already "loomed" as an "ugly and pervasive presence" in the life of many disadvantaged communities and that much greater expansion of the prisons would "decimate" those communities beyond recognition. In the 1990s, a fellow went to Tallahassee, Florida, to look at what mass incarceration was actually doing to the most heavily impacted, largely African-American, communities there. What Todd Clear and his colleagues discovered can be seen today in inner cities across the country. The number of American children with a parent behind bars has reached a level I never conceived of back in the day. We now have communities that have been so shaped by their links to the correctional apparatus that the boundaries between the community and the system have become permanently blurred. As the work of Bruce Western

09012-CrimJournal-Guts.indd 5 12/21/09 7:13:51 PM

and his colleagues has compellingly shown, mass incarceration indeed has contributed to the economic "writing off" of an entire generation of minority youth.

And what of crime? In 1984, in the introduction to *Confronting Crime* (Currie, 1985), I pointed out that our homicide rate recently had been anywhere from "seven to 10" times that of many other industrial societies. At last count, our homicide death rate was roughly 9 times that of France, 10 times that of Germany, and 12 times that of the United Kingdom. In *Confronting Crime*, I said that these statistics represented a "brutal and appalling affront" to any reasonable conception of civilized life. They still do.

Our rates of serious criminal violence continue to tower disgracefully over those rates of other advanced industrial countries, despite the fact that some of these countries tragically have begun to follow many of the same social and economic policies that have made us the world leader in serious violent crime in the first place. Many of our inner cities continue to be places where the level of insecurity and routine tragedy should shame us as a nation—places trapped in what I have sometimes have called a state of "permanent emergency."

These hard realities often are obscured by the widespread celebration of the "crime drop" that took place in the early to the late 1990s. The decline from the horrific peak we reached in the early 1990s was surely significant and most welcome. But it should not obscure what might be the more significant phenomenon—the *continuity* in serious violent crime; the persistence of outsized rates of serious violence throughout rises and declines together. In the mid-1980s, I said we might see a reduction in some serious violent crimes of up to 25% from the effects of incapacitation alone. That is about what, for whatever mix of reasons, we got reported rates of homicide and robbery between 1984 and 2007. We often hear it said, with much fanfare, that crime is, in some places, "at its lowest point in 40 years," or words to that effect. But 40 years ago, those levels of crime were considered outrageous, and they helped to elect a "law and order" president.

This trend helps put the meaning of those declines in perspective. The temptation to believe that we have found that elusive "fix" for the crime problem always has been a part of the American worldview, and it recently has resurfaced with a vengeance. Not many people still believe that building more prisons is that fix, but a surprising number—mainly in law enforcement circles with some criminological sympathizers—now argue that certain police strategies are indeed the fix we have been waiting for. It even is argued that the magnitude of the crime declines in New York City and a few other places proves that we can virtually eliminate crime without addressing the social conditions that breed it (if indeed social conditions matter much at all). The fact that American levels of serious violence remain far higher than in other advanced societies seems not to trouble proponents of that view nor has the fact that no credible evidence links the nationwide crime decline to police strategies that were launched in no more than a handful of cities. A good argument can be made that smart policing can have an effect on the safety of a city (I said as much in the 1980s and have said so ever since). Indeed, *Confronting Crime* (Currie, 1985) made a point of saying that investing in front-end preventive policing strategies made a lot more sense than throwing more money at the prisons. But believing that

09012-CrimJournal-Guts.indd 6 12/21/09 7:13:52 PM

good policing can make a difference is a far cry from believing that there is a policing solution to the American crime problem.

The national-level figures as well as the striking declines that took place in some cities have tended to lull us into ignoring the sobering *unevenness* of the crime decline. Walk into the meaner streets of Philadelphia, and you will hear few cheers about how low the crime rate is. The same is true for the streets of Oakland, Chicago, New Orleans, and many other places. Moreover, as the Police Executive Research Forum pointed out in a recent series of reports, violent crime has risen sharply in several hard-hit American cities; in at least two of these cities, homicide reached its highest levels *ever* in the last few years. And this after the eight-fold national increase in the incarcerated population since the early 1970s.

Again, when I wrote *Confronting Crime* (Currie, 1985), we were in the middle of an earlier crime drop. At the time, I noted that although the decline was welcome, how long it would last was uncertain—which turned out to be true in spades as crime rates went through the roof shortly thereafter. I also pointed out that the decline was disturbingly uneven; places like Detroit, East St. Louis, and Oakland were seeing rises in violence just when the overall national rates were declining. Once again, that concern was on target. The crime drop of more recent years has left some communities still in dreadful shape. More than ever, serious violent crime is an affliction of the already afflicted—another aspect of the impoverishment of the most impoverished communities (which seems to be true in some other countries as well).

This summer, CNN produced a television news report on what it called the *Chicago's Deadly Streets*. As someone who is both of a certain age and who grew up in Chicago, it is difficult to avoid a startling (and discouraging) sense of déjà vu. The streets of my home city were deadly when I was a kid; they are more so now, especially for the young. And the faces in the program are familiar—both the perpetrators and the victims. I've seen them for 40 years. We were failing them then, and we are failing them now.

These realities suggest a more accurate, nuanced, and sobering understanding of what has happened to the rates of violence in recent years. Epidemics have come and gone since the 1960s, and we will continue to study the reasons. But the disease of violence remains endemic throughout the most impacted communities across the United States. That was the prediction in *Confronting Crime* (Currie, 1985), and I am afraid that it has been borne out dramatically. Again, we did more than triple the incarcerated population in the intervening years. We stuffed the prisons to the point that we have radically transformed large parts of urban America and have virtually obliterated the boundaries between the prison and the many poor minority neighborhoods, but we did not solve the problem. Not even close.

And the figures on the persistence of horrific levels of violence in the poor communities of Philadelphia, Chicago, or Oakland at any one point in time tell only part of this story. What we rarely factor in when we assess the effects of the anticrime strategies of the last few decades is the magnitude of the *accumulated* damage that has been inflicted already on the poorest communities in America by 25 years of failed policy. That is partly, as I have argued elsewhere, a function of how we talk about and how we measure crime. When young men are dead and

09012-CrimJournal-Guts.indd 7 12/21/09 7:13:52 PM

off the streets, we no longer count them; they fall down the national memory hole. When they are locked behind bars, we no longer count them as part of the crime problem because they no longer contribute to the "crime rate." There is no systematic accounting of the number of lives lost, disabilities suffered, and futures erased or stunted as a result of the policy choices we have made. The accumulating numbers of lives lost, disabilities suffered, and futures erased or stunted cannot be accounted for systematically as a result of the policy choices we've made.

But by any honest or moral accounting, what has been wrought in low-income minority communities throughout the past few decades must rank among the greatest social catastrophes ever to strike an advanced industrial nation. Four years ago, Hurricane Katrina exposed to a properly shocked world the conditions in which many people lived in America's Third World: it also killed a little over 1,000 people. In the years since I first began making these arguments, criminal violence has given us the equivalent impact of several *hundred* Katrinas—concentrated, like that disaster, among the most vulnerable and disadvantaged people in our society.

What might in fact be most striking about the way we think about crime after those 25 years is the extraordinary complacency with which we now respond to that grim social reality. And this, it must be said, is something I did not anticipate. Twenty-five years ago, I thought that both our level of violence and our level of incarceration represented preventable social failures that should be cause for outrage and concerted public action. Today, the level of incarceration is stunningly higher, and the level of violence is still disastrous. Yet there is little outrage, and not much action.

Back in the early 1990s, Daniel Patrick Moynihan (1993) wrote an influential article called "Defining deviancy down." His argument was that many behaviors that until recently had been considered unacceptable were now widely (if tacitly) accepted as being okay they had, in that sense, been normalized. Moynihan worried about the social implication of that "downward" definition of deviance. Not everyone agreed with him, but I think we might be seeing something similar happening today regarding the state of our social and economic health, especially the deep and chronic problems of the inner cities. We might call this, for want of a less cumbersome term, "defining disintegration down." Simply put, we now accept as routine a level of social decline and community disintegration that would have been unacceptable not long ago. Our tolerance of the continuing slaughter of minority youth in the streets is part of that "downward" definition. So is our apathetic response to the metastatic, resource-draining growth of the prisons. So, more broadly, is the fact that both persistently high rates of poverty and the catastrophic decline of legitimate work for minority youth largely have fallen off the radar screen of public policy. Today, there is little urgency about these critical, socially devastating, and costly problems in curious contrast with the 1960s—when most of them were less serious and less entrenched than today. We have, thus, "defined down" our sense of concern and ratcheted upward our level of tolerance.

The reasons for this lack of concern are complex. One of them, I think, lies precisely in the growing concentration of the risks of both violence and incarceration in America's poorest communities, which also (not coincidentally) are those with the least political voice or influence.

09012-CrimJournal-Guts.indd 8 12/21/09 7:13:52 PM

For many Americans, serious crime (not to mention the social effects of our vast increases in imprisonment) has steadily become "their" problem, not "ours."

Another reason, I think, is that our view of the trajectory of these problems is increasingly a short-term one—or at least shorter term than my own. When I look out at the students in my classes today, I am struck by the realization that virtually none of them ever have known a *different* state of affairs than the one we have now in criminal justice in America. And for that matter, neither have most Americans younger than 35, who came of age after we already had gone far down the road to unprecedented levels of mass incarceration. So they have never known a time when there was *not* a prison roughly every 40 miles down the central core of California. For them, the contrast between now and then is not a lived experience but, if anything, a more abstract understanding. They have not seen firsthand the extraordinary transformation of our institutions of justice or of our communities.

I do not want to overstate the complacency. A lot of good work is being done across the country to address these issues with the seriousness and commitment they call for by community organizations, nonprofit advocacy groups, even some legislators—and, not least, criminologists. In fact, the brightest spot for me, in what is otherwise a deeply troubling picture, is that our profession continues to do fine and important work that powerfully illuminates the state we are in and lays essential groundwork for a new kind of policy. The recent work of scholars like Todd Clear and his colleagues and Bruce Western and his, represents the best traditions of scrupulous, skilled, and engaged social science. The bad news is that the realities this careful scholarship reveals are bleak and dispiriting. And even in a time of encouraging political change, there is still no concerted or sustained effort to address them on the national level. I think it is fair to say that what I called the "conservative model" has lost credibility and momentum by now. But we have not yet put anything sufficiently compelling in its place on the level of social action and public policy.

We have an important role to play here—one that so far we have not been very good at. Now, more than ever, criminologists need to learn to become more forceful and more effective advocates for different and better policies, not just criminal justice policies. To do that, we have to push back deliberately against the forces that often pull us into a weary pessimism that can be corrosive. In the face of the magnitude of the American social crisis, it is easy to feel overwhelmed. It is seductive to narrow our concern to fine-tuning policies in and around the justice system where we think we can influence how things work, even if it is in small ways. There is nothing wrong with small changes, especially if they accumulate over time. But that cannot be all we do.

Twenty-five years ago, I argued that what kept us from seriously confronting the tragedy of American violence was neither lack of knowledge nor lack of resources but a combination of ideological resistance and political timidity and inertia. That is another piece of the argument I think was right, and I believe it is even *more* true today. We knew then a lot about the roots of the problem and a lot about what it would really take to make an enduring difference. We know much more today after 25 more years of first-rate research and the accumulation of hard-won

09012-CrimJournal-Guts.indd 9 12/21/09 7:13:52 PM

experience. And we are an even richer country with even vaster resources that we can—and have—put to use to tackle those problems we choose to take seriously. After the $1 trillion (or $3 trillion) war in Iraq—and even more definitively, after the stunning sums we just spent to bail out wayward banks deemed "too big to fail"—surely the era of claiming that we do not have the money to launch a serious attack on longstanding domestic ills is over. We managed to find the money for these dubious and almost unbelievably costly enterprises. We can find it for job creation, the alleviation of demoralizing poverty, and the reliable provision of family supports: for child abuse prevention, early education, and real help for young offenders. But I do not see anybody making that happen *for* us; you and I have got to take the lead in putting these issues, at long last, firmly on the national agenda.

I am delighted to receive the Vollmer Award and am deeply grateful to the American Society of Criminology for it. But if I get another award in another 25 years, I would like to be able to stand here and cheerfully announce that I have become obsolete…

Thank you.

References

CNN. 2009, August 22. *Chicago's Deadly Streets*.

Currie, Elliott. 1985. *Confronting Crime*. New York: Pantheon.

Moynihan, David Patrick. 1993. Defining deviancy down. *American Scholar*, 62: 17–30.[1]

Yeats, William Butler. 1921. The second coming. In *Michael Robartes and the Dancer*. Whitefish, MT: Kessinger.

Elliott Currie is a professor of criminology, law, and society at the University of California, Irvine. He also has taught in the Legal Studies Program at the University of California, Berkeley, and in the Board of Studies in Sociology at the University of California, Santa Cruz. Currie is a graduate of Roosevelt University in Chicago and received his Ph.D. in sociology from the University of California, Berkeley. Currie is the author of many works on crime, juvenile delinquency, drug abuse, and social policy that include *Confronting Crime* (1985), *Dope and Trouble: Portraits of Delinquent Youth* (1991), *Reckoning: Drugs, the Cities, and the American Future* (1993), and *Crime and Punishment in America* (1998), which was a finalist for the 1999 Pulitzer Prize in General Nonfiction. He is coauthor of *Whitewashing Race: The Myth of a Colorblind America* (2003), a finalist for the C. Wright Mills Award from the Society for the Study of Social Problems in 2004 and winner of the 2004 Book Award from the Benjamin L. Hooks Institute for Social Change. His book *The Road to Whatever: Middle Class Culture and the Crisis of Adolescence*, a study of troubled middle-class youth in America, appeared in 2005. His newest book, *The Roots of Danger: Violent Crime in Global Perspective*, was published by Prentice-Hall in 2008.

1. I am grateful to Henry Pontell for reminding me of this argument.

09012-CrimJournal-Guts.indd 10 12/21/09 7:13:52 PM

Elliott Currie's contribution to public criminology

An appreciation and a lament

Malcom M. Feeley
University of California at Berkeley

The first article I read by Elliott Currie might have been the first thing he ever published—"Crimes without criminals: Witchcraft and its control in Renaissance Europe," published in the *Law & Society Review* (Currie, 1968). It fascinated me then, and it continues to fascinate me now. It is impressive for many reasons; erudition—a study of witchcraft in the Renaissance; scope—a comparative historical study that covers changes in two locations, which span more than three centuries; acknowledgments—Howard Becker, Egon Bittner, Fred Dubow, David Matza, Shelly Messinger, and Philip Selznick; and theoretical mentors—Becker, Erikson, Kitsuse, Goffman, and Lemert. It is impressive indeed.

In the introduction, Currie (1968: 8) wrote, "In this paper, I attempt to add to the rather small body of research on kinds of social control systems and their impact. The subject is witchcraft in Renaissance Europe, and in particular, the way in which the phenomenon of witchcraft was different in England than in Continental Europe, as a result of differences in their legal systems." He then explained (1968: 8), "I will show that the English and the continental legal systems during this period represented the two ends of a continuum along which different social control systems may be placed, and I will suggest some general ways in which each kind of control system affects the deviant behavior systems in which it is involved." The article goes on to take us on a tour of witchcraft law enforcement and trial proceeding on the Continent and in England between the 15th and 18th centuries that reveals and explains why witches were much more common and treated much more harshly on the Continent than in England.

But it was Currie's (1968) conclusions and his assessment of these topics that has interested me. One conclusion was that "witchcraft in Renaissance Europe strongly reinforces . . . the argument that deviance is what official say it is, and that deviants are those so designated by officials." But his conclusion went beyond reconfirming this key insight of labeling theory. His study examined variations and sought to explain why witchcraft as well as the prosecution and

Direct correspondence to Malcom M. Feeley, Jurisprudence and Social Policy Program, School of Law, University of California at Berkeley, Berkeley, CA 94620 (e-mail: mfeeley@law.berkeley.edu).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 11 12/21/09 7:13:53 PM

execution of witches was much more prevalent (two to five times so) on the Continent than in England. In England, "the response to witchcraft took place within a framework of effective limitations on the repressive power of the legal order and a relatively advanced conception of due process" (Currie, 1968: 11). In contrast, on the Continent, "the response took place within a framework of minimal limitations on the activity of the legal system, in which the due process and legal restraint tended to go by the board" (1968: 11). But these issues were only some of the structural factors that accounted for the differences. Currie (1968: 21) continued, "[o]n the Continent the convergence of a repressive control system with a powerful economic motive created something very much like a large scale industry based on the mass stigmatization of witches and the confiscation of their property." The result was "[s]everal hundred thousand witches were burned in continental Europe during the main period of activity" (1968: 22). In contrast, witch prosecutions in England were initiated by private persons who stood to make a small profit in a precarious enterprise. As a result, Currie maintained, fewer witches were in England than on the Continent (1968: 25). Currie then goes on to theorize about "repressive" and "restrained" control systems and to generalize about what types of deviants each is likely to focus on.

After so many years, Currie (1968) probably would see labeling theory in a somewhat different light. And I suspect that recent scholarship on the prosecution of crime in the Renaissance by John Langbein (1974) and others would alter his ideas of continental justice somewhat. Of course, following Jim Whitman (2003), he would want to puzzle on the turn of events; continental justice is now much less harsh than English justice.

But this shift is not my emphasis. Here, I want to call attention to what was present from the outset in Currie's scholarship. For more than 40 years, he has focused on the following three things: the destructive effects of labeling someone as deviant, structural factors that account for various types and amounts of deviance, and structural factors that account for various forms of social control as well as the variable effectiveness of social controls. These concerns that infused Currie's (1968) first article on witchcraft have been carried over into all the countless presentations, reviews, commentaries, articles, and books on contemporary social problems, deviance, crime, and social control he has produced since. Indeed, I think it might have been this early experience as a historian examining a distant social problem that helped him develop such analytic clarity and to search for the generic and underlying social dynamics that often are not seen clearly close up. Whatever the case, Currie consistently has had an unerring eye for describing and diagnosing contemporary social problems.

Currie's (1968) first article (if it was his first) was written at a crucial moment for criminal justice scholarship. When this piece was written, Currie represented the criminology mainstream. It fit nicely within the spirit of the guiding philosophy of reports and recommendations from the President's Commission on Law Enforcement and Administration of Justice (1967) (known hereafter as the Commission), and it fit nicely with the prevailing sentiment of the leading scholars and practitioners of criminal justice administration in the United States. The Commission (1967) focused on structures and institutions; it suggested that long-term and

09012-CrimJournal-Guts.indd 12 12/21/09 7:13:53 PM

lasting efforts were needed to combat crime. Crime—at least much serious and violent crime—was understood as part of a larger complex of social evils that were a product of dysfunctional lives, institutions, processes, and structures. Crime was perceived as but one of several interrelated social problems that plagued a resource-rich America. The crime policy envisioned in the Commission's (1967) reports and the policy that informed Currie's vision flowed from the perspective that had informed the governing framework of the Great Society, which was an expansion and adaptation of the emerging social contract that had been constructed in the New Deal and sketched out even earlier in the Progressive Era. That is, it was part of a long and largely consensual tradition.

By the time the Commission's (1967) reports were published, they were dated. Similarly, by the time Currie's (1968) article was published, the landscape had been altered radically. They appeared in the midst of, or just after, a long presidential campaign that witnessed the early successes of Eugene McCarthy in the Democratic primaries, the assassination of Robert Kennedy on the night of his primary victory in California, George Wallace's strong showing throughout the country, another escalation of the Vietnam War, and the election of Richard Nixon based on his emphasis on law, order, and an appeal to racism.

A sea of change had occurred almost instantaneously. The dominant views and voices that had informed the findings and reports of the Commission (1967) were drowned out or discarded to the dust bins of history before the ink had dried on them. The election of Nixon and his War on Crime took root in the way it had not in Barry Goldwater's losing presidential campaign 4 years earlier. Its covert appeal to racism to discredit the Civil Rights Movement and Great Society programs drove a wedge in the working class and played on racial resentments and insecurities that were a response to the Civil Rights Movement and the failings of the war in Vietnam. The war that could not be won, however, stoked a War on Crime that promised victory. Progressive social scientists and their structural analyses were shunted aside, or they retreated from their roles as public intellectuals as well as from their access to major news media and the councils of power. They ended up talking to themselves or taking on roles of social critics and largely were marginalized. In Berkeley, it was even bleaker; criminologists in the forefront of war resistance faced an even fiercer resistance.

Writing in another context, political scientist Stuart Scheingold (1975) characterized this transformation in more genteel terms. He noted that two traditions exist in American criminology and that they play out in different ways. One emphasizes structural factors and conditions that foster crime, poverty, and social problems in a way that suggests that, to be effective, social policy must attend to these root causes. The other is the "volitional" school of social thought that emphasizes individual freedom and choice, which in turn, suggests that one might affect social policy by altering the incentives of individuals.

In Scheingold's (1975) view, the transformation in social thought that was stroked by the social transformations of the late 1960s was the shift away from structural explanations of social problems to volitional explanations—an approach that focused on individual choice not on social ecology. Individuals making bad choices, not "structural" factors, led to decisions to burgle

09012-CrimJournal-Guts.indd 13 12/21/09 7:13:53 PM

houses, pull triggers, snatch purses, depend on welfare, and eschew employment. Scheingold (1975: 4) argued that "[s]tructural explanations focus on social disorganization with its roots in hierarchy, deprivation, coercion, and alienation. [In contrast, volitional criminology] associates street crime with individual pathologies—be they moral, emotional, or genetic." More generally, Scheingold (1975) maintained that volitional explanations were more compatible with the cultural conservatism that began to take deep root in the United States in the late 1960s and accelerated throughout the 1970s and 1980s.

So, just as Elliott Currie had developed his voice, his audience—or a huge slice of it—that was attuned to a social-factors approach, was lost. A new language suddenly was being spoken. The advocates of the new volitional social theory were the Neo-Conservatives who coalesced around support for the Vietnam War and opposition to the Great Society. However, their new volitional theory of criminal behavior might have been shaped by their acceptance of the war. Under the inspiration of Secretary of Defense, Robert McNamara, the enemy was individualized. If we could inflict enough pain on them, then the theory held; they would retreat and sue for peace. Robert McNamara assured us that if the kill ratios and body counts could be increased sufficiently, then the will of the enemy would be broken, and we would triumph. Although this way of thinking was soon discarded as an obvious failure in Vietnam, Secretary McNamara's ideas nevertheless were picked up by some of his whiz kids who transferred their skills to crime fighting. The Law Enforcement Assistance Administration emerged and laid the groundwork for a new kind of smart war on crime—one anchored in volitional as opposed to structural criminology. They reasoned that if salience and celerity of the criminal sanction were increased (i.e., increase body counts), then we could prevail in our domestic war against crime. Just as the Pentagon had declared that a dead enemy could not fight back, prosecutors of the new war on crime argued that an incapacitated person could not commit street crimes. As I mentioned, Pentagon officials gave up on this approach after a few hopeless years in Vietnam, but it has remained the dominant crime-fighting strategy for 40 years.

Thus, Currie's (1968) debut as a scholar advanced an outlook that was out of public fashion at the outset. With the consolidation of the new war mentality, the legitimization of racism, the demonization of the Great Society, and the lowered expectations of what government can do, volitional criminology emerged in full force within the public arena. Of course, structural criminology did not disappear, but it had lost most of its public audience.

Despite these challenges, Currie has continued to warn of the futility of the volitional approach, and across the years has come to seem like a prophet, especially now with his shock of white hair to match his resonant baritone voice. And like many prophets, he has been less honored at home than abroad. As sonorous as it is, his voice, for the most part, has fallen on the deaf ears of policy makers here in the United States, whereas abroad, he is regarded as the leading American spokesperson for a reasonable crime policy.

09012-CrimJournal-Guts.indd 14 12/21/09 7:13:53 PM

The Debate that Never Was

If Currie's voice was not heard and if structural theories of social problems were suddenly displaced by volitional theories, then whose voice did capture the public sphere? What theorists and public intellectuals emerged to replace structural theorists, like Currie, whose views had been part of a rough national consensus since the Progressive Era? Who were the theorists of lowered expectations for social policy and the policeman state for the Thatcher–Reagan Era? In his analysis of the structural and volitional theories, Stuart Scheingold (1975) argued that reason was largely set aside as the War on Crime heated up. He believed that, at the national level, reason gave way to political exploitation of growing insecurities about social change among the middle class and resentful white working class. Still, he argued that the political and cultural resonance of volitional criminology assigned Currie as a representative of the structural approach, which lost its dominance in the 1960s, and his exemplar of volitional criminology was political scientist James Q. Wilson. Scheingold's (1975) insight hit the mark. Wilson's (1975) collected essays on crime, *Thinking about Crime*, were published about the time Currie was hitting his stride. But it was Wilson's voice that ascended. In his essays, Wilson (1975) emphasized the limited effectiveness of treatment, the failures of probation, as well as the futility of looking for root causes and structural solutions. He emphasized the value of increasing the likelihood of apprehension, conviction, and incapacitation. Indeed, he took issue with those (like Currie) who would emphasize the value of improved education, increased employment opportunities, and so on. He acknowledged that these improvements might be good policies but claimed his focus was on what criminal justice officials—cops, prosecutors, and judges—could do within their limited range of responsibilities. His answer was to increase the likelihood of catching and punishing criminals.

I suspect that Currie would agree with much of what Wilson (1975) had to say on this subject. Indeed, Currie's views on poverty and crime were subjected to an informative exchange with Christopher Jenks in the *New York Review of Books* where the limited roles of law-enforcement officials and the complex relationship between poverty and crime were recognized. And Wilson's acknowledgment suggested that it would be beneficial to find ways to abate poverty and assimilate the marginalized into the mainstream. So although stark disagreements occur, zones of agreements and points of commonality do exist.

However, an arena for public debate has never been made available—an arena of public criminology performed in public spaces and attended by public intellectuals, policy makers, and opinion leaders. For the past 40 years, in almost all areas of social policy, the debate has been largely one-sided about variants of Neo-Conservative ideas. The differences and the overlaps between Currie and Wilson—and a dozen other social theorists in both camps—never were joined and never were debated. No dialogue has taken place, nor have policy makers searched for common ground. It is as if Vietnam drew lines too sharply. The war drove a wedge between public intellectuals. It hardened positions. People chose sides. Friendships were lost. Points of view that once overlapped were restated as exclusive. The middle ground was lost. Neo-Conservatives broke with the liberals. Daniel Patrick Moneyhan, the gregarious former

09012-CrimJournal-Guts.indd 15 12/21/09 7:13:53 PM

Harvard Professor turned New York Senator, was one—maybe the only—well-placed public intellectual who continued to collaborate with both camps. For years, at least on matters of crime and social policy, Neo-Conservative volitional criminology has been dominant. Wilson wrote for the *Public Interest, Commentary*, and the *New York Times Magazine*, publications that were read in the White House and by Cabinet secretaries. Currie wrote for *Dissent, the Nation*, and *Theoretical Criminology*—good publications that were read by a small group of the committed. Wilson counseled presidents and was fêted in the White House; Currie advised local governments and was feted by the American Society of Criminology. Currie has not had the bully national pulpit his ideas command.

This division led to a lost opportunity—an inability to search for common ground between the children of the Great Society and some of the most thoughtful and hard-headed realists who questioned the scope of long-standing approaches to social policy. It paved the way for the Right—especially for the extremism of the Right in the War on Crime—to jettison all rational discourse about crime and justice, and to "govern through crime" (Simon, 2007).

This claim is an overstatement, and I do not want to end on such a pessimistic note, especially at this time in American political life. Of course, Currie has made an important difference in many settings. His advocacy for sensible crime policies and programs in the Bay Area as well as in many other areas has made a difference. He has educated generations of students to think about the problem of crime and what to do about it in a more sensible manner. His work with the Milton Eisenhower Foundation has affected hundreds, if not thousands. Even if his work is not familiar to officials in the U.S. Department of Justice, his work has influenced thinking in the Home Office in London and ministries of justice elsewhere. His many honors, which include his near-miss at the Pulitzer Prize, are ample testimony to his many contributions and stature. But still, I cannot help thinking how things might have been different if Currie had had the bully national pulpit his work deserves.

But take note: Although Currie's approach is anchored in the old-fashioned public sociology of the Progressive-Era Chicago School; it is also enlightened and advanced. It will have an influence. Earlier, I said that Currie is something of a prophet, and I meant it. Like many prophets, he might not be honored in his own time, but it does not mean that he will not be honored. Like August Vollmer, the namesake of the award he has received, Currie is a thinker filled with common sense and practical vision but, nevertheless, is a thinker ahead of his time. One looks forward to when this fact is recognized widely beyond the confines of the American Society of Criminology and embraced generally. Certainly, when we do get serious about structural approaches to solve our many social problems, Currie's works will be pored over and the many lessons he taught will be appreciated more widely. In the meantime, we do what we can. We salute Elliott Currie within the Society for his many and lasting contributions to public criminology.

09012-CrimJournal-Guts.indd 16 12/21/09 7:13:54 PM

References

Currie, Elliott. 1968. Crime without criminals: Witchcraft and its control in Renaissance Europe. *Law & Society Review*, 3: 7–32.

Currie, Elliott and Christopher Jenks. Genes and Crime: An Exchange. New York Review of Books, 34 (10): 36-37, June 11, 1987.

Langbein, John. 1974. *Prosecuting Crime in the Renaissance: England, Germany, France*. Cambridge, MA: Harvard University Press.

President's Commission on Law Enforcement and Administration of Justice. 1967. *Crime in a Free Society*. Washington, DC: U.S. Government Printing Office.

Scheingold, Stuart. 1975. *The Politics of Law and Order*. Philadelphia, PA: Temple University Press.

Simon, Jonathan. 2007. *Governing through Crime*. New York: Oxford University Press.

Whitman, James Q. 2003. *Harsh Justice: Criminal Punishment and the Widening Divide between America and Europe*. New York: Oxford University Press.

Wilson, James Q. 1975. *Thinking about Crime*. New York: Basic Books.

Malcolm M. Feeley is the Claire Sanders Clements Dean's Chair Professor of Law (Boalt Hall) at the University of California—Berkeley. Since 1984, he has been associated with the Jurisprudence and Social Policy Program there. From 2005 to 2007, he was president of the Law & Society Association, and he currently serves as coeditor, with Jonathan Simon, of the journal *Punishment & Society*. The author or editor of numerous books and articles on the judicial process and the criminal justice system, Feeley's 1979 book, *The Process is the Punishment*, received the ABA's Silver Gavel Award for best book in law. His most recent books are (with Edward Rubin) *Judicial Policy Making and the Modern State* (Cambridge); *Fighting for Political Freedom: Comparative Studies of the Legal Complex and Political Liberalism* (with Terry Halliday and Lucien Karpik, Hart); and *Federalism: Political Identity and Tragic Choice* (with Ed Rubin, Michigan). Feeley received his Ph.D. in political science in 1969 from the University of Minnesota.

09012-CrimJournal-Guts.indd 17 12/21/09 7:13:54 PM

Elliott Currie

In tribute to a life devoted to confronting crime

Francis T. Cullen
University of Cincinnati

I have known Elliott Currie for nearly 25 years. In 1986, I had the good fortune to be invited to attend a conference carrying the title of my book with Karen Gilbert, *Reaffirming Rehabilitation* (Cullen and Gilbert, 1982). This was heady stuff for a still-young scholar. But the conference also meant that I would meet two of my heroes—the gathering's organizer Jerome Miller, known for his bold experiment in deinstitutionalizing virtually all youths in Massachusetts's juvenile reformatories, and Elliott Currie, the recent author of *Confronting Crime: An American Challenge* (1985). In his somewhat rundown car, Jerry Miller shuttled Elliott and me around to dinner, a gathering at the National Press Club, and a radio show interview—in addition to the Reaffirming Rehabilitation conference. In short, I got to "hang out" with Elliott for a couple days.

I have had occasion to know scholars that are simply beyond my capacities—Robert Merton, a professor of mine at Columbia; Richard Cloward, my dissertation mentor; and Gresham Sykes, a coauthor on a textbook. I am not naturally brimming in confidence, but I have enough self-efficacy not to be easily star struck. Still, these scholars think on a different level and have the ability to communicate ideas with inordinate clarity and power. I soon learned from my time with Elliott—watching his presentation, hearing his answers to provocative questions from reporters, and listening to him handle callers to the radio show—that he was one of these special scholars. It was like playing on a team with Michael Jordan; sometimes it was best simply to sit back and marvel at the performance.

Why was he so special? Eloquence, substance, and heart—it is this trinity of characteristics, I believe, that combine to make Elliott perhaps the most important criminologist of my generation. I invite anyone to listen to Elliott speak or to read his many works. The words flow, the images vividly jump into the mind of the audience, and the message is readily understood. But eloquence without substance is slickness, an invitation to hucksterism. Elliott is a gifted sociologist, someone who knows American society and how the nation's stubborn realities af-

Direct correspondence to Francis T. Cullen, School of Criminal Justice, PO Box 210389, University of Cincinnati, Cincinnati, OH 45221-0389 (e-mail: cullenft@ucmail.uc.edu).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 19 12/21/09 7:13:54 PM

flict us with crime generally and concentrate violence among our most disadvantaged brethren. His criminological substance is indisputable. Underlying it all, however, is his heart—his caring. Elliott's criminology is not simply a product of his mind but of his passionate belief that needless human suffering should not be tolerated. A restrained anger informs and enlivens his writing—almost a disbelief that in a nation as affluent as the United States, we permit young inner-city boys to die day after day and treat this behavior as "normal." He is moved to deconstruct conservative crime ideology, showing how it comprises comforting techniques of neutralization that allow the banality of evil to persist.

I have no doubt that *Confronting Crime* (1985) is the greatest criminological work ever published. Let me hasten to say that I am not offering inflated praise as part of my task to celebrate Elliott's selection for the Vollmer Award. I truly mean what I say. When *Confronting Crime* was published, criminology was on the verge of irrelevance. The Reagan Administration and the bold trumpeting of conservative ideas were flourishing. More disturbing, James Q. Wilson's (1975) critique of criminologists had merit. Our search for the root causes of crime had produced few feasible and effective crime control policies. We preached to our criminological choir and cheered ourselves for doing so, but our thinking was often sloppy and we did little to save offenders or their victims from criminal harms. We criticized Wilson for suggesting that the government should use available resources to manage the crime problem, but then we would claim that nothing the state did worked to control crime (Cullen and Gendreau, 2001). In fact, it seems that the failure of criminal justice interventions was preferred because it would, by implication, prove the point that state actions were inherently ineffectual and that crime was, after all, a matter of root causes.

Confronting Crime (1985), however, escaped this ideological trap. Elliott wisely understood—as many scholars now do today—that a root cause analysis of crime is not inconsistent with practical strategies to reduce offending and victimization. It was not so much that he proved Wilson (1975) wrong—again, I think Wilson's views should be taken seriously—but that he showed that a managerial criminology was not the only way to understand and blunt the effects of crime. In so doing, Elliott revitalized a sociological analysis of crime and opened up fresh vistas for confronting crime in more effective and humane ways.

When I initially read *Confronting Crime* (1985), I was simply astounded by its brilliance. In a very real way, I felt rescued from my deep criminological depression—my sense that, as a field, our answers about the origins of crime were stale and that our visions for crime control lacked relevance. Elliott, I suppose, gave me hope—he inspired me to believe that, as a discipline, we retained the intellectual power and human spirit to make a difference. Revisiting his work, I am nourished in the same way. His insights are as true today as they were in 1985. His warnings to avoid certain policy roads in favor of others not only were prescient in the 1980s but also should be heeded now. *Confronting Crime* thus has enduring value; it is a book about America—about our social failures, about our misguided thinking, and about ways to

09012-CrimJournal-Guts.indd 20 12/21/09 7:13:54 PM

lift ourselves out of this policy morass. Anyone who has not studied this volume should rush out and do so immediately.

On the first page of *Confronting Crime*, Elliott noted that "this book is about why there is so much crime in America and what we can do about it" (1985: 4). These two issues—high crime and how to reduce it—have remained central to Elliott's criminology during the past 25 years. Below, I will comment on these matters and the core lessons Elliott teaches.

Why There Is So Much Crime in America

Elliott was one of the first criminologists to understand the power of a comparative perspective. He was quick to observe that advanced Western societies—such as those in Japan, Canada just to the north, and Western Europe—were similar to the United States in many ways. They were democracies, had market economies, were modern, and were pushing into post-industrialism. Despite this fact, these nations' violent crime rate—especially homicide—comprised a small fraction of the U.S. crime rates. How could this be? What was distinctive about the United States that made us criminologically anomalous?

Starting in *Confronting Crime* (1985) and developed further in later works, Elliott highlighted the distinctiveness of American capitalism. His argument was not that capitalism causes crime—a crude and falsifiable claim—but that a certain kind of market economy was insidiously criminogenic. There was nothing inevitable about this economic arrangement; capitalism is not a natural phenomenon but is produced and shaped by policy decisions. Countries get the brand of capitalism they choose.

In the United States, we have chosen an economic system that is marked by two traits. It produces high levels of inequality between the rich and the poor, and it is stingy in the supports it gives to those who are at the bottom of this hierarchy. In short, our market economy provides enormous rewards for those who succeed and few buffers for those who fail. Notably, this societal recipe also prevails in many less advanced or third world nations. By contrast, most advanced Western nations limit levels of inequality and provide more supports for the disadvantaged (e.g., child care, health care, and monies for the unemployed). The punch line is, of course, that nations that are more equal and more generous in their social supports have low rates of violent crime, whereas the United States and countries that embrace inequality and stingy social welfare disproportionately jeopardize the lives and health of their citizens. Elliott has an incomparable way of making his point (Currie, 2009: 28–29):

> Indeed, the United States' homicide rate puts us above some of the countries in the developing world—including Chile, Costa Rica, and Argentina…. If the United States had the homicide rate of the average European high-income country—about 1 per 100,000—we would avoid 5 out of every 6 of our homicides, thus saving more than 13,000 American lives every year, or more than four times the number who died in the terrorist attacks of September 11, 2001…. An American female ages 15 to 19 is three times more likely to be murdered as an English male, nearly five times as likely as a French male, and nine times as likely as a Japanese male.

09012-CrimJournal-Guts.indd 21 12/21/09 7:13:54 PM

Importantly, Elliott is not some crass or lazy Marxist who decries the criminogenic effects of inequality and then assumes that the theoretical heavy lifting is complete. Rather, much of his work has been devoted to illuminating what it is about American-style inequality that is socially pernicious and, in turn, criminogenic. As Elliott has long understood, the market economy is not an isolated social institution that can be examined in and of itself (see also Messner and Rosenfeld, 1994). Rather, a market economy that breeds inequality is propped up by policies that benefit some Americans more than others and, given its dominance, has dire consequences for other social institutions—most notably labor, community, and family. Elliott has brilliantly demarcated these institutional linkages as well as how this fosters high crime rates, especially in his under-read 1997 article in *Theoretical Criminology*, "Market, Crime and Community" (Currie, 1997; see also Currie, 1985, 1989, 2009). I will not repeat them here, but I will outline what I see to be his core insight.

To wit: What is life in any society, including in the United States? We work, we reside in communities, and we rise and return home each day to our family. This is the "stuff" of life. It is what intimately affects our children's development, our daily routines, our feelings about ourselves, our sense of the goodness or the unfairness of our society, our visions of what the future holds, and, ultimately, the behavioral choices we make. When these fundamental elements of a meaningful and healthy life go awry—when the quality of our work, community, and family is eroded—sustaining a "regular" or conventional existence is difficult. When most of those around us also suffer a similar fate, then structures of social dysfunctions concentrate and create environments that are criminogenic.

In the United States, economic policies have exposed many people to work that lacks meaning and reward—that is low paying, has few benefits, is of uncertain stability, and offers little prospect of future advancement. Some communities, especially those populated by minorities, have lost their industrial base and have labor markets unable to employ able-bodied adults. Families, especially those disrupted and located in areas of concentrated disadvantage, are left largely to fend for themselves. Economic supports are short lived, and the value of collective responsibility for one another—the sense that we are all part of the same "village"—is derided as "socialist." As Elliott shows, the cost of marginalized work, of communities unable to employ its populace, and of families lacking breadwinners and teetering on the brink of despair is all but too predictable. It is an American recipe for high rates of violent crime.

What We Can Do About Crime

Elliott is a careful and thus reasonable scholar. He understands that opposing scholars often "have a point"—that is, that they capture a kernel of truth that is foolish to ignore. Thus, he would not suggest that prisons have no effect on crime or that offenders are normal, which would be an inappropriate "denial of pathology" (Currie, 1985). Still, he takes issue with an approach to crime—which he labels the "conservative model"—that has unfettered confidence that we can punish our way out of the crime problem (Currie, 1985, 1998).

09012-CrimJournal-Guts.indd 22 12/21/09 7:13:54 PM

I have this sense that Elliott is simply bewildered by such thinking. The conservative model has this unerring faith that if we punish kids enough in the home when they are naughty and throw them into jail when they are older and naughty, then things will be much safer. What about middle-class kids who become strung out on drugs or miss the train leading to upward mobility? Well, they need "tough love," including parental abandonment so that they can "hit bottom" before rising like a phoenix from the ashes to a renewed life of self-regulation (Currie, 2004). Again, even if some truth might reside in such thinking, the narrowness of this view of human development and of human change is astounding. Indeed, when in trouble, few of us say: Please hit me and threaten me with bodily harm so I can go straight. Please toss me onto the street so I can hit bottom. Please put me in a boot camp so I can be broken down and built back up. Please lock me up for 5 or 10 years so I can learn that crime does not pay. And if not me, then do it to my wayward child or to my troubled brother, sister, or high school chum.

Beyond control, there is another essential ingredient of healthy development: love, caring, and social support (Cullen, 1994). We do not just discipline our kids but love them; we do not just take responsibility for our misconduct but hope that others will help us through our difficulties; we do not let family members go hungry and homeless but take care of them. Why? Because we understand that humans are nourished by intimacy, affection, and assistance. The conservative model, however, assumes that meanness is what ensures conformity. It privileges the value of individualism so much that it denies that any kind of social welfare can benefit those in trouble or in need. It strips people of their social nature and treats them as though they respond only to the threats we hang over their heads.

In his many writings, Elliott has taken great pains to deconstruct logically and empirically the conservative model. His critique of the conservatives' "prison experiment"—the idea that mass incarceration achieves mass reductions in crime—has perhaps been unrivaled in its clarity and persuasiveness (Currie, 1998). But he has critically evaluated the conservative model not only to show its faults but also to create the space needed for us to consider alternative crime control strategies. Elliott thus has realized that the conservative embrace of punishment and of imprisonment is dangerous both for what it privileges (locking up offenders) and for what it precludes (socially attuned interventions). Thus, showing the model's limits is essential if we are to appreciate that mass imprisonment is not the only policy option—that the possibility exists to build a more progressive approach to mitigating America's crime problem (see also Listwan, Jonson, Cullen, and Latessa, 2008).

James Q. Wilson's (1975) critique of the policy irrelevance of root-cause thinking about crime should be taken seriously. Again, Wilson claimed that even if criminological theories have some merit, they identify causal factors that cannot reasonably be changed through governmental policies and programs. For Wilson, one of the few governmental resources that could be manipulated was the use of imprisonment, which is why he advocated using available prison cells to selectively incapacitate the "wicked." Thus, in one devastating stroke, Wilson had dismissed criminology as policy irrelevant and elevated imprisonment to the centerpiece of any feasible effort to reduce the victimization of innocent citizens.

09012-CrimJournal-Guts.indd 23 12/21/09 7:13:55 PM

Elliott's (1985) rebuttal to Wilson's (1975) perspective and to the conservative model more generally was three-fold. First, he showed that its major policy prescription—mass imprisonment—was only marginally effective but hugely costly. In short, it was bad social policy. Second, as noted, he showed how violent crime in America was fundamentally intertwined with the constitutive elements of daily life in America—the quality of our work, our community life, and our family well-being. Whether or not we wish to label these "root causes" is beside the point. They are the social features of our existence that shape our life-course trajectories. And third, he unmasked as foolish the idea that calling a source of crime a "root cause" meant that it was beyond intervention.

Let us take, for example, the claim that inequality is a root cause of violence in America. Elliott would not, of course, assert that to reduce the criminogenic effects of inequality, we need to have a socialist revolution in which all wealth is evenly distributed. Rather, the key is to identify the conduits through which inequality leads to crime. It might be possible, therefore, to blunt inequality's criminogenic effects by raising the minimum wage, providing disrupted families with quality child care, creating early intervention programs for youngsters in impoverished at-risk communities, or implementing quality rehabilitation programs for juveniles entering the justice system. In fact, Elliott has conveyed a systematic blueprint for a multilevel approach to reducing crime through correctional, prevention, and social welfare programs. Demonstrably, he has done so by identifying interventions that, either in the United States or in other nations, have been shown empirically to be effective. That is, Elliott has revealed an evidence-based, practical agenda for reducing crime through nonpunitive means (e.g., Currie, 1985, 1989, 1993, 1998).

In short, Elliott's special contribution is that he has shown how good criminology is the basis for good crime control policy. In our discipline, we at times attempt to treat criminology and criminal justice as though they are autonomous scholarly domains. But in the real world, such distinctions are not only foolish but also dangerous. Why crime occurs and what we can do about it are not, as Elliott so fully understands, separate problems but inextricably linked. A narrow or incorrect explanation of crime, such as the one that informs the conservative model, risks justifying limited, overly punitive policies that can exact huge costs in lives and public treasure. By contrast, an explanation of crime that is multifaceted, integrated, and grounded in our best data allows us to implement policies that target crime's true causes for change.

I will not claim that Elliott has "all the answers." As Merton (1973) noted some years ago, organized skepticism is a key scientific value, and it should lead us to scrutinize all scholars' work—including those we find most convincing. Still, Merton also reminded us that scholars can see farther by standing on the shoulders of giants in their field of inquiry. It is clear that Elliott is indeed a scholarly giant and that the horizon of our criminological vision is extended greatly by reading his work and thus by standing on his shoulders.

09012-CrimJournal-Guts.indd 24 12/21/09 7:13:55 PM

Conclusion: Beyond Antiseptic Criminology

I am guilty of being a practitioner of "antiseptic criminology." In part, this involves being an armchair or office-based scholar in which I avoid talking with offenders and going onto dangerous streets or into dreary prisons. But more than this is involved. I also embrace a brand of criminology that devotes itself to theory testing. When ambitious, this practice means developing my own surveys that operationalize competing theories of crime (and their core variables) and venturing out into high schools to collect data (or doing so vicariously by making my graduate students, not me, go to the schools). When less ambitious, this practice means finding secondary data sets that contain a four-item measure of self-control and perhaps a one-item measure of how many delinquent peers a sample member has—all so I can undertake the one hundredth "critical test" of control theory versus learning theory.

My intent is not to mock my own scholarly predispositions or those of most other criminologists who engage in the normal science practice of theory testing through primary or, more often, secondary data sets. This criminology is worthwhile and, beyond earning many of us tenure and promotion, has allowed for the accumulation of a wealth of informative data. Still, the weakness in this criminological style is that it privileges the analytical over the real world. That is, it seeks to decontextualize individuals and decompose them into a list of variables that the data set allows to be measured.

Again, this scholarly enterprise has some functional use. But it is *antiseptic* in that it removes people from the real world. As a result, in most studies, we know nothing about the labor market in the areas where our sample members reside, the level of child care and health care supports in their neighborhoods, the housing policies that shape where they live and what toxins they are exposed to, the noise and incivility they might encounter each day, the athletic leagues that may or may not be available, or the enduring cultural traditions that are not chosen but welcome them when they enter the world. In many of our studies, we would have no idea from reading the results and discussion sections where and when the research was undertaken. We might learn that low self-control and antisocial peers both are related to delinquent participation—a finding of some note—but this empirical tidbit would be shorn of any sense of social, political, or historical context.

By contrast, Elliott has avoided what C. Wright Mills once called "abstracted empiricism." Wiser and less regulated by the discipline's normal science paradigm, he has chosen the road less traveled. Indeed, Elliott has avoided my path—*antiseptic criminology*—and walked a different path, what I call *American criminology*. By "American criminology," I mean that his writings start with an understanding of how crime, especially violent crime, is distributed within the United States and in comparison with other nations. He then asks us to consider why this is so—not to see it as a natural occurrence but to see it as the predictable product of how we choose to arrange our society. In particular, he pushes us to probe the nature and effects of America's core institutions—our unequal class structure, our labor markets, our families, our communities, and our politics and the policies that undermine many people's quality of life. Elliott reminds

09012-CrimJournal-Guts.indd 25 12/21/09 7:13:55 PM

us that we live in a particular socio-historical time in which the lives of Americans are shaped by forces beyond their control. Yes, pathological individuals exist who are, in important respects, the architects of their own failure and harmful criminal acts. But their pathology is not all of their own choosing, and the choices they make are always bounded by the structure of opportunities and provision of social supports that have been present or absent from their first breath of life. To ignore this reality is to commit what Elliott calls the "fallacy of autonomy"—the tendency to focus on individual defects and to ignore that people are deeply enmeshed in a web of social contexts that delimits who they are and what they can become.

Although Elliott does not criticize our discipline for being antiseptic, his writings during the past 25 years remind us that more scholars should have followed his example—spending less time conducting self-report studies of high-school students and more time learning about the social forces shaping and reshaping our society. In the end, Elliott's work has relevance precisely because it illuminates the real America in which we and those we purport to study live. Much like a Michael Moore documentary film, his writings force us to pay attention to uncomfortable truths about the United States—about, as a people, how we choose high inequality, how we marginalize many workers, how we turn a blind eye to children living in dire poverty, how we try not to talk about the close tie between race and enduring concentrated disadvantage, and how we have used mass imprisonment as a solution for our social failures. These are the realities that an American criminology would place at the center of its inquiries and take into account as it articulates a coherent and feasible crime control agenda. In this context, I thus will leave you with one final prescient quote from Elliott, which appropriately also ended his classic book, *Confronting Crime* (1985: 278):

> If we wanted to sketch a hypothetical portrait of an especially violent society, it would surely contain these elements: It would separate large numbers of people, especially the young, from the kind of work that could include them securely in community life. It would encourage policies of economic development and income distribution that sharply increased inequalities between sectors of the population. It would rapidly shift vast amounts of capital from place to place without regard for the impact on local communities, causing massive movements of population away from family and neighborhood supports in search of livelihood. It would avoid providing new mechanisms of care and support for those uprooted, perhaps in the name of preserving incentives to work and paring government spending. It would promote a culture of intense interpersonal competition and spur its citizens to a level of material consumption many would not lawfully sustain.
>
> Does that model fit any societies in the real world? I think it does. It closely resembles some of the developing countries of the Third World—Brazil, the Philippines, and parts of Mexico—which have among the worst rates of criminal violence in the world. But we don't have to go that far: It looks uncomfortably like the United States in the 1980s—and like one version of our own future.

09012-CrimJournal-Guts.indd 26 12/21/09 7:13:55 PM

References

Cullen, Francis T. 1994. Social support as an organizing concept for criminology: Presidential address to the Academy of Criminal Justice Sciences. *Justice Quarterly*, 11: 527–559.

Cullen, Francis T. and Paul Gendreau. 2001. From nothing works to what works: Changing professional ideology in the 21st century. *The Prison Journal*, 81: 313–338.

Cullen, Francis T. and Karen E. Gilbert. 1982. *Reaffirming Rehabilitation*. Cincinnati, OH: Anderson.

Currie, Elliott. 1985. *Confronting Crime: An American Challenge*. New York: Pantheon.

Currie, Elliott. 1989. Confronting crime: Looking toward the twenty-first century. *Justice Quarterly*, 6: 5–25.

Currie, Elliott. 1993. *Reckoning: Drugs, the Cities, and the American Future*. New York: Hill and Wang.

Currie, Elliott. 1997. Market, crime and community: Toward a mid-range theory of post-industrial violence. *Theoretical Criminology*, 1: 147–172.

Currie, Elliott. 1998. *Crime and Punishment in America*. New York: Metropolitan Books.

Currie, Elliott. 2004. *The Road to Whatever: Middle-Class Culture and the Crisis of Adolescence*. New York: Metropolitan Books.

Currie, Elliott. 2009. *The Roots of Danger: Violent Crime in Global Perspective*. Upper Saddle River, NJ: Prentice Hall.

Listwan, Shelley Johnson, Cheryl Lero Jonson, Francis T. Cullen, and Edward J. Latessa. 2008. Cracks in the penal harm movement: Evidence from the field. *Criminology & Public Policy*, 7: 423–465.

Merton, Robert K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago, IL: University of Chicago Press.

Messner, Steven F. and Richard Rosenfeld. 1994. *Crime and the American Dream*. Belmont, CA: Wadsworth.

Wilson, James Q. 1975. *Thinking about Crime*. New York: Vintage.

Francis T. Cullen is Distinguished Research Professor in the School of Criminal Justice at the University of Cincinnati. Recently, he has co-authored *Unsafe in the Ivory Tower: The Sexual Victimization of College Women*, *The Origins of American Criminology*, and the *Encyclopedia of Criminological Theory*. His current research interests include the impact of race on public opinion about punishment, rehabilitation as a guide for correctional policy, and building a social support theory of crime. He was president of both the American Society of Criminology and the Academy of Criminal Justice Sciences.

09012-CrimJournal-Guts.indd 27 12/21/09 7:13:55 PM

Elliott Currie

August Vollmer Award winner, 2009

Jerome H. Skolnick
New York University

I have known the ever youthful Elliot Currie for more than 40 years. President Lyndon B. Johnson established The National Commission on the Causes and Prevention of Violence after the assassination of Robert Kennedy on June 6, 1968. The Commission, chaired by Milton Eisenhower, appointed Marvin Wolfgang and Jim Short as its social-science directors and me as the head of a research "task force" that focused on anti-Vietnam war protest, student protest and civil-rights protest. Six months later, the task force submitted a report that was published by Simon and Schuster and Ballantine Books in June 1969 as *The Politics of Protest*. With a new introduction, *The Politics of Protest* will be published approximately 40 years later in February 2010 by New York University Press.

Why is this relevant to Elliott Currie? I needed an assistant to help with the research and writing of this hurry-up report. Sheldon Messinger strongly recommended that I hire a graduate student whom he described as "brilliant" to help prepare the report. The contribution of this graduate student became so valuable that the report acknowledges him not as a research assistant but as assistant director of the task force. Thus began a collaboration (with a book of readings on America's social problems, *Crisis In American Institutions*, now in its 13th edition) and a friendship that has lasted for more than 40 years.

Elliott's contributions to *The Politics of Protest* (National Commission on the Causes and Prevention of Violence, 1969) and to *Crisis in American Institutions* (Skolnick and Currie, 2006), however valuable, have been a side project to his career as an internationally renowned theorist and critic of American crime, drug, and imprisonment policies as well as a researcher of middle-class culture and adolescence within it.

Currie on Crime and Punishment

Currie's writings on crime and punishment have had consistent themes. Foremost is the idea that the severity of punishment cannot, and should not, be the principal public strategy to achieve public safety. It was not entirely surprising when his book *Crime and Punishment in America*

Direct correspondence to Jerome H. Skolnick, New York University School of Law, 40 Washington Square South, 330B, New York, NY 10012 (e-mail: Jerome.Skolnick@nyu.edu).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 29 12/21/09 7:13:55 PM

(Currie, 1998) was chosen as a finalist for the 1999 Pulitzer Prize in nonfiction. Currie writes with the dexterity of a Pulitzer nonfiction nominee—yet with the authority of a criminological scholar. He is a renowned comparativist and is admired as much by criminologists in Great Britain and Ireland—perhaps even more—as he is by those in the United States. And he has drawn on international experience to critique the punitive crime and imprisonment policies of the United States.

The U.S. incarceration rate was low and relatively stable between 1920 and 1970 (Blumstein and Cohen, 1973). California's prisons, which operated under a system of indeterminate sentencing, maintained empty beds. But as protest and civil disorders were forcibly suppressed, crime rates began to rise, although not nearly as much as prison populations. These populations doubled in the 1970s and tripled in the 1980s as sentencing laws became increasingly punitive across the nation. In 1970, the federal government and most states—not California alone—maintained indeterminate sentencing systems in which legislatures formulated penalties, judges interpreted the penalties, and correctional and parole officers made recommendations to the parole boards who decided on release. These decision makers could be arbitrary and unfair, but on the whole, sentences were lighter and inmates—as well as the public—were relatively safe. Since then, sentences have been lengthened, especially (but not only) for violent offenders. State and federal prosecutors have become the main arbiters of the fate of defendants as three strikes laws and federal sentencing guidelines have come to dominate criminal punishment.

At both the federal and state levels, defendants overwhelmingly plead guilty before trial and bargain with prosecutors regarding the severity of sentences. Prosecutorial authority at the federal level has diminished since the U.S. Supreme Court decided *United States v. Booker* (2005), an opinion that restored to federal judges much of the discretion that Congress had taken when it introduced mandatory sentencing guidelines. Nevertheless, most judges follow the sentencing guidelines that they are obliged to consider before deviating. Moreover, most federal judges have been appointed since the passage of the guidelines and have become familiar with its complex schema.

Prison policy is often attributable to interest groups. Unionized correctional officers have become powerful legislative actors and are supported by local legislators who seek prison growth to stimulate the economy in their districts. Citizen groups are also influential. Most interest groups have supported harsh justice, whereas others have less successfully argued to reduce mandatory minimum sentences. California has become an ironic exception—prison population growth has begun to threaten the safety of correctional officers as well as of prisoners, and correctional officers themselves are seeking sentence reduction.

Although California maintains the largest prison population in the United States, it is by no means the only state confronting overcrowded prisons. After 40 years, from the protest era of 1968 to 2008, the United States has fashioned its own gulag archipelago. For the first time in United States history, more than 1 out of every 100 American adults are confined in a jail or prison—and the prisoners are disproportionately Black and male. One out of every

09012-CrimJournal-Guts.indd 30 12/21/09 7:13:56 PM

21 Black males is imprisoned, whereas the number for Whites is 1 out of 136 (Pew Center on the States, 2008).

The imprisonment binge of the United States will affect the capacity of states to address public sector needs for years to come, especially their ability to fund education. Indeed, five states—Vermont, Michigan, Oregon, Connecticut, and Delaware—recently have spent more on corrections than on education. This choice has not been explicit. But funds allocated to corrections cannot be drawn for other public needs, especially education (Pew Center on the States, 2008).

Some good news has been reported. Even as imprisonment rates rose for young Black males, opportunities increased for segments of the African American population. Affirmative action policies in colleges, universities, and major corporations brought African Americans into the mainstream. An intelligent, motivated African American girl from Chicago's South Side could follow her star-athlete brother to Princeton and move on to graduate cum laude from Harvard Law School. A poor Latina raised in Bonfire of the Vanities country (the south Bronx) could become a *Yale Law Journal* editor and the first Latina Supreme Court justice—a phenomenon that the historian Michael Katz labels "selective incorporation."

Selective incorporation has led to social stratification within the Black and Brown communities. The losers were those with only a high-school education, White, Brown, or Black (Western, 2006). In 1993, Elliot concluded that, "much greater numbers of people of all races, even if they successfully complete high school, realistically have very little to look forward to in the world of work" (Reckoning, 1993: 132). Although the relationship among crime, education, and employment is complex, there is scarce question that a relationship does exist. Legitimate employment and a steady income are a form of social control. The duties of work discipline one's daily life and encourage family responsibilities. By contrast, jobless—or school dropout—young men tend to be impetuous, to join gangs, and are more likely than other jobless men to be free of ordinary social constraints.

Too many of these young men turned to crime and drug dealing. Those at the bottom are unlikely to be deterred by the threat of imprisonment partly because they are not rational cost-benefit calculators and the threat is not always credible, but mostly because they have few alternatives in the legal economy. And when they are arrested, even if not convicted, their chances of obtaining legitimate work are diminished even more. The alternative is to develop an underground economy of illegitimate work—like selling drugs, stealing, and selling property or prostitution.

Currie's research reflects many of these themes. Moreover, he is a superb interviewer of troubled teenagers from every social background as in *Dope and Trouble: Portraits of Delinquent Youth* (Currie, 1991) and *The Road to Whatever: Middle Class Culture and the Crisis of Adolescence* (Currie, 2005). He is an insightful student of American race relations as he demonstrates in his co-authorship of *Whitewashing Race: The Myth of a Colorblind America* (Brown et al., 2003), which was a finalist for the C. Wright Mills Award from the Society for the Study of

09012-CrimJournal-Guts.indd 31 12/21/09 7:13:56 PM

Social Problems and winner of the 2004 Book Award from the Benjamin L. Hooks Institute for Social Change.

But as in his most recent book, *The Roots of Danger: Violent Crime in Global Perspective* (Currie and Pontell, 2008), or in his earlier work, *Crime and Punishment in America* (Currie, 1998), Currie also has been a theoretical internationalist and has used the experiences of other nations to review our own crime policies and practices. A nation's crime, in Currie's view, is not attributable to genetic defect or inherently twisted character but as a response to the social and economic conditions into which criminals are born. Poverty, inequality, and the absence of social supports correlate with crime across the world.

Elliot's most recent book, *The Roots of Danger: Violent Crime in Global Perspective* (Currie, 2009), reviews, summarizes, and updates his research and conclusions about the causes of violent crime that were introduced in his Pulitzer-nominated 1998 book. Where violence is high, Currie writes, social and political harshness prevails in many ways. Crime and violence become culturally acceptable when the economic and political policies of nations fail to provide a social safety net. The rise in crime erodes a sense of community, which in turn leads to higher crime.

Elliott's contributions to our knowledge of crime and its causes demonstrate that the brilliant graduate student has fulfilled his promise. He has become an internationally renowned writer, thinker, social critic, public intellectual, and superb criminologist well deserving of our Society's August Vollmer Award. Congratulations Elliott!

References

Blumstein, Alfred and Jacqueline Cohen. 1973. A theory of the stability of punishment. *Journal of Criminal Law and Criminology*, 64: 198–207.

Brown, Michael K., Martin Carnoy, Elliott Currie, Troy Duster, David B. Oppenheimer, Marjorie M. Schultz, and David Wellman. 2003. *Whitewashing Race: The Myth of a Colorblind America*. Berkeley: University of California Press.

Currie, Elliott. 1991. *Dope and Trouble: Portraits of Delinquent Youth*. New York: Pantheon Books.

Currie, Elliott. 1993. *Reckoning: Drugs, the Cities and the American Future.* New York: Hill and Wang.

Currie, Elliott. 1998. *Crime and Punishment in America*. New York: Holt Paperbacks.

Currie, Elliott. 2005. *The Road to Whatever: Middle Class Culture and the Crisis of Adolescence*. New York: Holt Paperbacks.

Currie, Elliott. 2009. *The Roots of Danger: Violent Crime in Global Perspective*. London, UK: Pearson.

Currie, Elliott and Henry N. Pontell. 2008. *The Roots of Danger: Violent Crime in Global Perspective*. Upper Saddle River, NJ: Prentice Hall.

The National Commission on the Causes and Prevention of Violence. 1969. *Politics of Protest.* New York: Simon & Schuster.

09012-CrimJournal-Guts.indd 32 12/21/09 7:13:56 PM

The Pew Center on the States. 2008. *One in 100: Behind Bars in America 2008*. Retrieved October 1, 2009 from pewcenteronthestates.org/uploadedFiles/8015PCTS_Prison08_FINAL_2-1-1_FORWEB.pdf.

Pager, Devah. 2003. The mark of a criminal record. *American Journal of Sociology*, 108: 937–975.

Skolnick, Jerome H. and Elliott Currie. 2006. *Crisis in American Institutions*, 13th edition. Upper Saddle River, NJ: Allyn & Bacon.

Western, Bruce. 2006. *Punishment and Inequality in America*. New York: Russell Sage Foundation.

Cases Cited

United States v. Booker, 543 U.S. 220 (2005).

Jerome H. Skolnick teaches at the New York University School of Law, is the codirector of its Center for Research in Crime and Justice, and is the Claire Clements Dean's Chair Emeritus at the University of California, Berkeley, School of Law. He has been honored for his scholarship by virtually every major criminal law and justice organization for his distinctive sociological work on crime and criminal justice administration.

09012-CrimJournal-Guts.indd 33 12/21/09 7:13:56 PM

EDITORIAL INTRODUCTION

ALCOHOL POLICY IMPLEMENTATION EVALUATION

Evaluation of the Licensing Act of 2003

A look inside the black box

Andrew Treno, *Senior Editor*
Prevention Research Center

In late autumn of 2005, the Licensing Act of 2003 (henceforth, the Act) was implemented throughout England and Wales with the "intentions of reducing crime and disorder, enhancing public safety, preventing public nuisance, and protecting children from harm" (Humphreys and Eisner, 2010, this issue). The Act essentially "removed restrictions on trading times and encouraged licensed premises (should they wish) to apply for extended hours of trade" (Humphreys and Eisner, 2010). The decisions to grant such extensions were relegated to local licensing authorities. In light of evidence linking increased alcohol availability to increased drinking and alcohol-related crime and disorder, the efforts to decrease such problems through increased alcohol availability and increased hours of sales might seem a bit curious. As indicated by earlier research, increasing hours for serving alcohol generally has been associated with increased violence, not less (Chikritzhs and Stockwell, 2002; D'Abbs, 1994; Duailibi, Ponicki, Grube, Pinsky, Laranjeiro, and Raw, 2007; Haines and Graham, 2005; Norstrom and Ole-Jorgen, 2005; Ragnarsdottir, Kjartansdottir, and Davidsdottir, 2002). In fact, the broader alcohol-availability literature has linked increases in the physical, social, legal, and geographic availability to increases in violence and assaultive injuries (Gruenewald, Freisthler, Remer, LaScala, Treno, and Ponicki, in press; Gruenewald, Remer, and Treno, 2009; Holder et al., 2000; Lipton and Gruenewald, 2002; Maguire and Nettleton, 2003; Norstrom, 2000; Treno, Gruenewald, Lee, and Remer, 2007; Zhu, Gorman, and Horel, 2004). The logic behind the new law, however, involved the consideration of alcohol-related violence as not simply the result of drinking but as the result of the interaction between drinking and "social mixing" that occurs when large numbers of establishments close, which concentrates large numbers of individuals in small areas. As described by Humphreys and Eisner (2010), the Act was to serve as a trigger by removing fixed closing times and installing staggered closing times, which would lead to an

Direct correspondence to Andrew Treno, Prevention Research Center, 1995 University Avenue, Suite 450 Berkeley, CA 94704 (e-mail: Andrew@PREV.org).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 35 12/21/09 7:13:56 PM

easing of crowding and stress on the urban infrastructure with resulting reductions in violent and disorderly interactions.

As noted by Humphreys and Eisner (2010), although this intervention has been evaluated by several studies, none have focused on whether the Act changed retail practice. Thus, evaluation of the Act has adopted a "black box" approach that considers the overall effects of the intervention. In summarizing these studies, Humphreys and Eisner noted that evaluations have produced inconsistent results. Some found postintervention decreases, whereas others found postinterventions increases in assaults and other violent events. However, as they note, because an examination of actual closing-time changes is absent, any attribution of these changes to the Act remains questionable.

Humphreys and Eisner's (2010) article is the first attempt to look inside the black box and focus on the one aspect of the logic model underlying the intervention—the link between the implementation of the Act and subsequent changes in closing times. The article, thus, poses three questions. First, did the Act increase the availability of alcohol? Second, did it lead to staggering closing times as intended? Third, did changes in closing times differ between regions and neighborhoods? This last question seems particularly relevant because most evaluations of the Act have not addressed the issue of an appropriate unit of analysis—an important issue to consider where a local variation of dosage occurs. Thus, Humphreys and Eisner conclude (a) that licensing practices did change the availability of alcohol, although this change varied from place to place, (b) that staggering did occur to some extent at the city level, and (c) within smaller units of analysis, peaks in closing times remained even with staggering.

The two companion policy essays that accompany the Humphreys and Eisner (2010) article have focused on different aspects of the issue addressed. Frattaroli's (2010, this issue) essay "Why implementation matters: Recent experience with the U.K. Licensing Act" correctly notes that evaluating policy impact "is predicated on some assurance that the policy was implemented with some fidelity to the policy goal." The essay also describes the haphazard decentralized implementation motivated largely by broader political concerns and market forces. Although the Humphreys and Eisner article attempts to document the staggering phenomenon, the Frattaroli essay highlights other aspects of the "black box" and noted that a consideration of more "qualitative" data consisting of "interviews with government officials, residents, and trade professionals [which] have added the perspectives and experiences of the people affected by the change in policy." Three lessons are drawn from this point. The first is that varied methods, which include the use of qualitative data, can inform our consideration of implementation questions. The second is that those factors influencing implementation should be considered. The final lesson pertains to revisiting the Act in light of what is now known about the broader implementation process.

Hadfield and Measham (2010, this issue) approach the Act in terms of social and political developments that occurred during the 1990s in England and Wales—specifically the culture of intoxication that emerged during the period, particularly among youth and young adults. Hadfield and Measham interpret the Act as a result of a desire to bring about cultural change

09012-CrimJournal-Guts.indd 36 12/21/09 7:13:56 PM

through the implementation of one piece of policy legislation. Moreover, it addressed the significance of the licensing placement in the Department of Culture, Media, and Sport—with no particular focus on issues of public safety and health—as opposed to the Home Office and Ministry of Health. In this broader context, both incompetence and interests contrary to the public health might have held sway. The implication of this issue for the science is clear. If local conditions determine implementation, then one might find less staggering of closing times at the local than the aggregate level. The essay concludes by noting that the major contribution of the Humphreys and Eisner (2010) article is to "remind us that minutiae matters; too often, previous research in this field has involved superficial retrospection, or simple 'before and after' comparisons of aggregate data that are not nuanced sufficiently to provide sound bases for public policy."

Across the original Humphreys and Eisner (2010) article as well as Frattaroli (2010) and Hadfield and Measham (2010), the following points are made:

1. At the most general level, implementation matters. Program implementation levels might vary based on either local political circumstances or the vagaries of free-market operations.
2. Local variation in implementation levels might mask intervention effects when measured in aggregate. Thus, local effects must be considered relative to local, and not overall, levels of implementation.
3. Local political and market forces might intervene to undermine the intent of policy interventions.
4. Ultimately, black-box approaches, which characterize many program intervention evaluation efforts, must be replaced by evaluations that incorporate local variations and all aspects of the logic model underlying the intervention.

The implications of this collection for the field of preventive interventions might be read narrowly in the context of extended hours of sale, more broadly in terms of the study of the impact of increased alcohol availability, or more broadly still in terms of the development of a gold standard for the evaluation of preventive interventions. At the broadest level, the attribution of effects to any given intervention requires that several conditions be fulfilled. These include (a) a logic model developed based on prior research findings; (b) a mapping of indices based on model constructs that include process, intermediary, and outcomes measures; (c) a collection of preintervention data across all model constructs; (d) a collection of postintervention data across all model constructs; (e) statistical controls for potential confounds; (f) a determination of the appropriate scale at which model constructs are to be considered; and (g) data-analysis strategies and statistical procedures appropriate for the unique characteristics of the distributional properties of the data elements. Unfortunately, such a gold standard is seldom attained. Most certainly, it has not been attained in extent evaluations of the Act. Fortunately, articles such as those presented here move us in the desired direction.

09012-CrimJournal-Guts.indd 37 12/21/09 7:13:57 PM

References

Chikritzhs, Tanya and Tim Stockwell. 2002. Impact of later trading hours for Australian public houses (hotels) on levels of violence. Journal of Studies on Alcohol, 63: 591–599.

D'Abbs, Peter. 1994. *Darwin Nightclubs: A Review of Trading Hours and Related Issues*. Casuarina, Australia: Menzies School of Health Research.

Duailibi, Sergio, William Ponicki, Joel Grube, Ilana Pinsky, Ronaldo Laranjeiro, and Martin Raw. 2007. The effect of restricting opening hours on alcohol-related violence. *American Journal of Public Health*, 97: 2276–2280.

Frattaroli, Shannon. 2010. Why implementation matters: Recent experience with the U.K. Licensing Act. *Criminology & Public Policy*. This issue.

Gruenewald, Paul J., Bridget Freisthler, Lillian G. Remer, Elisabeth A. LaScala, Andrew J. Treno, and William R. Ponicki. 2010. Ecological associations of alcohol outlets with underage and young adult injuries. *Alcoholism: Clinical and Experimental Research*. In press.

Gruenewald, Paul J., Lillian G. Remer, and Andrew J. Treno. 2009. Alcohol outlets, crime, and disorder in the United States of America. In (Phil Hadfield, ed.), *Nightlife and Crime: Social Order and Governance in International Perspective*. Oxford, UK: Oxford University Press.

Hadfield, Phil and Fiona Measham. 2010. After the Act: Alcohol licensing and the administrative governance of crime. *Criminology & Public Policy.* This issue.

Haines, Ben and Kathryn Graham. 2005. Violence prevention in licensed premises. In (Tim Stockwell, Paul J. Gruenewald, John W. Toumbourou, and Wendy Loxley, eds.), *Preventing Harmful Substance Use: The Evidence Base for Policy and Practice*. New York: Wiley.

Holder, Harold D., Paul J. Gruenewald, William R. Ponicki, Andrew J. Treno, Joel W. Grube, Robert F. Saltz, et al. 2000. Effect of community-based interventions on high risk drinking and alcohol-related injuries. *Journal of the American Medical Association*, 284: 2341–2347.

Humphreys, David K. and Manuel P. Eisner. 2010. Evaluating a natural experiment in alcohol policy: The Licensing Act (2003) and the requirement for attention to implementation. *Criminology & Public Policy*. This issue.

Lipton, Robert and Paul J. Gruenewald. 2002. The spatial dynamics of violence and alcohol outlets. *Journal of Studies on Alcohol*, 63: 187–195.

Maguire, Mike and Hilary Nettleton. 2003. *Reducing Alcohol-Related Violence and Disorder: An Evaluation of the TASC Project*. London Home Office.

Norström, Thor. 2000. Outlet density and criminal violence in Norway, 1960–1995. *Journal of Studies on Alcohol and Drugs*, 61: 907–911.

Norström, Thor and Skog Ole-Jorgen. 2005. Saturday opening of alcohol retail shops in Sweden: An experiment in two phases. *Addiction*, 100: 767–776.

09012-CrimJournal-Guts.indd 38 12/21/09 7:13:57 PM

Ragnarsdottir, Porgerour, Asgerour Kjartansdottir, and Sigurlina Davidsdottir. 2002. Effect of extended alcohol serving hours in Reykyavik, Iceland. In (Robin Room, ed.), *Effect of Nordic Alcohol Policies*: *What Happens to Drinking and Harm When Alcohol Controls Change?* Helsinki, Finland: Nordic Council for Alcohol and Drug Research.

Treno, Andrew J., Paul J. Gruenewald, Juliet P. Lee, and Lillian G. Remer. 2007. The Sacramento Neighborhood Alcohol Prevention Project: Outcomes from a community prevention trial. *Journal of Studies on Alcohol and Drugs*, 68: 197–207.

Zhu, Li, Dennis M. Gorman, and Scott Horel. 2004. Alcohol outlet density and violence: A geospatial analysis. *Alcohol and Alcoholism*, 39: 369–375.

Andrew Treno is a research scientist at the Prevention Research Center in Berkeley, California, a National Alcohol Research Center. He has worked in the area of environmental prevention for the past 20 years. During that period he has worked on two major projects, the Community Trials Project (Harold D. Holder, Principal Investigator) and the Sacramento Neighborhood Alcohol Prevention Project, a project designed to reduce alcohol access, drinking, and related problems in two low income largely minority neighborhoods in Sacramento, California (Paul J. Gruenewald, Principal Investigator). As a member of the research team on the Community Trials Project he assumed responsibility for the evaluation of community mobilization and media advocacy and developed a surrogate measure for alcohol involved injury. He currently serves as project director on the SNAPP project and maintains management, budgetary, and scientific responsibility for the daily conduct of the project. Additionally, he has served as Project Director on a project funded by the National Institutes of Health investigating the effects of Alcohol Advertising on Youth (Joel Grube, Principal Investigator), which has involved conducting both focus groups and self-administered questionnaires in school settings and has published using those data. Treno received his Ph.D. in Sociology from the University of California at Berkeley in 1983.

09012-CrimJournal-Guts.indd 39 12/21/09 7:13:57 PM

RESEARCH ARTICLE

ALCOHOL POLICY
IMPLEMENTATION EVALUATION

Evaluating a natural experiment in alcohol policy

The Licensing Act (2003) and the requirement for attention to implementation

David K. Humphreys
Manuel P. Eisner
University of Cambridge

Research Summary

This study examines the effects of a dramatic change in alcohol licensing policy on the practice of liquor retailing across a metropolitan urban region. The policy was enacted to remove fixed closing times and to install staggered closing times across regions where it potentially would reduce violent crime and disorder. We found great variation in the extent to which the provisions of the policies were installed as intended.

Policy Implications

The findings suggest that prior consideration for how policy interventions will be applied to real-world settings is important to the validity of research evaluations. Understanding and monitoring variation in the implementation of an intervention is a vital prerequisite to the evaluation of outcomes. Without knowledge of the real-world application of an intervention, researchers risk attributing causal effects to a prevention initiative that might have been absent or only partially received.

Keywords

alcohol regulation, implementation, crime prevention

The authors would like to thank Professors David P. Farrington and Lawrence W. Sherman for helpful comments on an earlier draft as well as Caroline Moul for technical assistance. Direct correspondence to David K. Humphreys, Institute of Criminology, University of Cambridge, Sidgwick Avenue, CB3 9DA Cambridge, UK (e-mail: dkh25@cam.ac.uk); Manuel Eisner, Institute of Criminology, University of Cambridge, Sidgwick Avenue, CB3 9DA Cambridge, UK (e-mail: mpe23@cam.ac.uk).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 41 12/21/09 7:13:57 PM

Over the last twenty years, concern for the prevalence of alcohol-related crime and disorder has reached fever pitch in the United Kingdom. Alcohol-related crime costs the U.K. government £12 billion ($19.5 billion) annually. The British Crime Survey (BCS, 2007–2008) reported that 58% of victims of violent crime perceived their attackers to have been under the influence of alcohol, which has prompted several political responses to problematic alcohol consumption. In late autumn 2005, the Licensing Act (2003) was enacted with the intentions of reducing crime and disorder, enhancing public safety, preventing public nuisance, and protecting children from harm. The Licensing Act is the largest liberalization of alcohol regulations in England and Wales since the beginning of the 19th century. Contrary to the traditional emphasis on regulatory control, the Licensing Act (referred to hereafter as "the Act") is a rare example of a measure that tries to achieve crime reduction by *removing restraints* rather than by adding them.

Since the Act's inception, several academic- and government-funded research evaluations have been published (Babb, 2007; Bellis, Anderson, and Hughes, 2006; Hough and Hunter, 2008; Newton, Jala-Sarker, Pahal, Van den Bergh, and Young, 2007; Newton, Hirschfield, Armitage, Rogerson, Monchuk, and Wilcox, 2008; Pike, O'Shea, and Lovbakke, 2008) that attempt to assess the impact of the Act on crime and disorder. However, they have struggled to arrive at firm conclusions about the Act's impact. These studies have largely adopted black-box before-and-after designs that examine patterns of criminal event data before and after the regulation change. Unfortunately, they omit important discussions of how or whether the Act affected retail practice; without which, questions of effectiveness as well as of impact cannot be answered purely by observing fluctuations in crime outcomes.

Much attention has been placed in recent years on establishing "what works" in crime prevention (Sherman, Gottfredson, MacKenzie, Eck, Reuter, and Bushway, 1997; Welsh and Farrington, 2001, 2007). However, despite emphasis on the success of specific interventions for specific settings, rarely has attention been paid to understanding the mechanics of *how* such interventions work (Domitrovich and Greenberg, 2000; Durlak and Wells, 1997; Michalic and Irwin, 2003). This is not to say that "monitoring" prevention programs does not occur—it does. Many examples are available in which "what was done" is taken into consideration in evaluations and systematic reviews (see Farrington and Welsh, 2007; Welsh and Farrington, 2007). However, investigating the impact of "what was done" on program outcomes is an often ignored question (Dumas, Lynch, Laughlin, Phillips Smith, and Prinz, 2001; Durlak and Wells, 1997). As discussed in Lipsey (1993), breaking away from conventional black-box before-and-after outcome evaluations can help construct more sophisticated models of what works, in what forms, and for different settings. As Eisner (2009) found, this information becomes important when attempting to understand why replications of evidence-based interventions fail to show similar effects in different social contexts.

In cases where programs produce null effects, it is common to direct questions at implementation variation (Lipsey, 1993; Petrosino and Soydan, 2005) in which concepts, such as

09012-CrimJournal-Guts.indd 42 12/21/09 7:13:57 PM

implementation *fidelity*, *integrity*, or *adherence*, have become an important consideration for program evaluators (Dumas et al., 2001). Despite the growing emphasis on these terms in program evaluation, little attention has traditionally been focused on *how* public policy interventions work, whether their implementation fulfills their desired aims, whether variation occurs within the application of policies, and whether such variations can impact differently on the relevant outcomes.

The purpose of this article is to investigate the application of a natural experiment in crime prevention policy.[1] Although we acknowledge that issues of implementation cannot be applied to public policy interventions in the same way they are applied to program evaluations, we share Lipsey's (1993) view that understanding mediating processes between inputs and outputs benefits the transfer of knowledge between practice and research. This step is important in the process *and* development of public policy and is central to the way researchers evaluate interventions prior to analysis of the outcomes. Researchers *and* policy makers are required to understand *why* desired outcomes occur, not just *whether* they occur.

The Licensing Act (2003)

The Licensing Act was implemented throughout England and Wales on November 24th. It removed restrictions on trading times and encouraged licensed premises (should they wish) to apply for extended hours of trade. The Act is a multidimensional regulatory intervention aimed at achieving improvements in social order throughout the nighttime economy but also has larger goals of instilling a "modernized" nighttime environment and a safer drinking culture equivalent to central European attitudes to alcohol. The Act uses several initiatives, which (through various mechanisms familiar to criminological theory) aim to achieve the aforementioned statutory goals (see Table 1). However, this discussion will be focused on the situational prevention method of the deregulation of fixed trading hours. [2]

1. Opinions vary somewhat of what does and does not constitute a "natural experiment." This article understands a "natural experiment" to be an experiment in which an intervention is catalysed by a naturally occurring event, like an earthquake or assassination, where the manipulation is outside of the control of researchers (Shadish, Cook, and Campbell, 2002; Tabachnick and Fidell, 2006). The implementation of the Act occurred outside the control of researchers, which represents an abrupt and significant change to a regulatory system. We believe this change can be used to make causal inferences, and therefore, it is acceptable to refer to this legislative change as a "natural experiment."

2. For more discussion about other facets of the Act, readers should consult Hadfield (2006) and Hadfield and Measham (2009).

09012-CrimJournal-Guts.indd 43 12/21/09 7:13:58 PM

TABLE 1

Summary of New Initiatives under the Licensing Act (2003)

Initiative	Purpose	Proposed Effect
A. Unified system of licensing	Procedural	To reduce the burden of unnecessary regulation
B. Introduction of personal licenses as well as place licenses	Procedural/ accountability	To increase individual accountability for premises and throughout the industry
C. Switching licensing governance from magistrate courts to local authorities	Procedural/local governance and partnership empowerment	To increase the participation of key local bodies in licensing decision making
D. Deregulation of fixed closing times	Situational prevention	Removing fixed closing times to cause staggering and reduce violent and disorderly flashpoints
E. Additional police powers to close disorderly premises	Deterrence	To encourage premises to manage orderly and crime-free environments

Rationale

The primary feature of the new legislation was the removal of the fixed closing times that previously limited the sale of alcohol on licensed premises. Under the previous legislation, licensed premises were allowed to sell alcohol between the hours of 11 a.m. and 11 p.m. (10:30 p.m. on Sundays) with 20 minutes of "drinking up" time.[3] In addition, a smaller proportion of premises also held "Special Hours Certificates," which could allow for the service of alcohol until 2 a.m.[4] The new legislation introduced much greater flexibility in licensing, which was believed to be more suitable for alcohol retailers and the modern 24-hr urban consumer (Hadfield, 2006).

Three studies (Hope, 1986; Marsh and Fox-Kibby, 1992; Tuck, 1989) were central in identifying the role that licensing could play in the prevention of violent crime and disorder. In a study based in a city in Northeast England, Hope (1986) found that disorder was concentrated in both space and time. He found that 25% of all disorder that occurred between the hours of 6 p.m. and 6 a.m. was concentrated in the 30 minutes after the first peak in closing time. This finding was supported in a Home Office study compiled by Tuck (1989) that spanned several urban, suburban, and rural sites. Tuck, like Hope, hypothesized that problems with violence and disorder were triggered as individuals left bars and nightclubs, as opposed to occurring within venues. According to Tuck, after leaving licensed premises, intoxicated individuals congregate at take-away food outlets and around public transport nodes. Furthermore, a study by Marsh and Fox-Kibby (1992) also noted a concentration of violent crime and disorder around both the initial peak in closing time (11 p.m.) and the secondary peak (2 a.m.). They found that

3. The hours included sale of alcohol for on- and off-premises consumption.

4. For space considerations, we will not discuss the nature of other licensing permutations; readers should see Hadfield (2006) and Raistrick, Hodgson, and Ritson (1999) for a more detailed summary of previous licensing practice.

09012-CrimJournal-Guts.indd 44 12/21/09 7:13:58 PM

47% of incidents of disorder related to alcohol consumption occurred within only 2 hrs during the week—between 11 p.m. and 12 a.m. on Friday and Saturday.

From these observations, these researchers proposed that the temporal concentration of closings times could be an important factor that fuels acts of violent crime and disorder. They argued that individuals become involved in fractious and sometimes violent interactions after acute alcoholic intoxication and prompt ejection from licensed premises at closing time. Furthermore, competition for scarce resources (food, more alcohol, transport, or sexual intercourse) within confined urban environments was a contributing factor in problematic interactions. These authors suggested that the current framework of licensing caused *unintended negative consequences* by putting undue restraints on closing hours. They argued that the removal of fixed hours of trade could serve to stagger closing times and ease the burden placed on the local environment and infrastructure.

Mapping the Causal Chain

A primary concern for those researchers evaluating the Act is that legal change is assumed to trigger a sequence of causal factors central to the achievement of reduced violent and disorderly acts. As the causal model in Figure 1 demonstrates, the Act has enabled premises to change operating hours to install staggered closing times (inputs) to result in less crowding and stress on the urban infrastructure (the mechanism)[5] with the result of reduced incidents of violent and disorderly crime (the outcome). However, the provisions of the Act do not guarantee that the essential inputs (staggered closing times) will occur (Guidance, 2006). Provisions such as license extensions and staggered closing times were not planned formally but instead were left to market forces to "naturally" organize.[6] Therefore, to evaluate the impact of the Act, it is important to know the extent of the changes to practice and how variations impact outcomes. Without understanding what the Act's inputs *were*, it is impossible to say, with any confidence, whether the Act has had a positive or negative impact.

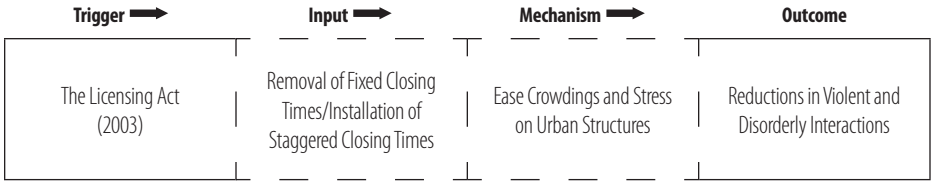
5. Despite advances in scientific technology that enable the measurement and analysis of patterns in crowd behavior (see Moore, Flajšlik, Rosin, and Marshall, 2008), this study will be unable to make valid post hoc inferences about the Act's ability to alter crowding and stagger street-level interactions.

6. Licensing authorities can lodge a motion not to grant further extended licenses with a Cumulative Impact Policy. This might occur when it is believed (and where evidence exists) that the cumulative impact of licensed premises in a geographic area is detrimental to the Act's statutory objectives. However, as Hadfield and Measham (2009) correctly observe, this motion only can act as one of intention. Licensing authorities cannot act alone in limiting licensing extensions; they must be supported by representations from an "interested party." In the absence of such representations, licenses and extensions must be granted.

09012-CrimJournal-Guts.indd 45 12/21/09 7:13:58 PM

FIGURE 1

The Act's Pathway to Prevention



Existing Studies

Since the introduction of the Act, several evaluations have been published that seek to examine the effects of the Act on crime and disorder (see Appendix A). These evaluations consist of publicly and privately funded reports; at present, only two peer-reviewed evaluations have been published (Hough and Hunter, 2008; Newton et al., 2007).[7] Hough and Hunter's study differs somewhat from the other evaluations because it summarizes the empirical findings from several other impact evaluations. However, the authors were able to conduct a series of interviews with licensees (premises managers) stressing the importance of the *implementation* of licensed extensions in hours of trade. However, the authors did not relate these findings to the analysis of crime outcomes to examine the potential impact of differential implementation.

Findings from other evaluations have been muted because of problems with accurately examining the practical changes in licensing after the Act's implementation. Three of the six studies shown in Appendix A (Babb, 2007; Bellis et al., 2006; Pike et al., 2008) found modest decreases in recorded violent crime after the implementation of the Act. In contrast, three studies observed increases in violent crime (Humphreys, 2007; Newton et al., 2007, 2008). However, the ability of these studies to identify the specific *effect* of the Act is questionable. These studies largely represent analyses of crime patterns before and after the introduction of the Act, but they do not assess the more detailed changes in licensing practices or whether these changes were related to crime outcomes.

Limitations of Previous Studies

In discussions of evaluation methodology for policy interventions, various scholars (Chen and Rossi, 1983; Horney and Spohn, 1990; Lipsey, 1993) have argued that researchers should be mindful of the nature of interventions in the construction of appropriate research designs. Notably, that attention should be paid to the target(s) of an intervention, the aims of the intervention, and the unit at which the intervention should be measured. An important limitation of existing evaluations of the Act is the failure to examine variations in its central mechanism (i.e., staggered hours of closing). In addition, crucial design limitations, such as the misspeci-

7. This study is not included in the summarized table in Appendix A because it draws largely on the same empirical findings of a related evaluation (Newton et al., 2008) that is listed in the same table.

09012-CrimJournal-Guts.indd 46 12/21/09 7:13:58 PM

fication of units of analysis exist in most of the aforementioned evaluations. These limitations make it difficult to establish, with any degree of confidence, whether staggering had any effect on crime and disorder.

Intervention Targets and Units of Analysis

Analyses of crime prevention interventions use a variety of different units of analysis such as individuals, networks, schools, and communities (Weisburd, Bruinsma, and Bernasco, 2008). For examination of sociolegal prevention policies, larger units often are used by researchers (such as national, state, or regional jurisdictions)[8] and comparative jurisdictions are used by researchers as natural controls. However, for many situational interventions such as closed circuit television (CCTV) and improved street lighting, experimental and quasi-experimental evaluation designs often use smaller units of analysis to disentangle treatment and comparison areas (Farrington, Bennett, and Welsh, 2007). Analysis at smaller units is required because the crime prevention initiatives (CCTV cameras or street lighting) might vary across geographical areas or political jurisdictions.

The provisions of the Act differ substantially from traditional alcohol policy because of the focus on intervening in the situational factors involved in crime and disorder, as opposed to previous controls on consumption. There is some disagreement concerning the extent to which the Act has brought about change in either licensing practices or levels of crime. For instance, variation might have occurred in the application of the new legislation between areas. Therefore, the rate at which premises applied for extended trading and the prevalence with which closing became staggered is unlikely to be fixed in either magnitude (lengths of extensions or extent of staggering) or geographic coverage (Rossi, Freeman, and Lipsey, 1999). Consequently, analysis and evaluation of effects at larger units of analysis (i.e., cities) can blur variation in the treatment variable, which obscures possible treatment effects.

As identified in other discussions (Hobbs, Hadfield, Lister, and Winlow, 2003; Shepherd and Brickley, 1996), nighttime economies and drinking circuits vary in concentration, clientele, atmosphere, and problems with crime throughout urban areas. The rationale of the Act suggests that fixed closing times lead to complications with overcrowding and overburdening local environments (Homel and Clark, 1994). As is noted in a later discussion (referring to Figure 2), concentrations of licensed premises are different in many aspects: they might vary spatially, they might be spread across smaller or larger geographic areas, the type of clientele that is attracted to any area might vary, and socioeconomic factors of the immediate surrounding environment might be vastly different between areas. These factors are important to understand the effect of any intervention targeted at reducing violence and disorder in licensed premises. Therefore, the Act's impact must be examined at smaller units, such as bar concentrations.

Until now, studies of the Act have neglected this aspect and have measured effects on larger units of analysis such as national crime patterns (Babb, 2007), local authority units (Bellis et

8. For examples of studies of sociolegal interventions at larger units of analysis, see Campbell and Ross (1968), Horney and Spohn (1989), and Norstöm and Skog (2003, 2005).

09012-CrimJournal-Guts.indd 47 12/21/09 7:13:58 PM

al., 2006), or central business districts (Humphreys, 2007; Newton et al., 2007; Pike et al., 2008). In a study by Newton et al. (2008), outcomes were measured across five study sites of varying size. Within this study, attention *was* paid to clustering of licensed premises. However, no analysis of staggering by cluster area took place, and therefore, the dose–response relationship was omitted.

Process of Implementation

Policy interventions, such as increased sanctions for drunk-driving offenses (Campbell and Ross, 1968) or changes to rape laws (Horney and Spohn, 1989), often are fixed in the nature in which they are applied and in the coverage of their application (Lipsey, 1993; Rossi et al., 1999). To use epidemiological terminology, these interventions are equivalent to the entire population receiving the same treatment in the same dose. This pattern makes the analysis of effects somewhat more straightforward; the most important design aspect is the point at which the systemic changes were made, such as change to a law or procedure. At this point, a researcher can compare the patterns of behavioral trends before the sanction against patterns after the sanction to assess the extent of changes in outcomes. This analysis can be performed in several rigorous methodological designs that will not be discussed here (see Cook and Campbell, 1979; Shadish, Cook, and Campbell, 2002).

However, in other areas of prevention research (see Lipsey, 1993; Bernfeld, Farrington, and Leschfield, 2001; Dobson and Cook, 1980; Dumas et al., 2001; Sanchez et al., 2007), the onset of a policy does not necessarily equate to the effective implementation of evidence-based prevention initiatives. Several reviews of implementation integrity (Dane and Schneider, 1998; Domitrovich and Greenberg, 2000; Dusenbury, Brannigan, Falco, and Hansen, 2003; Moncher and Prinz, 1991) have identified important issues in the maintenance of integrity in prevention programs, such as adherence to program goals, dose, *or* amount of treatment received, the quality of program delivery, and the participant responsiveness and program differentiation. In several meta-analyses, researchers found that attention to these issues was varied and in many cases rare (Dane and Schneider, 1998; Domitrovich and Greenberg, 2000; Michalic and Irwin, 2003; Moncher and Prinz, 1991; Peterson, Homer, and Wonderlich, 1982).

This variation underlines a critical issue in the evaluation of effects of the Act. Although the prevention initiatives like extended trading and staggered closing times were not *implemented* as such, but were triggered by deregulation, assessment of local reactions in licensed trading are essential for understanding changes in patterns of violence and disorder. None of the previous evaluation studies of the Act have effectively examined the degree of change in licensing that occurred after the change of law. This article will argue that the following series of important questions should be considered prior to analysis of outcomes: Did the implementation of the Act succeed in increasing the availability of alcohol? Did the implementation of the Act successfully lead to the staggering of closing times in licensed premises? Did the extent to which these protocols were achieved vary between regions or neighborhoods? The following sections describe the data and the method used to systematically observe the integrity of the legal change.

09012-CrimJournal-Guts.indd 48 12/21/09 7:13:59 PM

Methods

Data and Variables

The data used in this study were collected from numerous sources of local government in a large metropolitan city in the North of England. This study used a sample of 611 on-license public houses (which includes nightclubs and discotheques),[9] and this figure included 526 premises that were open throughout the 4-year study period as well as 85 premises that opened after the implementation of the Act. The data collected were extracted from two sources. *Pre-implementation licensing data* comprised opening and closing times for licensed premises for a 2-year period prior to the Act's implementation. This information was collected from manual registers held within the courts of local magistrates. Variables computed from these data included venue name, street address, seven opening and closing hour variables,[10] daily durations (times 7), and weekly duration. The *post-implementation licensing data* comprised opening and closing times for a 2-year period after the Act's implementation. These data were compiled from a database held by the newly formed Licensing Authorities, which under the new legislation were to be located within local authorities. The variables included were venue name, address, XY geocode, date of license conversion,[11] date of license variation,[12] date of secondary variation (if present), licensed alcohol service (times 7 [opening]), licensed alcohol service (times 7 [closing]), daily duration (times 7), and weekly duration. Obtaining data for licensed opening hours under both the old and the new licensing protocols would allow researchers to examine the extent and magnitude of licensing change (the input, shown in Figure 1).[13]

9. This figure is compiled of 611 premises that were clearly identifiable as "on-license" premises from the local authority licensing records. Since the administrative changes took place, the distinction between "on" and "off" licensed premises was removed from the recording standards of the local authority. For more information, a Department of Culture, Media and Sport (DCMS) statistical bulletin (2004) published national and local counts of licensed premises. Prior to the change in licensing, these findings showed 911 on-licensed premises.

10. Seven opening and closing hour variables were computed because times of closure could vary across the weekly period.

11. A license conversion is a procedural requirement that demands all premises to convert their existing license to be compatible with procedural requirements under the new legislation. This conversion requires the completion of updated paperwork and a small fee for each premise.

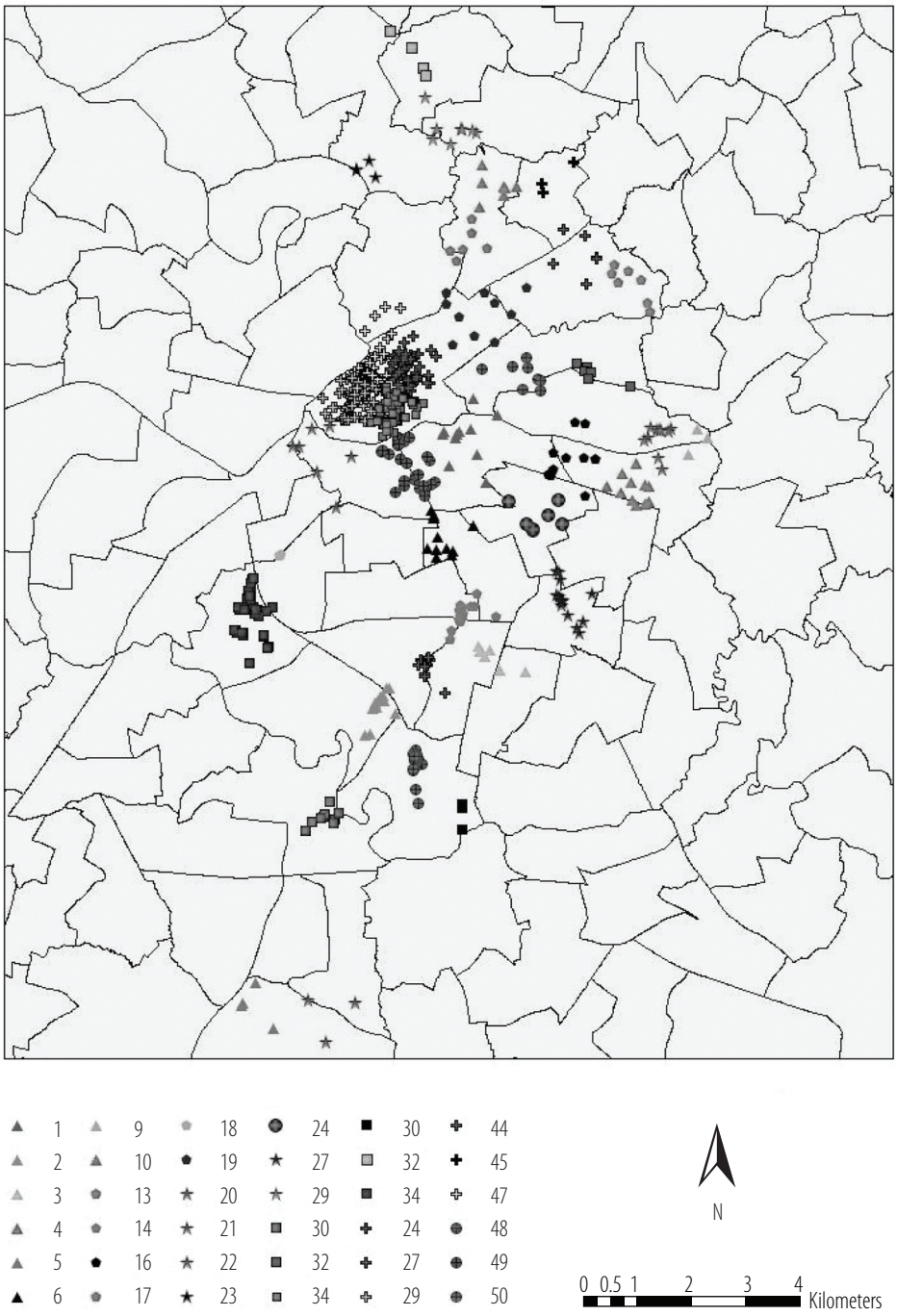
12. A license variation indicates a change to a licensable activity at the request of a premises' licensee. However, it should not be assumed that such variation is one of extended hours; it could, for example, be for any one of nine licensable activities, such as a performance of live music, playing recorded music, dance performance, live performances, facilities for making music, dancing facilities, entertainment facilities for making of music or dancing, the provision of late-night refreshment, or the sale of alcohol. Vigilance was required during data collection to count only licensed opening and closing times and the licensed sale of alcohol.

13. One limitation important to draw attention to here is that the Act does not enforce licensed trading times. These trading times represent the latest times that premises are allowed to trade. However, it would be possible, on some occasions, for premises to close before their licensed trading time, thus, creating some potential inaccuracy within the data. However, discussions with licensing practitioners suggested that this problem was likely to be infrequent and limited to a small proportion of premises sampled. Despite this potential limitation, the data collected provide the most representative picture of licensing for a post hoc research design.

09012-CrimJournal-Guts.indd 49 12/21/09 7:13:59 PM

FIGURE 2

Map of Premises Clusers



09012-CrimJournal-Guts.indd 50 12/21/09 7:14:00 PM

Units of Analysis and Measurement

The first step in analyzing the data was to determine units of analysis adequate for examining local effects of the Act on the staggering of closing times. To do this we first used the nearest neighbor index (NNI) in CrimeStat Version III to assess whether the geographic points were distributed evenly on the basis of chance or whether geographic clustering could be observed (Bailey and Gattrell, 1995; Ned Levine & Associates, 2005). These analyses indicated that licensed premises were clustered highly under both the previous and the new legislative protocols.[14] To identify the clustered patterns, we experimented with several spatial analytical tools.[15] When other methods proved unsatisfactory, a *k*-means clustering technique (SPSS Version 15; SPSS Corporation, Chicago, IL) was applied. The *k*-means algorithm uses an iterative process to assign data points (in this case, licensed premises) to clusters that minimizes within-cluster squared error (SSE; Fortin and Dale, 2005). A problem with this clustering technique is that it requires some subjectivity in the decision about the *k* number of clusters to identify (Bailey and Gattrell, 1995; Fortin and Dale, 2005; Legendre and Legendre, 1998; Ned Levine & Associates, 2005). However, as Ned Levine & Associates (2005) suggested, the use of *k*-means clustering often is useful when the researcher wants to control partitioning. With a substantial qualitative knowledge of the city in question, we were interested in identifying, as accurately as possible, clusters that represented known drinking circuits within the city. As Figure 2 demonstrates, the use of *k*-means clustering was an adequate tool in this instance.

Using this method, we ran a search for 48 clusters. A large number of clusters were chosen in order to identify isolated premises. When presented using GIS software (ArcMap Version 9.1; ERSI, Redlands, CA), we identified 12 clusters (comprising 18 premises) that contained less than 3 premises. For the purposes of these analyses, we decided that clusters should consist of at least 3 premises; therefore, we chose to exclude isolated premises at the later stages.[16] The remaining 36 clusters of licensed premises ranged in size with an average of 16.5 premises (maximum = 88 premises; minimum = 3 premises). As Rengert and Lockwood (2009) suggested, traditional administrative boundaries are ineffective in the measurement of many modern research problems. Therefore, constructing geographic units based on the proximity and concentration of licensed premises helps to define areas of risk where changes to licensing policy might be best observed. From Figure 2, readers will note that clusters differ in size, spatial concentration of

14. NNI measures the clustering of events in space and the significance of those clusters; when the NNI < 1, it indicates spatial clustering. For the spatial distributions of bars in this sample, we found that the spatial distribution before the Act was NNI = 0.4371, $p < .0001$, Z-score = –24.6964, and after the Act was NNI = 0.4309, $p < .0001$, Z-score = –26.9138. These results indicate that both before and after the change in law, licensed premises were clustered spatially; after the Act, several new premises contributed to increasing fractionally the degree to which these premises were clustered.

15. In attempting to identify meaningful cluster groups of licensed premises, we also conducted nearest-neighbor hierarchical clustering (Nnh) in CrimeStat Version III. However, after running the analysis under numerous parameter settings, the results of this analysis (based on local knowledge of the city) did not represent local drinking circuits accurately.

16. Therefore, the full sample of premises was reduced to 593 for the cluster-level analysis.

09012-CrimJournal-Guts.indd 51 12/21/09 7:14:00 PM

their points, and proximity to other clusters. These issues could prove to be important avenues for consideration at later stages of the evaluation.

To observe whether the Act succeeded in cultivating staggered closing times and to evaluate whether the Act's provisions were implemented with integrity, a measure of staggering would be required. To accomplish this measurement, we used the index of dispersion (*D*), which measures the dispersion of counts of events that occur across a set of time or space intervals and is expressed as follows:

$$\text{Index of Dispersion (D)} = \frac{\sigma^2}{\mu}.$$

First, we calculated the observed premises closing within any 1 of 11 hourly units that spanned the nighttime period (10 p.m.–8 a.m.). We then calculated an expected frequency, in other words, the amount of premises likely to close in each hourly unit if closing was dispersed equally across the nighttime units. To calculate the variance, we then squared the differences (expected–observed) for each nighttime hourly unit for both before and after distributions. Then, we divided the summed variance by the average number of premises expected to close if the distribution was dispersed equally. We repeated this process for before and after distributions for both the full sample (611 premises) and the geographically clustered groups of premises (593 premises), which resulted in a *D* score per unit of analysis for the dispersion of closing before and after the Act. The *D* score ranged from 0 to 110, with 0 representing equally dispersed closing across the 11 hours sampled, and a score of 110 representing 100% concentration of closures (per unit of analysis) within one hourly unit of the night. A graphical example of the shape of the *D* scores is provided in Appendix C.

To examine the extent of change, we took the ratio of DBefore and DAfter. Values of one or more indicate that the events are underdispersed and represent a more even distribution of events across time or space. Values 1 more than indicate that events are overdispersed and, therefore, clustered in either time or space. In the following analyses, we will compute the index of dispersion for closing times at both city and cluster levels. In addition, the following results will observe general descriptive changes to licensing and temporal displacement.

Results

Did Licensed Practice and Alcohol Availability Change under the New Legislation?

Table 2 shows the distribution of licenses both before and after the implementation of the Act for the full sample of premises. Section A of the table shows the distribution of the 526 premises that were open throughout the study period. Of these premises, 66% had standard opening times (11 a.m. until 11 p.m., with 20 min of "drinking-up" time) before the Act, and 34% were premises that specialized in late-night entertainment holding special hours certificates (11 a.m. until 2 a.m.). After the implementation of the Act, the premises opted to alter their licenses in several ways, as shown in Section B of Table 2. A substantial proportion (78%) of the sample opted to extend their hours of trade after the liberalization of licensing regulations. However, 23% of these premises delayed extending their hours until some time after the Act's implementation. A smaller proportion of premises either did not change (19%) or reduced the hours under which

09012-CrimJournal-Guts.indd 52 12/21/09 7:14:01 PM

they had previously traded (4%). These actions suggest that licensed premises differed in the extent that they adopted practical changes under the Act's new regulatory framework.

TABLE 2

Licensing Distributions Before and After Implementation of the Act

		Full Sample ($n = 611$)			Cluster Sample ($n = 593$)		
	License Categories	***n***	**%**	**%[a]**	***n***	**%**	**%[a]**
A. Before	Standard timings	347	66	—	335	66	—
	Special hours certificates	179	34	—	176	34	—
	Total	526	100	—	511	100	—
B. After	Immediate extension	287	48	55	287	48	56
	Immediate reduction	15	3	3	14	2	3
	No change	99	16	19	97	16	19
	Delayed extension	121	19	23	109	18	21
	Delayed reduction	4	1	1	4	1	1
	New premises	85	14	—	82	14	—
	Total	611	101	101	593	99	100

[a] Percentages exclude newly opened premises and might not equal 100 due to rounding.

Given the observation that licensing practice *did* change substantially under the new licensing legislation, it is important to ask to what extent the availability of alcohol was extended alongside these changes? To do this, we compared the pre- and post-weekly trade durations of premises in the full sample. Prior to the change of regulation, the distribution of licensed closing was concentrated primarily within two hourly periods, 11 p.m. for standard timings and 2 a.m. for special hours certificates. As a result, the durations for which licensed premises were open across a weekly period were equally concentrated; 66% of premises in the pre-implementation period were open for 82.5 hours a week (standard timings), and a smaller proportion (21% of premises) were open across longer periods with durations of 102.5 hours a week.[17] After the implementation of the Act, the bimodal distribution described was removed and reinstalled with longer trading durations that were spread normally around an average weekly duration of 99.3 hours (SD = 15.3). When these deviations are broken down by post-implementation licensing decisions, it is possible to observe how these changes took place. Immediate extensions of trading times increased by 15.7 (SD = 17.1) hours a week, immediate reductions reduced trading times on average by –12.4 (SD = 12.4) hr, delayed extensions by 15.7 (SD = 8.1) hours a week, and delayed reduction by –6.5 (SD = 5.3) hours a week. These findings indicate that,

17. The 13% remaining premises were characterized as Special Hours Certificates with some hourly variations at 1 or 2 days of the week.

09012-CrimJournal-Guts.indd 53 12/21/09 7:14:01 PM

on average, licensing practices did change alcohol availability substantially after the deregulation. However, they also indicate that these changes varied from place to place, which is an important consideration for measuring effects on crime and disorder.

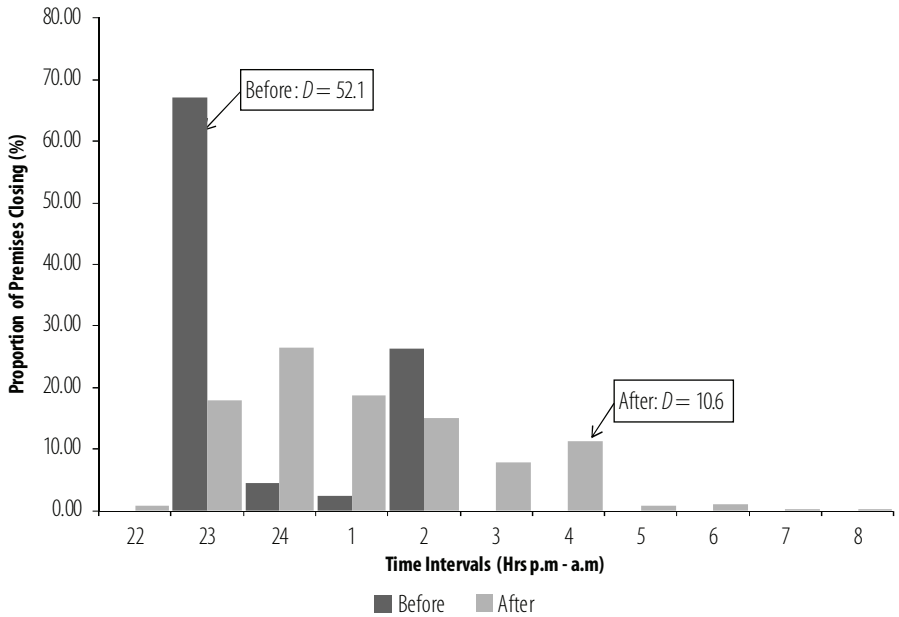
Did Staggering of Closing Times Occur?

Although the removal of fixed closing times might have spread availability through increased average sales durations, it does not mean that staggered closing times were effectively installed. In the next step, we explored the extent of staggering after the Act's implementation. Figures 3–5 show the staggering effect across the full sample (N = 611) of premises for different week periods (Friday, Saturday, and pooled weekdays). These results are restricted between the hours of 10 p.m. and 8 a.m. when 100% of premises were observed to close.

These distributions of closing hours suggest that the Act has achieved increased staggering across week and weekend periods. Licensing practices have been characterized by a substantial peak in closing times of 67% for Friday and Saturday and 71% of premises for weekdays at 11 p.m. followed by a smaller, secondary peak of 26% (Friday and Saturday) and 23% (weekdays) at 2 a.m. After the Act's implementation, primary peaks in closing have had an approximately 50% reduction for all week periods. Although deflated, peak closing has undergone some displacement; for example, Friday's peak closing (26% of premises) now occurs at 12 a.m., which is an hour later than under the previous legislative process. This displacement of 1 hr in primary peaks is notable in the other week periods. Furthermore, under the new licensing

FIGURE 3

Staggering Effects for Friday Closing (Ratio of *D* = .20)



09012-CrimJournal-Guts.indd 54 12/21/09 7:14:02 PM

FIGURE 4

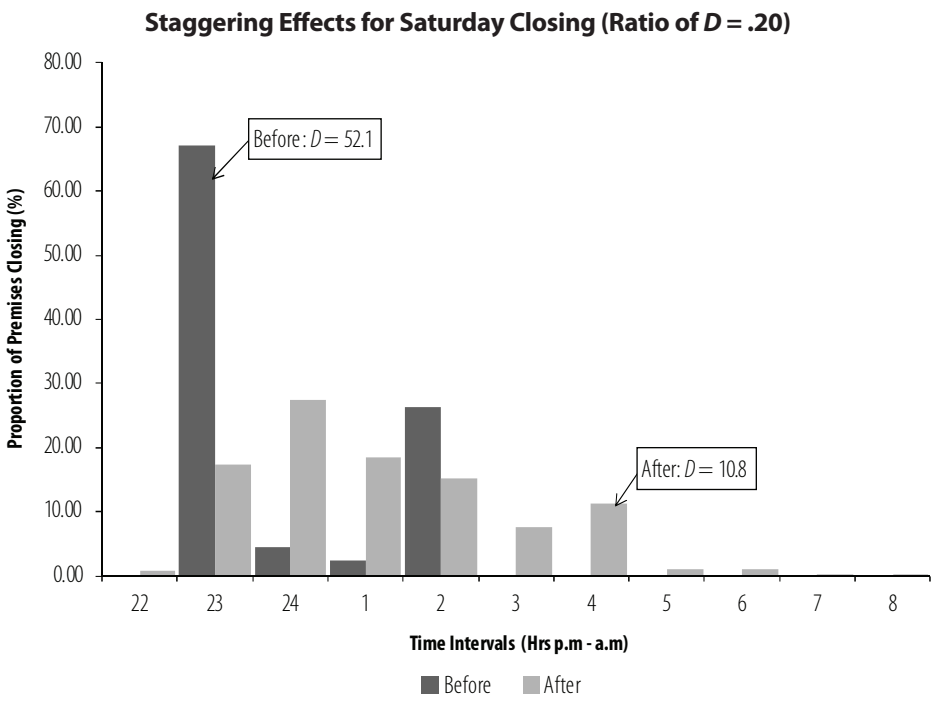
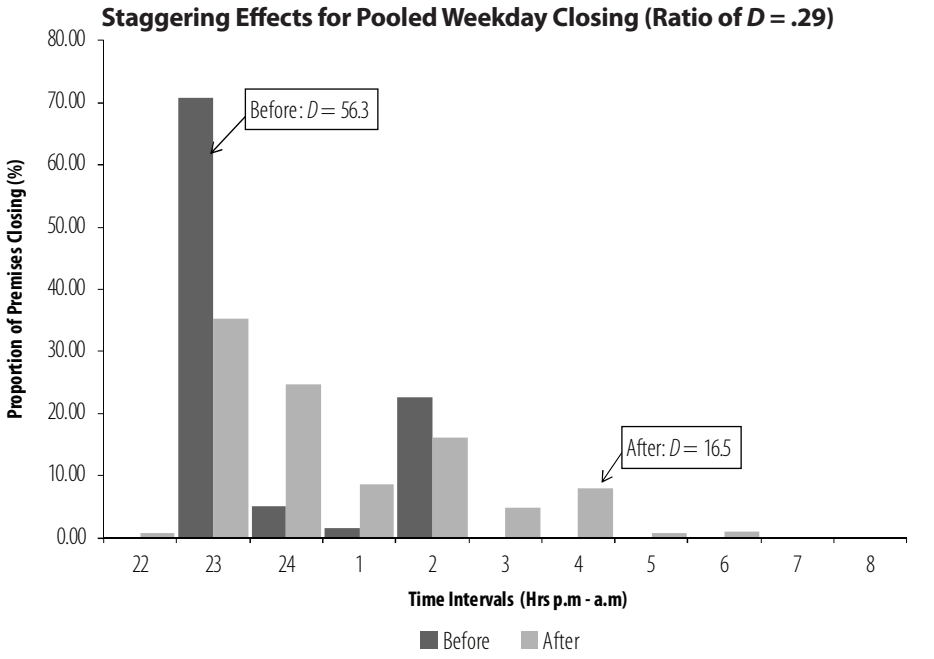


FIGURE 5



09012-CrimJournal-Guts.indd 55 12/21/09 7:14:03 PM

system, it is observed that (for each week category) three peaks occur in closing as opposed to two. For Friday, Saturday, and weekdays, the second, less pronounced peak follows at 2 a.m., and is followed by a third peak later at 4 a.m. This trend can be observed in Figure 5.

These findings suggest that staggering at the city level has occurred, albeit with several important caveats. The spread of closing times has been hypothesized as a mechanism to reduce the burden on nighttime economies as people make their way home. However, results show that, to some extent, staggering *has* occurred, yet peaks in closing seem to have remained. An important question to consider is whether staggering itself will help to reduce net rates of violence or whether staggered peaks in closing times have contributed to other problems, such as putting extra strain on police and Accident and Emergency resources during early morning periods. Consideration of staggering at the aggregate level is an important factor for resourcing and tactical decision making for emergency services and policing departments that would be required to plan where and when resources are deployed to deal with nighttime crime and disorder.

Staggering Within Premises Clusters?

Although measuring staggering at aggregate levels is useful for those individuals and agencies charged with planning and maintaining order in the nighttime economy, it is less useful as a step toward observing the possible impact of the intervention. Thus, smaller units of analysis are required to isolate the effects of the regulatory change to the localities at which they are most relevant. For example, increased dispersal of closing times at the larger unit of analysis does not dispel the concern that licensed premises in the closest proximity to one another will not have changed their hours in the way the legislation desired. Appendix B shows the results of raw analyses of staggering undertaken at cluster units of analysis. These results are summarized in the following discussion.

To begin looking at the results for staggering by cluster, we will first look at the "Ratio of *D*" columns, which is the same measure used to compare staggering in Figures 3–5. The ratio measures the two values for the index of dispersion both before and after the Act's implementation. This measure helps to simplify into one score whether change took place after the implementation of the Act (< 1 >) as well as in what direction that change occurred. If clusters scored greater than 1, then closing times became less dispersed (i.e., more concentrated). If clusters scored less than 1, then closing times within clusters became more dispersed (i.e., less concentrated). And if clusters scored exactly 1, staggering did not change.

As Table 3 shows, most clusters were observed to have staggered their closing times. However, a substantial proportion of clusters were noted *not* to have altered their closing times in a fashion consistent with the aims of the Act. For weekend evenings, a small proportion of clusters failed to stagger times between premises, 14% for Fridays (34 premises in total) and 17% for Saturdays (37 premises in total). In most (60% for Fridays, 67% for Saturdays) of these clusters, the dispersion of closing times remained the same. Despite extending hours of trade, the premises within these clusters opted to extend by the same proportions as their neighboring premises, thus negating the potential effects of staggering. The most concerning

09012-CrimJournal-Guts.indd 56 12/21/09 7:14:03 PM

finding was that a small number of clusters (such as 14 and 36) actually showed evidence of an increased concentration of closing times. According to the Act's rationale, this increase actually predicts a *greater* likelihood that violent crime and disorder might occur. Similar to the previous example, it is likely that premises made choices about closing times exclusive of staggering considerations and consequently opted to close at similar times. It should be noted that the cumulative percentage of bars that were affected by concentrating closing times was relatively small (approximately 6% of the total sample of bars on average, 7% for weekdays).

TABLE 3

Installing Staggered Closing and Removing Peak Closing

	Failure to Stagger Closing			Failure to Reduce Peak Closing Times		
	Weekdays	Friday	Saturday	Weekdays	Friday	Saturday
Sum of clusters *(n)*	7.0	5.0	6.0	26.0	19.0	20.0
Percent of clusters	19.4	13.9	16.7	72.2	52.7	55.6
Cumulative percent of premises at risk	7.4	5.7	6.2	44.1	26.9	27.8

The magnitude of the staggering effect can be observed in two columns of Appendix B. First, the *D* difference indicates the difference in scores on the index of dispersion: The greater the negative number, the greater the staggering effect; similarly, the greater the positive number, the greater the concentrating effect. In addition, it is necessary to observe how concentrated closing might have been at clusters prior to the implementation of the Act. Differences in *D* might seem small; however, this finding could be explained if the cluster was adequately staggered prior to the regulatory changes. Therefore, as a proxy measure of staggering prior to the regulatory change, the column "% at Peak Closing Before" represents the proportion of premises closing at the clusters' busiest closing times. To understand how successful staggering has been in less abstract terms, we have computed another variable that shows the percentage of premises closing at peak times after the Act's implementation. Using the Act's rationale, we could suggest that increased staggering should cause the removal of large peaks in closing times, which facilitates a spread of closing across time intervals. Despite the evidence that clusters *have* staggered their closing times, evidence is also available that improvements in staggering have made only modest impacts on reducing peaks in closing.[18] With this analysis, we found that 72% of clusters on weekday evenings failed to remove large peaks in closing times and, for Friday and Saturday, 53% and 56% (respectively) of clusters failed to remove large peaks in closing.

18. Here, the failure to remove large peaks in closing times is defined as clusters where 50% of premises closed at one time interval after the implementation of the Act.

09012-CrimJournal-Guts.indd 57 12/21/09 7:14:04 PM

This evidence suggests that the Act was successful in cultivating change in licensing routines; however, it is less clear whether the Act successfully installed systems of staggered closing (the prevention mechanism) across the study area. The present findings suggest that although staggering was achieved in most clusters, the magnitude of these changes was relatively modest. Future research needs to investigate the optimum amount of staggering required to achieve meaningful crime prevention targets. With a greater understanding of the process of regulatory change, these findings are useful for policy makers, practitioners, and academics to understand how changes in regulation can be observed and how they should be related to eventual analysis of outcomes.

Discussion

In this article, we examined the impact of a dramatic policy change aiming to prevent peaks of violent crime and disorder by altering problematic concentrations in the closing times of premises licensed to serve alcohol. The implementation of the Licensing Act (2003) deregulated fixed trading times with the hopes that a more relaxed drinking pattern could be cultivated and that peak times of closing could be removed, which were believed to contribute to concentrations in violence and disorder. According to the rationale of the Act, removing fixed closing times should activate a natural staggering of closing times that interrupts situational triggers for violence and disorder. However, the implementation of this mechanism was not planned or guided by the bodies charged with issuing licenses; these bodies were instructed to let staggering occur "naturally" through market forces (Guidance, 2006). Unlike other studies, our interest in this article has not been to examine changes in crime trends before and after the policy implementation but, instead, to investigate whether the Act's primary mechanism was implemented in accordance with its policy goals. We believed that this question is of great importance in applying meaning to analyses of crime outcomes and, therefore, should be studied before conclusions of effectiveness are made from black-box before-and-after designs.

The findings of this study have indicated that the Act *did* change patterns of opening hours among a large sample of licensed premises during the first 2 years of implementation. These findings indicate that changes in licensing were diverse with licensed premises that opted to extend, reduce, and delay changes to the hours with which they sold alcohol. A substantial proportion of the sample also chose to continue trading at the hours specified under the previous legislation. At the aggregate level, staggering of closing times was evident, yet the magnitude of this change was modest. Although closing times were spread across the nighttime period, important peaks in closing were diluted but not entirely removed. Moreover, such peaks were, for most periods of the week, temporally displaced later into the night, which *could* serve to intensify the burden on police and emergency services to stretch shifts out longer.

In observing change caused within clusters of premises, the results indicate that some degree of staggering had been achieved for most clusters. Instances in which no change or increased concentration of closing times occurred was limited to a small proportion of the sample. Despite the prevalence of increased staggering, the magnitude of this staggering effect varied between

09012-CrimJournal-Guts.indd 58 12/21/09 7:14:04 PM

clusters. Most clusters in this sample failed to reduce peak closing times for all periods of the week. As mentioned, little is known of the optimum amount of staggering required to achieve meaningful preventions in crime, or the contextual characteristics or composition of licensed concentrations that might be important in understanding problems related to crime and disorder. This research should be an avenue for future investigations.

Conclusion

This study has demonstrated the importance of observing how policy interventions are applied in real-world contexts. It shows that greater attention to implementation issues will help improve understandings of impact and effectiveness of prevention interventions. These findings have implications for both policy and evaluation practitioners. For policy makers, these findings indicate the importance for procedural, as well as theoretical, considerations to be taken into account when drafting new policies. Like other prevention interventions, ensuring that the program (or, in this case, policy) is implemented with integrity to the rationale upon which it is based is important to ensure that accurate claims can be made about the effectiveness of the intervention after evaluation. Without any knowledge of how the policy affects the targets, or if it does so at all, little can be contributed to the knowledge base for prevention. In the present example of the Act, it is feasible that Licensing Authorities could have overseen the implementation of staggered closing to ensure integrity and to assist evaluation, although practical guidance passed down from the Department of Media, Culture, and Sport (DCMS) prevented them from doing so (Guidance, 2006).

The second implication of these findings is aimed at evaluation researchers. Acknowledging that public policies will continue to be drafted and enacted without prior consideration for implementation issues, such as integrity, fidelity, and adaptation, passes some responsibility to evaluators to try to examine these issues. It seems appropriate that researchers should attempt, as much as possible, to observe the extent to which targets of interventions received the treatments intended by the policy intervention and to do so at appropriate units of analysis to avoid blurring important variations in implementation. Although these basic considerations will assist a general understanding of how a policy is applied to a real-world situation, they also are useful when incorporated into the analysis of outcomes to help strengthen the validity of research findings. Greater specification of independent variables, identification of variation in interventions (across people and places), and plotting strengths (or dose) of interventions all are important factors that can increase the reliability of findings and illuminate important practical considerations.

To conclude, we suggest that greater attention to policy implementation can help evaluation researchers avoid making incorrect causal assumptions by concluding that outcomes are attributable to an intervention that is absent or only partially implemented, which Dobson and Cook (1980) referred to as the Type III error. As Visher and Weisburd (1998) argued, a common assumption made throughout the history of crime prevention evaluation has been that "nothing works"; yet in reality, "nothing was in fact happening in crime prevention programs" (1998:

09012-CrimJournal-Guts.indd 59 12/21/09 7:14:04 PM

238). With greater attention to development and monitoring of policy interventions, perhaps the same misinterpretations can be avoided in the context of public policy interventions.

References

Babb, Penny. 2007. *Violent Crime, Disorder, and Criminal Damage since the Introduction of the Licensing Act 2003*. London, UK: Home Office.

Bailey, Trevor C. and Anthony C. Gatrell. 1995. *Interactive Spatial Data Analysis*. Essex, UK: Longman Scientific & Technical.

Bellis, Mark A., Zara Anderson, and Karen Hughes. 2006. *Effects of the Alcohol Misuse Enforcement Campaigns and the Licensing Act 2003 on Violence: A Preliminary Assessment of Accident and Emergency Attendances in Wirral*. Liverpool, UK: Northwest Regional Alcohol Strategic Group.

Bernfeld, Gary A., David P. Farrington, and Alan W. Leschfield. 2001. *Offender Rehabilitation in Practice: Implementing and Evaluating Effective Programs*. Chichester, UK: Wiley.

Campbell, Donald T. and H. Laurence Ross. 1968. Connecticut crackdown on speeding: Time-series data in quasi-experimental analysis. *Law & Society Review,* 3: 33–54.

Chen, Huey-Tsyh and Peter H. Rossi. 1983. Evaluating with sense: The theory-driven approach. *Evaluation Review,* 7: 283–302.

Cook, Thomas D. and Donald T. Campbell. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago, IL: RAND.

Dane, Andrew V. and Barry H. Schneider. 1998. Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review,* 18: 23–45.

Dobson, Douglas and Thomas J. Cook. 1980. Avoiding type III error in program evaluation: Results from a field experiment. *Evaluation and Program Planning,* 3: 269–276.

Domitrovich, Celene E. and Mark T. Greenberg. 2000. The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation*, 11: 193–221.

Dumas, Jean E., Anne M. Lynch, James E. Laughlin, Emilie Phillips Smith, and Ronald J. Prinz. 2001. Promoting intervention fidelity: Conceptual issues, methods, and preliminary results from the EARLY ALLIANCE prevention trial. *American Journal of Preventive Medicine,* 20: 38–47.

Durlak, Joseph A. and Anne M. Wells. 1997. Primary prevention mental health programs for children and adolescents: A meta-analytic review. *American Journal of Community Psychology,* 25: 115–152.

Dusenbury, Linda, Rosalind Brannigan, Mathea Falco, and William B. Hansen. 2003. *A Review of Research on Fidelity of Implementation: Implications for Drug Abuse Prevention in School Settings*. Oxford, UK: Oxford University Press.

Eisner, Manuel P. 2009. No effects in independent prevention trials: Can we reject the cynical view? *Journal of Experimental Criminology*, 5: 163–183.

09012-CrimJournal-Guts.indd 60 12/21/09 7:14:04 PM

Farrington, David. P., Trevor H. Bennett, and Brandon C. Welsh. 2007. The Cambridge evaluation of the effects of CCTV on crime. In (Graham Farrall, Kate J. Bowers, Shane D. Johnson, and Michael Townsley, eds.), *Imagination for Crime Prevention: Essays in Honour of Ken Pease*. Cullompton, UK: Willan.

Farrington, David P. and Brandon C. Welsh. 2007. *Closed-Circuit Television Surveillance and Crime Prevention: A Systematic Review*. Stockholm, Sweden: Swedish National Council for Crime Prevention.

Fortin, Marie-Josée and Mark Dale. 2005. *Spatial Analysis: A Guide for Ecologists*. Cambridge, UK: Cambridge University Press.

Hadfield, Phillip. 2006. *Bar Wars: Contesting the Night in Contemporary British Cities*. Oxford, UK: Oxford University Press.

Hadfield, Phillip and Fiona Measham. 2009. England and Wales. In (Phillip Hadfield, ed.), *Nightlife and Crime: Social Order and Governance in International Perspective*. Oxford, UK: Oxford University Press.

Hobbs, Dick, Phillip Hadfield, Stuart Lister, and Simon Winlow. 2003. *Bouncers: Violence and Governance in the Night-Time Economy*. Oxford, UK: Oxford University Press.

Homel, Ross and Jeff Clark. 1994. The prediction and prevention of violence in pubs and clubs. In (Ronald V. Clarke, ed.), *Crime Prevention Studies*. Monsey, NY: Criminal Justice Press.

Hope, Tim. 1986. Liquor licensing and crime prevention. *Home Office Research and Planning Unit Research Bulletin*, 20: 5–8.

Horney, Julie and Cassia Spohn. 1989. *The Impact of Rape Reform Legislation: Final Report to the National Institute of Justice and the National Science Foundation*. Washington, DC: National Institute of Justice.

Horney, Julie and Cassia Spohn. 1990. Issues in legal-impact research. In (Kimberly Kempf-Leonard, ed.), *Measurement Issues in Criminology*. New York: Springer-Verlag.

Hough, Mike and Gillian Hunter. 2008. The 2003 Licensing Act's impact on crime and disorder: An evaluation. *Criminology and Criminal Justice*, 8: 239–260.

Humphreys, David. K. 2007. *Alcohol-Related Violence in the Urban Environment: A Preliminary Assessment of the Effects of the Licensing Act (2003) on the Prevention of Violence and Disorder*. Cambridge, UK: Institute of Criminology.

Legendre, Pierre and Louis Legendre. 1998. *Numerical Ecology*. Amsterdam, the Netherlands: Elsevier.

Lipsey, Mark. 1993. Theory as method: Small theories of treatments. *New Directions for Program Evaluation*, 57: 5–38.

Marsh, Peter E. and Kate Fox-Kibby. 1992. *Drinking and Public Disorder.* London, UK: The Portman Group.

Michalic, Sharon F. and Katherine Irwin. 2003. Blueprints for violence prevention: From research to real-world settings—Factors influencing the successful replication of model programs. *Youth Violence and Juvenile Justice*, 1: 1–23.

Moncher, Frank J. and Ronald J. Prinz. 1991. Treatment fidelity in outcome studies. *Clinical Psychology Review,* 11: 247–266.

09012-CrimJournal-Guts.indd 61 12/21/09 7:14:04 PM

Moore, Simon C., Mario Flajšlik, Paul L. Rosin, and David Marshall. 2008. A particle model of crowd behavior: Exploring the relationship between alcohol, crowd dynamics and violence. *Aggression and Violent Behavior*, 13: 413–422.

Ned Levine & Associates. 2005. *User Manual for CrimeStat III*. Washington, DC: National Institute of Justice.

Newton, Alastair, Shah Jalal Sarker, Gurjinderpal Pahal, Eric Van den Bergh, and Charles Young. 2007. Impact of the new U.K. licensing on emergency hospital attendances: A cohort study. *Emergency Medical Journal*, 24: 532–534.

Newton, Andrew, Alex Hirschfield, Rachel Armitage, Michelle Rogerson, Leanne Monchuk, and Aidan Wilcox. 2008. *Evaluation of the Licensing Act: Measuring Crime and Disorder in and around Licensed Premises*. Huddersfield, UK: University of Huddersfield Applied Criminology Centre.

Norström, Thor and Ole-Jørgen Skog. 2003. Saturday opening of alcohol retail shops in Sweden: An impact analysis. *Journal of Alcohol Studies*, 64: 393–401.

Norström, Thor and Ole-Jørgen Skog. 2005. Saturday opening of alcohol retail shops in Sweden: An experiment in two phases. *Addiction*, 100: 767–776.

Peterson, Lizette, Andrew L. Homer, and Stephen A. Wonderlich. 1982. The integrity of independent variables in behaviour analysis. *Journal of Applied Behaviour Analysis*, 15: 477–492.

Petrosino, Anthony and Haluk Soydan. 2005. The impact of program developers as evaluators on criminal recidivism: Results from meta-analyses of experimental and quasi-experimental research. *Journal of Experimental Criminology*, 1: 435–450.

Pike, Sarah, Jo O'Shea, and Jorgen Lovbakke. 2008. *Early Experiences of the Licensing Act (2003) in the East of England and Yorkshire and the Humber Regions*. London, UK: Home Office.

Raistrick, Duncan, Ray Hodgson, and Bruce Ritson. 1999. *Tackling Alcohol Together: The Evidence Base for A UK Alcohol Policy*. London, UK: Free Association Books.

Rengert, George and Brian Lockwood. 2009. Geographical units of analysis and the analysis of crime. In (David Weisburd, Wim Bernasco, and Gerben J. N. Bruinsma, eds.), *Putting Crime in its Place: Units of Analysis in Geographic Criminology*. New York: Springer.

Rossi, Peter H., Howard E. Freeman, and Mark Lipsey. 1999. *Evaluation: A Systematic Approach*. Thousand Oaks, CA: Sage.

Sanchez, Victoria, Allan Steckler, D. Hallfors, H. Cho, P. Brodish. 2007. Fidelity of implementation in a treatment effectiveness trial of Reconnecting Youth. *Heath Education Research,* 20: 95-107.

Shadish, William. R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin.

Shepherd, Jonathan and Mark Brickley. 1996. The relationship between alcohol intoxication, stressors, and injury in urban violence. *British Journal of Criminology*, 36: 546–566.

09012-CrimJournal-Guts.indd 62 12/21/09 7:14:04 PM

Sherman, Lawrence, Denise C. Gottfredson, Doris L. MacKenzie, John E. Eck, Peter Reuter, and Shawn D. Bushway. 1997. *Preventing Crime: What Works, What Doesn't, What's Promising*. Washington, DC: U.S. Department of Justice.

Tabachnick, Barbara G. and Linda S. Fidell. 2006. *Using Multivariate Statistics*. Boston, MA: Pearson.

Tuck, Mary. 1989. *Drinking and Disorder: A Study of Non-Metropolitan Violence*. London, UK: Home Office.

Visher, Christy A. and David Weisburd. 1998. Identifying what works: Recent trends in crime prevention strategies. *Crime, Law, and Social Change*, 28: 223–242.

Weisburd, David, Gerben J.N. Bruinsma, and Wim Bernasco. 2008. Units of analysis in geographic criminology: Historical development, critical issues, and open questions. In (David Weisburd, Wim Bernasco, and Gerben J.N. Bruinsma, eds.), *Putting Crime in its Place: Units of Analysis in Geographic Criminology*. New York: Springer.

Welsh, Brandon C. and David P. Farrington. 2001. Toward an evidence-based approach to preventing crime. *ANNALS of the American Academy of Political and Social Science,* 578*:* 158–173.

Welsh, Brandon C. and David P. Farrington. 2007. *Preventing Crime: What Works for Children, Offenders, Victims, and Places*. New York: Springer.

Statutes Cited

Guidance Issued Under Section 182 of the Licensing Act 2003. 2006. London, UK: Department for Culture, Media and Sport.

David K. Humphreys is a Ph.D. candidate at the Institute of Criminology at the University of Cambridge, UK. His doctoral thesis investigates the impact on violent crime and disorder of a natural experiment in changes to alcohol regulation. His wider research interests include methodological issues in the evaluation of experimental prevention programs, situational crime prevention, geographical criminology, and substance abuse.

Manuel P. Eisner is a professor of developmental and comparative criminology at the Institute of Criminology, University of Cambridge, UK. His research interests include historical trends in violent crime, developmental criminology, and experimental prevention research. He is the director of the Zurich Project on the social development of children, a longitudinal study of 1,200 children in the city of Zurich.

09012-CrimJournal-Guts.indd 63 12/21/09 7:14:05 PM

APPENDIX A

Existing Evaluations of the Impact of the Licensing Act (2003)

Study	Location	Study Period	Design	Findings	Funding Body
Bellis (2006)	Merseyside	April 2003 to March 2006 (4 times, 3-month comparison groups, 3 control, 1 treatment)	Basic before/after natural experiment	Licensing intervention effect = Significant ($p < .001$) decrease when compare with a pooled "before" control; Significant ($p < .05$) decrease when compared with individual year control groups.	Northwest Regional Alcohol Strategic Group
Babb (2007)	National	October 2004 –November 2006 (1 year after, 1 year + 1 month before)	Descriptive analysis of temporal crime trends across time	Serious violent crime: Slight fall in total offenses but increase between 3:00 a.m. and 6:00 a.m.; less serious wounding: reduction between 11:00 p.m. and 12:00 a.m., rise between 3:00 a.m. and 6:00 a.m.; assault without injury: general decline in daytime and early evening hours, rise in offenses between 12:00 a.m. and 6:00 a.m. Harassment: general increase, most substantially between 12:00 a.m. and 2:00 a.m.; criminal damage: general decline during daytime hours, slight increase between 11:00 p.m. and 6:00 a.m.	Home Office
Humphreys (2007)	Large metropolitan city center	November 2003 to March 2006 (analysis of 3 groups, November to February 2003–2005)	Basic before/after natural experiment	Crime: modest increases in less serious violence (2.2%); burglary, theft, and handling of stolen goods (all < 1%) ($\chi^2 = 38.63$, $p < .0001$); alcohol-related violence: increased 2.3% in experimental period ($\chi^2 = 8.759$, $p < .005$, OR = 1.54, CI = 1.05–1.27); temporal patterns: remained consistent with comparison groups; spatial concentration (ARV): remained stable, 50% of ARV committed in approximately 10% of all alcohol-related crime locations.	Independent Police Review
Newton et al. (2007)					National Health Service (NHS)

09012-CrimJournal-Guts.indd 64 12/21/09 7:14:05 PM

Study	Location	Study Period	Design	Findings	Funding Body
Newton et al. (2008b)	a) Birmingham b) Blackpool c) Croydon d) Guildford e) Nottingham	November 23, 2003 to November 23, 2006	Multisite longitudinal status comparison	VATP: statistically significant increase of 11.5% ($p < .05$) in only one case study area; CD: no significant changes overall, slight reductions in 3 case study areas and fractional increase in 2 case study areas; DCS: overall decrease of 1.7%, substantial decrease in 1 case study area, subtle fluctuations in all other areas; SO: overall decrease of 5.6%, slight increases in 2 case study areas, substantial decreases in the remaining 3 areas.	Home Office
Pike et al. (2008)	6 market towns (East England)/4 towns (Yorkshire and Humber)	December 2004 to May 2006 (4 times, 6-month comparison groups)	Multisite before/after outcome analysis	Crime and disorder: East—overall= –0.8%, town 1 = –21.7%, 2 = 39.5%, 3 = 31.5%, 4 = –12.%, 5 = 17.8%, 6= –6.5%, Yorkshire and Humber: overall = –26.5%, town 1 = 9.6%, 2 = not reported, 3 = 0%, –33.4%. Temporal patterns: little significant change.	Home Office

Notes. AF = after; BF = before; CD = criminal damage; CI = confidence interval; DCS = disorder calls for service; OR = odds ratio; SO = sexual offenses; VATP = Violence Against the Person.

09012-CrimJournal-Guts.indd 65 12/21/09 7:14:05 PM

APPENDIX B

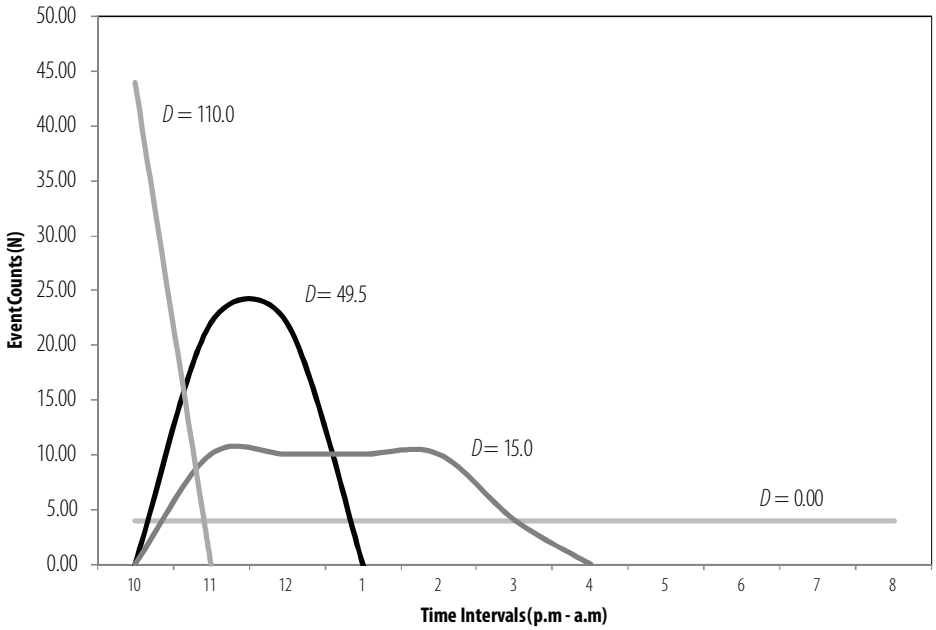
Results of Staggering for Premises Clusters

			Week				Friday				Saturday			
Cluster No.	No. Bars in Post Period	% of Sample	D Ratio	D Difference	% at Peak Closing Before	% at Peak Closing After	D Ratio	D Difference	% at Peak Closing Before	% at Peak Closing After	D Ratio	D Difference	% at Peak Closing Before	% at Peak Closing After
34	88	14.84	**.29**	–34.07	59.74	34.09	**.20**	–42.13	64.94	27.27	**.27**	–38.60	64.94	26.14
43	85	14.33	**.20**	–33.09	47.06	22.35	**.22**	–31.54	51.47	24.71	**.24**	–30.91	51.47	25.88
41	61	10.29	**.20**	–41.99	68.63	26.23	**.19**	–43.13	66.67	26.23	**.18**	–43.78	66.67	26.23
47	47	7.93	**.25**	–34.80	57.89	25.53	**.20**	–33.12	50.00	21.28	**.19**	–33.45	50.00	19.15
38	34	5.73	**.66**	–24.61	82.14	64.71	**.76**	.76	75.00	64.70	**.82**	–10.99	75.00	67.65
49	30	5.06	**.25**	–37.92	62.96	30.00	**.22**	–41.71	62.96	23.33	**.22**	–41.71	62.96	23.33
50	19	3.20	**.48**	–50.29	94.12	57.89	**.60**	–26.84	76.47	52.63	**.60**	–26.84	76.47	52.63
14	18	3.04	**1.15**	6.66	58.82	61.11	**1.21**	6.83	52.94	61.11	**1.51**	16.71	52.94	61.11
2	15	2.53	**.60**	–26.17	77.78	53.33	**.47**	–34.78	77.78	40.00	**.47**	–34.78	77.78	40.00
27	15	2.53	**.55**	–35.82	85.71	53.33	**.32**	–54.10	85.71	33.33	**.32**	–54.10	85.71	33.33
6	12	2.02	**.31**	–75.63	100.00	41.67	**.20**	–70.65	90.00	33.33	**.20**	–70.65	90.00	33.33
1	11	1.85	**.49**	–46.00	90.91	63.64	**.22**	–70.00	90.91	36.36	**.22**	–70.00	90.91	36.36
4	11	1.85	**.56**	–48.00	100.00	72.73	**.35**	–72.00	100.00	54.55	**.35**	–72.00	100.00	54.55
16	11	1.85	**.27**	–80.00	100.00	36.36	**.20**	–88.00	100.00	36.36	**.20**	–88.00	100.00	36.36
42	11	1.85	**.41**	–40.17	77.78	45.45	**.32**	–46.17	77.78	36.36	**.32**	–46.17	77.78	36.36
19	10	1.69	**.36**	–70.18	100.00	50.00	**.23**	–84.70	100.00	40.00	**.23**	–84.70	100.00	40.00
20	10	1.69	**.30**	–77.44	100.00	40.00	**.23**	–84.70	100.00	40.00	**.23**	–84.70	100.00	40.00
30	9	1.52	**.82**	–15.36	87.50	77.78	**.46**	–45.23	87.50	44.44	**.49**	–42.25	87.50	55.56
29	8	1.35	**.76**	–26.47	100.00	87.50	**.28**	–79.41	100.00	37.50	**.28**	–79.41	100.00	37.50
48	8	1.35	**.48**	–56.72	100.00	62.50	**.35**	–71.84	100.00	50.00	**.35**	–71.84	100.00	50.00
3	7	1.18	**.73**	–29.63	100.00	85.71	**.46**	–59.27	100.00	57.14	**.46**	–59.27	100.00	57.14
10	7	1.18	**1.00**	.00	100.00	100.00	**.55**	–49.39	100.00	71.43	**.73**	–29.63	100.00	85.71
13	7	1.18	**.73**	–29.63	100.00	85.71	**.37**	–69.14	100.00	57.14	**.33**	–74.08	100.00	42.86
21	7	1.18	**.46**	–59.27	100.00	57.14	**.33**	–74.08	100.00	42.86	**.33**	–74.08	100.00	42.86
24	7	1.18	**.69**	–24.69	85.71	85.71	**1.00**	.00	85.71	85.71	**1.00**	0.00	85.71	71.43
17	6	1.01	**1.00**	.00	100.00	83.33	**.69**	–33.61	100.00	83.33	**.69**	–33.61	100.00	100.00
32	6	1.01	**.69**	–33.61	100.00	83.33	**.33**	–73.94	100.00	50.00	**.33**	–73.94	100.00	50.00
44	6	1.01	**.51**	–53.78	100.00	50.00	**.45**	–60.50	100.00	50.00	**.45**	–60.50	100.00	66.67
5	4	.67	**1.00**	.00	100.00	100.00	**.45**	–60.50	100.00	50.00	**.59**	–45.38	100.00	75.00
23	4	.67	**.77**	–15.13	75.00	50.00	**.53**	–30.25	75.00	50.00	**.53**	–30.25	75.00	50.00
37	4	.67	**.77**	–15.13	100.00	50.00	**.31**	–75.63	100.00	50.00	**.31**	–75.63	75.00	50.00
9	3	.51	**.51**	–53.78	100.00	100.00	**1.00**	.00	100.00	100.00	**1.00**	.00	100.00	66.67
18	3	.51	**1.00**	.00	100.00	66.67	**.51**	–53.78	100.00	100.00	**1.00**	.00	100.00	100.00
22	3	.51	**1.00**	.00	66.67	66.67	**1.00**	.00	66.67	66.67	**1.00**	.00	66.67	66.67
36	3	.51	**2.22**	60.50	50.00	66.67	**1.14**	6.72	50.00	66.67	**1.14**	6.72	50.00	100.00
45	3	.51	**.51**	–53.78	100.00	33.33	**.27**	–80.67	100.00	33.33	**.27**	–80.67	100.00	66.67

09012-CrimJournal-Guts.indd 66 12/21/09 7:14:06 PM

APPENDIX C

Examples of Dispersion (*D*) Distributions



09012-CrimJournal-Guts.indd 67 12/21/09 7:14:06 PM

ALCOHOL POLICY IMPLEMENTATION EVALUATION

After the Act

Alcohol licensing and the administrative governance of crime

Phil Hadfield
University of Leeds

Fiona Measham
Lancaster University

Alcohol and crime have a long and checkered relationship that perfectly embodies some central dilemmas of good government. How should public policy balance the freedoms of the individual to consume as he or she wishes with those of public health and social order? The self-evident premise that underpins such questions surrounds the status of alcohol as a powerful psychoactive substance and, therefore, "no ordinary commodity" (Babor, Caetano, Casswell, and Edwards, 2003). Numerous responsibilities and limitations are placed on those who supply citizens with alcohol that befits a legally controlled drug. One key dimension of this drinking governance is the attempt by regulators to control the circumstances under which alcohol purchases are made. The typical policy levers that control "availability" are pricing and taxation (Meier, Booth, Stockwell, Sutton, Wilkinson, and Wong, 2008), the number and/or density of outlets (Livingston, Chikritzhs, and Room, 2007), and age restrictions and trading hours (Stockwell and Chikritzhs, 2009). Humphreys and Eisner's (2010, this issue) article is concerned with the Licensing Act of 2003 (henceforth referred to as "the Act"), which is a statute through which the British government is explicit in its intention to use the manipulation of trading hours for businesses licensed to supply alcohol as a situational mechanism to control crime. It also has a much more ambitious underlying aim of instigating a change in the longstanding British drinking culture that is characterized as "binge and brawl." The Act sought to introduce a "flexible" regime in which the previously "fixed" closing times that applied across England and Wales could be replaced with a regulatory framework that permitted alcohol sales up to 24 hours a day, 7 days a week. This Act was, in effect, a deregulatory measure that

Direct correspondence to Fiona Measham, Department of Applied Social Science, Bowland North Lancaster University, Lancaster, LA1 4YT, UK (e-mail: F.Measham@lancaster.ac.uk); Phil Hadfield, Centre for Criminal Justice Studies, School of Law, 20 Lyddon Terrace, University of Leeds, LS2 9JT, JK (e-mail: phil@philhadfield.co.uk).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 69 12/21/09 7:14:06 PM

potentially *increased* the availability of alcohol. Considerable faith was placed in the argument that alcohol-related public assault and nuisance offenses would decrease if drinkers in urban entertainment zones no longer flooded onto the streets from licensed premises at a nationally enforced closing time.

The thinking behind the belief that increased trading hours could reduce alcohol-related crime and disorder was that in the established system of England and Wales, bars with 11 p.m. closing times and nightclubs with 2 a.m. closing times, typically had a short period of rapid consumption just prior to closing time and, consequently, raised blood alcohol concentrations in customers leaving the licensed premises. An urban-management problem then ensued with tensions at congested and overburdened transport hubs and fast food outlets, which are identified as "flashpoints" for alcohol-related crime and disorder (Tuck, 1989). Official statistics confirm that one in five violent crimes and more than half of the alcohol-related violent crimes that take place between strangers and acquaintances occur in or around pubs and clubs, and that 75% of the alcohol-related violent crime that takes place between strangers occurs on evenings during the weekend (Walker, Flatley, Kershaw, and Moon, 2009).

As Humphreys and Eisner (2010) note, the Act, therefore, represents a "rare example of a measure that tries to achieve crime reductions by *removing restraints* rather than by adding them." This restraint removal was perhaps most evident in the intention to limit the powers of local regulators to engineer closing times and to trust in the ability of "market forces"—specifically, the commercial preferences of licensed businesses to seek and obtain a range of closing times for their outlets—to deliver the desired implementation effects and attendant crime reduction goals.

However, the Act represented far grander aspirations for the New Labour Government than mere crime prevention, and it was those aspirations that led to the contradictions and ultimate failure of the Act to deliver the politicians' promises. The broader timing of the Act was linked to heightened concerns regarding excessive alcohol consumption in the United Kingdom, in particular, a pattern of weekend "binge" drinking by young adults. Commentators have argued that a "new culture of intoxication" developed in the 1990s, whereby young people's increased sessional consumption and a "normalisation of determined drunkenness" (Measham, 2004: 321) was facilitated by the expansion of the British nighttime economy (NTE; Hobbs, Hadfield, Lister, and Winlow, 2003), the development of continental-style café bars, and a widening range of high-strength alcoholic beverages (such as "alcopops" and shots) that appealed to young women as well as young men (Measham and Brain, 2005). Weekly alcohol consumption by 11–15-year-olds and 16–24-year-olds doubled in the 1990s with images of drunken teenagers filling the press (Berridge, Thom, and Herring, 2007; Plant and Plant, 2006). It was against this backdrop that a growing consensus developed that something needed to be done about Britain's binge-drinking problem, which is characterized as the "British disease."

Despite a broad church of opposition from the policing and health professionals as well as from alcohol researchers, the British government persisted with a piece of legislation that it

09012-CrimJournal-Guts.indd 70 12/21/09 7:14:06 PM

claimed would reduce binge drinking and its associated problems. Moreover, the Culture Minister at that time ambitiously declared that the introduction of flexible closing times would result in more relaxed and "Mediterranean-style" drinking—a culture to be aspired to, which is in stark opposition to the "binge and brawl" Northern European model. In the United Kingdom and across the rest of Europe, this contrast between Northern and Southern European drinking cultures was played out in local council policies designed to promote café bar commerce at the city's center (Roberts, Turner, Greenfield, and Osborn, 2006). The British Government was not alone, therefore, in aspiring to reduce binge drinking through *increased* access to alcohol and romanticization of the Mediterranean way of life (Jayne, Valentine, and Holloway, 2008), although they were perhaps more explicit than other countries in pursuing so ambitious a target of cultural change through one piece of licensing legislation. Although the Act did not produce the devastation predicted by some, the folly of aspiring to such a deep-seated change in drinking behaviors through increased trading hours is evident in recent proposals by both the Prime Minister and the Leader of the Opposition to allow councils to reverse the Act less than 4 years after its introduction (Porter, 2009; Travis, 2009).

Humphreys and Eisner (2010) considered what implications resulted from the decision of the British Government not to adopt a robust approach to the implementation of their new policies. Using data from a large U.K. city, some of the resultant tensions between the policy pronouncements and the policy effects have been revealed, and the obstacles that prevent adequate evaluation in cases where implementation are limited or even nonexistent. In the remainder of this essay, we introduce readers to the policy context in which the Act might be situated.

Alcohol and Crime Reduction Policy in England and Wales

Given the objectives of the Act—the prevention of crime and disorder, promotion of public safety, prevention of public nuisance, and protection of children from harm—international readers might be surprised to learn that the arm of central government tasked with developing and administrating the Act is the Department for Culture, Media, and Sport (DCMS)—colloquially known as the "Ministry of Fun." This assignment has important implications, not only at a bureaucratic level but also symbolically; for many commentators in the research, policing, and public health community, it underlines a concern that the British government is aligned closely with business interests within the alcohol (manufacture and retail), leisure, and tourism sectors (Room, 2004; Shenker, Gilmore, Craik, and Rogers, 2009), which relegates the Home Office and Department of Health to the periphery of alcohol policy. Moreover, the bodies tasked with the day-to-day administration of licensing are not regional police forces but local government licensing committees. Standing outside of the criminal justice apparatus and having no special training in security or community safety matters, these tribunals nonetheless deploy civil and administrative mechanisms of central importance to the strategic governance of crime and disorder in British towns and cities.

09012-CrimJournal-Guts.indd 71 12/21/09 7:14:06 PM

Despite an avowed commitment to "evidence-based policy," the British government could not be described as having adopted a scientific approach to the formulation, development, or implementation of the Act. No pilot projects or trial locations were employed, although the absence of investment in prior research—even to a minimal stage of comprehensive literature review— prompted some to describe the proposals as a "leap in the dark" (Plant and Plant, 2005). In actuality, the weight of international evidence regarding the effects of extending trading hours for pubs and bars across a range of crime and public health impacts has been shown to run contrary to the claims of the British Government (see Stockwell and Chikritzhs, 2009). Yet, the resulting concerns run deeper than those of mere incompetence and encompass evidence of suppression and selective editing in the drafting of an official document that summarizes international knowledge on trading hours (Hadfield, 2006; Stockwell, 2004). This issue remains a serious charge for which the Labour Administration's Cabinet Office has yet to account for. It could be viewed in the context of a series of political actions that suggest a preference for economically protectionist and popular punitive solutions to a nation's alcohol problems (Hadfield, Lister, and Traynor, 2009).

Humphreys and Eisner (2010) used a city in Northern England as a case study and consider licensing trends (opening and closing times for licensed premises) during a 4-year period—2 years before and 2 years after the Act was introduced in November 2005. Humphreys and Eisner were not concerned with an attempt to measure the crime reduction outcomes of the legislation but with the extent to which the Act succeeds in applying its main crime reduction lever—a more staggered pattern of closing times for licensed premises. Humphreys and Eisner show that as a result of local licensing committees not being given sufficient powers to dictate or impose closing times in response to local circumstances alongside a lack of commercial motivation, outcomes varied and did not always result in the desired staggering effect.

Aggregate data for England and Wales shows that for outlets supplying alcohol for consumption on the premises, the overall effects of the Act were limited to an average Saturday night trading-time extension of 21 minutes with only 1% of premises that previously had not closed after midnight opting to close beyond midnight (DCMS, 2008). Such data add to the suggestion of limited implementation while masking significant local variations. Humphreys and Eisner (2010) report major divergences at the micro level—although most of the licensed premises clusters they identified had taken advantage of the new opportunities for extended hours, each premise often opted to extend by a similar proportion to their neighbors, which negated the potential effects of staggering (the prevention mechanism) and created new peaks in closing that were displaced temporally longer into the night. Humphreys and Eisner provide a map of the premises clusters, which demonstrates that clusters are not confined to the central core of the city. Humphreys and Eisner do not comment on how different types of neighborhoods (e.g., affluent suburbs and deprived inner cities) might have experienced the extensions and staggering effects in potentially different ways, which resulted in a geographical as well as temporal displacement rather than in a reduction in alcohol-related crime and disorder. Such questions

09012-CrimJournal-Guts.indd 72 12/21/09 7:14:07 PM

point to the importance of local contextual knowledge. Through statistical analysis, Humphreys and Eisner reveal some of the implementation factors that might be qualitatively apparent. A more confounding variable might be, for example, that in practice, licensed premises close at a time earlier than that allowed by their license (Hadfield, 2007; Home Office, 2008).

The degree of disparity within one city raises obvious questions that concern the implementation of national policy *between* cities. Recent studies from the United Kingdom that evaluate the Act report variations between sites in relation to overall crime impacts such as police-recorded violent crime and criminal damage (or vandalism), whereas at a national level, the studies found "no obvious impact" according to a range of aggregate measures, which include crime statistics, victim surveys, and medical statistics (Hough and Hunter, 2008: 239). The effects of a differential uptake of extended hours and the consequent divergences in achieving staggering within and between sites were not considered, which is a shortcoming Humphreys and Eisner (2010) have noted clearly. As Stockwell and Chikritzhs (2009: 165) concluded, "had the U.K. case studies been able to link violent incidents to particular licensed venues and to also compare those that took advantage of the extra hours and those which did not," they might well have reported similar results to studies from Australia and others jurisdictions, which have reported predominately more negative crime outcomes. As Stockwell and Chikritzhs went on to note, recent U.K. studies, like many others purporting to measure the impacts of extended hours, have failed to account for the impact of concurrent policy changes likely to produce confounding effects. For the Act, these changes included a significant investment by the British Government in increased police activity focused on alcohol-related problems between 2004 and 2007 and other major initiatives around alcohol-related harm, such as the launch of the *Alcohol Harm Reduction Strategy for England* in 2004 and the *National Alcohol Strategy* in 2007 as well as subsequent amendments (see Hadfield and Measham, 2009). One, of course, might speculate on the crime reduction efficacy of a policy that has failed to remove the peaking or clustering of closing times while concurrently extending the availability of alcohol and displacing crime and disorder incidents longer into the night and further out from city centers.

Humphreys and Eisner's (2010) central point is that in the absence of a concerted effort to ensure that the stated aims are achieved, the uptake and impact of policies might vary from place to place and, in some cases, might amount to implementation failure. The research community should be mindful of such factors when attempting to measure and evaluate the effects of policy interventions. The major contribution of Humphreys and Eisner is to remind us that minutiae matters; too often, previous research in this field has involved superficial retrospection, or simple "before and after" comparisons of aggregate data that are not nuanced sufficiently to provide sound bases for public policy. The article is suggestive of the need for researchers to pay attention to local contextual factors, measured most effectively through the use of mixed-method alcohol–crime nexus case studies within the cultural context of "naturally occurring" local leisure scenes and clearly specified regulatory landscapes.

09012-CrimJournal-Guts.indd 73 12/21/09 7:14:07 PM

As a tool for situational crime reduction, alcohol licensing operates within the context of multifaceted constraints such that the attempts of the British government at "social engineering" could be judged in hindsight as hopelessly optimistic. Alcohol licensing, thus, remains a fascinating topic of research. It demonstrates the limits of criminal justice processes in shaping crime policy and in attempting to change longstanding traditions as well as the need for criminal justice agents to engage fully with civil and administrative adjudicators, city governments, leisure businesses, and the cultural forces that shape nightlife and crime. Ultimately, it illustrates the complexities of the relationship between public policy and human behavior.

References

Babor, Thomas, Raul Caetano, Sally Casswell, and Griffith Edwards. 2003. *Alcohol: No Ordinary Commodity: Research and Public Policy*. Oxford, UK: Oxford University Press.

Berridge, Virginia, Betsy Thom, and Rachel Herring. 2007. *The Normalisation of Binge Drinking? An Historical and Cross Cultural Investigation with Implications for Action*. Unpublished final report for AERC.

Department for Culture, Media and Sport (DCMS). 2008. *Evaluation of the Impact of the Licensing Act 2003*. London, UK: Author.

Hadfield, Phil. 2006. *Bar Wars: Contesting the Night in Contemporary British Cities*. Oxford, UK: Oxford University Press.

Hadfield, Phil. 2007. A hard act to follow: Assessing the consequences of licensing reform in England and Wales. *Addiction*, 102: 177–180.

Hadfield, Phil, Stuart Lister, and Peter Traynor. 2009. This town's a different town today: Policing and regulating the night-time economy. *Criminology and Criminal Justice*, 9: 1–21.

Hadfield, Phil and Fiona Measham. 2009. England and Wales. In (Phil Hadfield, ed.), *Nightlife and Crime: Social Order and Governance in International Perspective*. Oxford, UK: Oxford University Press.

Hobbs, Dick, Phil Hadfield, Stuart Lister, and Simon Winlow. 2003. *Bouncers: Violence and Governance in the Night-time Economy*. Oxford, UK: Oxford University Press.

Home Office. 2008. *Review of the Social Responsibility Standards for the Production and Sale of Alcoholic Drinks*. Birmingham, AL: KPMG LLP.

Hough, Mike and Gillian Hunter. 2008. The Licensing Act's impact on crime and disorder: An evaluation. *Criminology and Criminal Justice*, 8: 239–260.

Humphreys, David K. and Manuel P. Eisner. 2010. Evaluating a natural experiment in alcohol policy: The Licensing Act (2003) and the requirement for attention to implementation. *Criminology & Public Policy*. This issue.

Jayne, Mark, Gill Valentine, and Sarah L. Holloway. 2008. Fluid boundaries—British binge drinking and European civility: Alcohol and the production and consumption of public space. *Space and Polity*, 12: 81–100.

Livingston, Michael, Tanya Chikritzhs, and Robin Room. 2007. Changing the density of alcohol outlets to reduce alcohol-related problems. *Drug and Alcohol Review*, 26: 557–566.

09012-CrimJournal-Guts.indd 74 12/21/09 7:14:07 PM

Measham, Fiona. 2004. The decline of ecstasy, the rise of "binge" drinking and the persistence of pleasure. *Probation Journal*, 51: 309–326.

Measham, Fiona and Kevin Brain. 2005. "Binge" drinking, British alcohol policy, and the new culture of intoxication. *Crime, Media, Culture: An International Journal*, 1: 263–284.

Meier, Petra, Andrew Booth, Tim Stockwell, Anthea Sutton, Anna Wilkinson, and Ruth Wong. 2008. *Independent Review of the Effects of Alcohol Pricing and Promotion*. London: U.K. Department of Health.

Plant, Emma J. and Martin Plant. 2005. A "leap in the dark?" Lessons for the United Kingdom from past extensions of bar opening hours. *International Journal of Drug Policy*, 16: 363–368.

Plant, Martin and Moira Plant. 2006. *Binge Britain: Alcohol and the National Response*. Oxford, UK: Oxford University Press.

Porter, Andrew. 2009. Gordon Brown to give councils new powers on 24-hour licenses. *Daily Telegraph*, September 29.

Roberts, Marion, Chris Turner, Steve Greenfield, and Guy Osborn. 2006. A continental ambience? Lessons in managing alcohol-related evening and night-time entertainment from four European capitals. *Urban Studies*, 43: 1105–1125.

Room, Robin. 2004. Disabling the public interest: Alcohol strategies and policies for England. *Addiction*, 99: 1083–1089.

Shenker, Don, Ian Gilmore, Michael Craik, and Alison Rogers. 2009. Measures to curb binge drinking must not be delayed. *The New York Times*, September 24. Retrieved December 1, 2009 from timesonline.co.uk/tol/comment/letters/article6846426.ece.

Stockwell, Tim. 2004. Lies, damned lies, and no statistics: A study of dysfunctional democracy in action. *Addiction*, 99: 1090–1091.

Stockwell, Tim and Tanya Chikritzhs. 2009. Do relaxed trading hours for bars and clubs mean more relaxed drinking? A review of international research on the impacts of changes to permitted hours and drinking. *Crime Prevention and Community Safety*, 11: 153–170.

Travis, Alan. 2009. Tories pledge steep tax rises for super-strength beer and cider. *The Guardian*, October 7.

Tuck, M. 1989. Drinking and Disorder: A Study of Non-metropolitan Violence. Home Office Research Study No. 108. London: HMSO.

Walker, Alison, John Flatley, Chris Kershaw, and Debbie Moon. 2009. *Crime in England and Wales 2008/09*. London, UK: Home Office.

09012-CrimJournal-Guts.indd 75 12/21/09 7:14:07 PM

Phil Hadfield is Senior Visiting Researcher at the Centre for Crimnal Justice Studies, School of Law, University of Leeds, UK. He is Director of philhadfield.co.uk a research and training consultancy specialising in crime prevention, community safety and urban cultural planning issues, with a particular reference to alcohol, drugs and the night-time economy. He is the author of numerous scholarly articles and three books. Recent publications include an edited collection entitled 'Nightlife and Crime' containing research reports from 17 countries—including 4 chapters from North America—published by Oxford University Press (2009).

Fiona Measham is Senior Lecturer in Criminology at Lancaster University, with 20 years' experience as a researcher in the field of drug and alcohol studies, gender, licensed leisure and the night time economy. She is co-author of *Illegal Leisure* (1998), *Dancing on Drugs* (2001) and *Illegal Leisure Revisited* (2010), and co-editor of *Swimming with Crocodiles* (2008). Her interests span cultural criminology, gender studies and the sociology of intoxication, with a particular interest in exploring the boundaries of transgression, the criminalisation of leisure and the problematic-recreational interface in leisure time consumption.

09012-CrimJournal-Guts.indd 76 12/21/09 7:14:07 PM

POLICY ESSAY

ALCOHOL POLICY IMPLEMENTATION EVALUATION

Why implementation matters

Recent experience with the U.K. Licensing Act (2003)

Shannon Frattaroli
The John Hopkins Bloombeg School of Public Health

If people are going to drink at bars, pubs, and clubs, then is it in the best interest of public safety to flood the streets with these drinkers by regulating a single closing time for businesses that serve alcohol? Or would public safety be served better if those same drinkers comprised a steady stream of fewer drinkers dispersed across a longer period of time? In 2003, the U.K. government signaled its response to these questions with the Licensing Act (hereafter referred to as "the Act"). The Act went into effect in 2005 and sought to reduce alcohol-related disorder and crime by removing the requirement that alcohol licensees close their shops by a centrally established time. By deregulating the operating hours that govern alcohol service establishments, policy makers predicted that the Act would result in varied closing times in communities throughout England and Wales. They also hypothesized that these staggered closing times would diffuse the previously large concentration of drinkers across time, thereby reducing opportunities for alcohol-related disorderly conduct and crime (Humphreys and Eisner, 2010, this issue).

According to Humphreys and Eisner (2010), the Act was a response to contemporary analyses, which demonstrated that a large proportion of disorder occurred during the hour after closing time for alcohol establishments set by the U.K. government. People departed en mass and entered the streets at the same time, which caused chaos. Advocates of the policy surmised that it was necessary to diffuse the population of drinkers as they made their way home. With staggered closing times, rowdy drinkers would trickle into the streets across an extended period of time and would be less likely to encounter one another as they sought cabs, space on trains, and other commodities that otherwise might lead to confusion and conflict when demand exceeds the available supply. To accomplish this aim, the Act removed the set closing time and empowered local licensing authorities to approve the hours of operation proposed by individual licensees through the licensing process.

Direct correspondence to Shannon Frattaroli, The Johns Hopkins Bloomberg School of Public Health, Center for Injury Research and Policy, 624 North Broadway, Baltimore, MD 21205 (e-mail: SFrattar@jhsph.edu).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 77 12/21/09 7:14:07 PM

Despite several evaluations of the Act, Humphreys and Eisner (2010) remained unconvinced that key evaluation issues had been accounted for in the literature to date. In particular, they were concerned that evaluation designs failed to connect the implementation of the Act adequately with the outcomes being measured. Variation in the implementation of the Act combined with city-level and country-level outcome measures could lead evaluators to miss important differences between areas where deregulation achieved staggering and areas where closing times remained uniform. By ignoring instances in which the Act produced little or no change in operating hours, the evaluators also were at risk of incorrectly attributing changes in disorder and crime to the Act. In their study, Humphreys and Eisner questioned how and to what extent the deregulation of closing hours was implemented. The resulting study was an ambitious set of analyses that provided a detailed understanding of the extent of the Act's implementation in one English city.

Considering Implementation of the Act

Whether a particular policy achieves its intended effect is a question of interest to researchers, policy makers, and the general public. Any expectation of answering questions about policy impact is predicated on some assurance that the policy was implemented with some fidelity to the policy goal. Despite this seemingly obvious logic, relatively few examples exist in which implementation questions have driven research (e.g., Frattaroli and Teret, 2006; Goggin, 1987; Pressman and Wildavsky, 1984). As Humphreys and Eisner (2010) have convincingly argued, we have many reasons to focus on the implementation process that surrounds the Act, which includes the latitude given to local authorities to specify operating hours through licensure and the option of maintaining pre-Act closing hours—both of which plausibly could lead to variation in the extent of staggering within and among communities. Humphreys and Eisner's approach to accomplish this aim has been instructive for discussing policy implementation research and the relationship between such research and policy making. Select observations from Humphreys and Eisner's study are discussed in this context.

Details and Nuances Are Essential to Implementation Research

Humphreys and Eisner (2010) designed a study to quantify changes in closing times that resulted from the Act, measure the staggering of closing times in the aggregate, and assess the extent of staggering that occurred within geographically clustered licensed alcohol establishments. With this approach, Humphreys and Eisner illustrated the complexity of implementation actions required to achieve staggered closing times. For example, under the new market-driven system of setting closing times, licensees request changes in their hours of operation, the local licensing authority approves the requests, and the result is supposed to be closing times staggered enough to affect the concentration of drinkers at levels that would reduce disorder and crime. Humphreys and Eisner's design also documented variations in implementation that occurred. By constructing clusters of premises based on geography, Humphreys and Eisner built a unit

09012-CrimJournal-Guts.indd 78 12/21/09 7:14:08 PM

of analysis that accurately reflected whether the consequences intended by the drafters of the Act were realized through implementation. Staggered closing times matter little to the theory underlying the Act if these closing times are the same within neighborhoods but different across neighborhoods throughout a city. This clustered staggering approach provided a detailed picture of implementation variation within the communities and offered important context for understanding impact measures. Compared with the pre- and post-evaluation designs in which differences in crime are compared without regard for whether or how implementation occurred, Humphreys and Eisner's analysis offered critical details with which to conduct a more informed evaluation.

Relationship between Policy Characteristics and Implementation

The Act includes mechanisms for accomplishing reductions in alcohol-related disorder and crime. These mechanisms represent decisions made during policy formulation, and they affect the implementation process. For example, the decision to move the regulatory authority for operating hours from the central government to the localities meant that the implementation process would be decentralized. This decentralized approach reflected a desire to invite local authorities and communities into the process of regulating neighborhood alcohol establishments. The process, with individual licensees proposing new operating hours, focused more on deregulation and market forces than on achieving staggered closing times. If the U.K. government had opted to take a more prescriptive approach to staggering hours, then a different implementation process would have evolved. The recognition that policy formulation involves a series of choices and that those choices have implementation implications is a potentially useful reminder for policy makers, advocates, and stakeholders that consideration of implementation should occur when developing policy.

Implementation Questions Can Be Informed by Varied Methods

How policies are implemented and the factors that influence implementation are important lines of inquiry that might be best answered through quantitative, qualitative, or mixed methods approaches. The findings of Humphreys and Eisner's (2010) study quantified the impact of the implementation process well and might prompt additional questions from readers about the factors that influence these implementation processes. Although Humphreys and Eisner's inquiry did not pursue this aspect of implementation, prior research (Hough, Hunter, Jacobson, and Cossalter, 2008) has offered some insight into the incentives and challenges associated with achieving staggered operating hours under the Act. Specifically, data collected through in-depth interviews with government officials, residents, and trade professionals have added the perspectives and experiences of the implementers as well as of the people affected by the change in policy. These qualitative interview findings have complemented the picture of implementation that emerged from Humphrey and Eisner's work.

09012-CrimJournal-Guts.indd 79 12/21/09 7:14:08 PM

Humphreys and Eisner's (2010) examination of the Act in an English metropolitan area is a fine example of the importance of applying rigorous research methods to answer questions about the implementation phase of the policy process. No doubt this work will inspire others to consider more carefully how to account for and measure the predictable variations in policy implementation and to inform evaluations designed to measure impacts of the policies.

Lessons for Improving the Implementation Phase of the Policy Process

The Act marked an important change in the approach of the U.K. government to alcohol and its regulation. Humphreys and Eisner's (2010) study provided important lessons for the general policy process as well as for the ongoing dialogue about alcohol policy in the United Kingdom. Select comments about each of these topics are offered below.

Considering Implementation during Policy Formulation

Humphreys and Eisner's (2010) examination of the Act made a powerful case for understanding how implementation affects outcomes, and described "a requirement for researchers *and* policy makers to understand *why* desired outcomes occur, not just *whether* they occur." Certainly one goal of implementation research should be to provide policy makers with tools and guidance for how to anticipate and address implementation challenges through policy development. For example, inviting stakeholders to comment during policy formulation can generate suggestions on how to accomplish the desired policy goal and provide an early test of perspectives of implementers on proposals being considered by policy makers. In the case of the Act, additional input from license holders could have generated suggestions on how to achieve staggered closing times better.

Understanding the Factors that Affect Policy Implementation

Quantifying changes in the operating hours of alcohol establishments and measuring the dispersion of closing times within neighborhood clusters of establishments provided a clear picture of how local level implementers translated the Act in one urban area. Humphreys and Eisner's (2010) study also offered a possible explanation for the inconsistent findings among the outcome evaluations of the Act to date. If the variation in implementation demonstrated in the study site is indicative of the implementation experiences throughout England and Wales, then the impact of the Act is also likely to vary.

Although Humphreys and Eisner's (2010) study answered important questions, it also raised new questions for policy makers and others interested in improving how policies are implemented. To affect implementation, understanding the factors that influence implementation is needed. In the case of the Act, Humphreys and Eisner's findings demonstrated that some licensees changed their operating hours, whereas others did not. Understanding how licensees considered this decision and the thought processes that ultimately led to a specific proposal for revised operating hours would offer important insight into how the licensees, who are the means to accomplish staggering under the Act, make decisions about their businesses.

09012-CrimJournal-Guts.indd 80 12/21/09 7:14:08 PM

Other aspects of the implementation of the Act (such as the role of local governments as well as the barriers to achieving greater variation in closing times) provide more support for Humphreys and Eisner's (2010) position that implementation is a complex process and needs to be better understood. Such understanding is crucial to making sound policy decisions and to building a research literature that is based on valid findings about the effect of prevention interventions.

Revisiting the Act

With an appreciation for the many ways in which the Act did or did not affect the closing times of drinking establishments and, therefore, did or did not alter the concentration of drinking patrons in the streets at closing, the conflicting evaluation results can be understood better as a reflection of variation in implementation. Assuming the theory underlying the Act is correct, that dispersing the population of drinkers across time as they leave bars, clubs, and pubs will reduce alcohol-related disorder, it becomes necessary to understand how to implement this law effectively. I would add to this charge the need to consider how implementation also can affect reductions in other alcohol-related health impacts (motor vehicle crash injuries, violence, cancer, and cirrhosis, to name a few). On this matter, the existing literature is instructive.

What remains unstated in the Humphreys and Eisner (2010) article is the controversy surrounding the deregulation strategy. Unlike many topics of policy debate, alcohol is well studied. Although significant challenges to realizing effective policy strategies to reduce the harmful effects associated with alcohol consumption have developed, we have authoritative evidence on many strategies that they are ripe for policy adoption. The health and safety benefits associated with decreased consumption and increased costs are well established (Anderson, Chisholm, and Fuhr, 2009; Cook, 2007; Toomey and Wagenaar, 1999). Similarly, several countries have tested their own natural experiments that involve regulatory strategies to restrict operating hours in bars, pubs, and clubs, and evaluations of those efforts offer convincing evidence of the benefits of such an approach (Duailibi, Ponicki, Grube, Pinsky, Laranjeira, and Raw, 2007; Plant and Plant, 2005). The decision of the U.K. government to eliminate the fixed closing time in favor of a liberal, locally based system of oversight presents an opportunity to revisit this decision with the benefit of implementation findings and a reconsideration of the evidence base for effective alcohol-control policies. Specifically, if a more directed policy of staggering can be achieved while simultaneously restricting operating hours, then a policy solution aimed at both removing the risks attributed to peak drinking hours and adopting a solid evidence-based strategy for reducing alcohol-related disease and injury would be expected as well as clear health and safety gains associated with the Act. The findings from Humphreys and Eisner's implementation study suggest that such a system is unlikely to develop without a clear plan, which provides more direction than relying on the market to realize this end.

09012-CrimJournal-Guts.indd 81 12/21/09 7:14:08 PM

Conclusion

Recent indications that the U.K. government remains open to considering new alcohol-control policies to address this persistent social problem are encouraging (O'Dowd, 2009). In particular, references by a top government official to an alcohol tax (a proven strategy for reducing the poor health and safety outcomes related to alcohol) suggest that the government is open to considering proposals that go beyond current law (O'Dowd, 2009). Perhaps this information also carries with it the hope that the work of Humphreys and Eisner (2010) and others on the Act will result in a careful review of the evidence, a full consideration of implementation, and a more comprehensive approach to alcohol policy that will achieve a safer nighttime environment and a healthier population overall.

References

Anderson, Peter, Dan Chisholm, and Daniela C. Fuhr. 2009. Effectiveness and cost-effectiveness of policies and programmes to reduce the harm caused by alcohol. *Lancet*, 373: 2234–2246.

Cook, Philip J. 2007. *Paying the Tab: The Costs and Benefits of Alcohol Control*. Princeton, NJ: Princeton University Press.

Duailibi, Sergio, William Ponicki, Joel Grube, Ilana Pinsky, Rolando Laranjeira, and Martin Raw. 2007. The effect of restricting opening hours on alcohol-related violence. *American Journal of Public Health,* 97: 2276–2280.

Frattaroli, Shannon and Stephen P. Teret. 2006. Understanding and informing policy implementation: A case study of the domestic violence provisions of the Maryland Gun Violence Act. *Evaluation Review,* 30: 347–360.

Goggin, Malcolm L. 1987. *Policy Design and the Politics of Implementation: The Case of Child Health Care in the American States*. Knoxville: University of Tennessee Press.

Humphreys, David K. and Manuel P. Eisner. 2010. Evaluating a natural experiment in alcohol policy: The Licensing Act (2003) and the requirement for attention to implementation. *Criminology & Public Policy*. This issue.

Hough, Mike, Gillian Hunter, Jessica Jacobson, and Stefano Cossalter. 2008. *The Impact of the Licensing Act of 2003 on Levels of Crime and Disorder: An Evaluation*. London, UK: Home Office.

O'Dowd, Adrian. 2009. U.K. government may consider minimum price for alcohol. *BMJ*, 339: b2929.

Plant, Emma J. and Martin Plant. 2005. A "leap in the dark?" Lessons for the United Kingdom from past extensions of bar opening hours. *International Journal of Drug Policy*, 16: 363–368.

09012-CrimJournal-Guts.indd 82 12/21/09 7:14:08 PM

Pressman, Jeffrey L. and Aaron Wildavsky. 1984. *Implementation: How Great Expectations in Washington Are Dashed in Oakland; Or, Why It's Amazing that Federal Programs Work at All, This Being a Saga of the Economic Development Administration as Told by Two Sympathetic Observers Who Seek to Build Morals on a Foundation of Ruined Hopes*. Berkeley: University of California Press.

Toomey, Traci L. and Alexander C. Wagenaar. 1999. Policy option for prevention: The case of alcohol. *Journal of Public Health Policy*, 20: 192–213.

Shannon Frattaroli, Ph.D., M.P.H., is an assistant professor in the Department of Health Policy and Management at The Johns Hopkins Bloomberg School of Public Health where she is affiliated with the Center for Injury Research and Policy. Frattaroli's research interests include injury prevention policy, effective strategies for translating research into policy and practice, and policy implementation.

09012-CrimJournal-Guts.indd 83 12/21/09 7:14:08 PM

EDITORIAL INTRODUCTION

LEGITIMACY PERCEPTIONS IN CORRECTIONS

Boot camps redux

What can they tell us about correctional legitimacy?

Susan Turner, *Senior Editor*
University of California, Irvine

When we hear the words "boot camps," many of us recall the great fanfare with which this model burst on the correctional landscape and was lauded by politicians, who recalled their own transformations in military boot camps. The boot-camp model, with its emphasis on strict discipline and physical training, blossomed across the country only to lose its luster when evaluations began to show that, despite positive changes in offender attitudes, no evidence showed that boot camps were not effective in reducing recidivism for juveniles or adults (MacKenzie 2006; Wilson, MacKenzie, and Mitchell, 2003).

Despite disappointing recidivism outcomes, boot camps could allow us to learn about different aspects of the incarceration experience. In Franke, Bierie, and MacKenzie's (2010, this issue) article, a boot camp was used as a laboratory to understand the role legitimacy plays in the correctional system (as opposed to policing and courts). They asked the following two main questions: Do inmate attitudes toward the justice system change during the course of confinement, and if so, what causes it? Perceptions of legitimacy in prisons are different from those in courts and of the police because, as Franke et al. stated, contact is longer in prison, and interactions are diverse and wider ranging.

At first blush, one might think that the boot-camp structure could be delegitimizing with a focus on corporal punishment within an authoritarian environment. At the same time, however, boot camps also have shown improvement in prosocial attitudes, which suggests the opposite might be true. To answer their questions, Franke et al. (2010) took advantage of a rare randomized experiment in which inmates served identical length sentences in either a boot camp or a traditional prison; attitudes were measured the week before inmates entered their programs and again 6 months later. Legitimacy measures were selected from a larger survey along themes of trust, allegiance, and confidence in the system.

Direct correspondence to Susan Turner, Department of Criminology, Law and Society, University of Calilfornia, Irvine, 336 Social Ecology, Irvine, CA 92697 (e-mail: sfturner@uci.edu).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 85 12/21/09 7:14:09 PM

Results showed that inmates who left the boot camp had higher measures of perceived legitimacy than those who left the regular prison, although the boot-camp perceptions had not increased from pre- to post-prison experience. Boot-camp participants perceived more positive than negative experiences. Negative experiences worsened perceptions of legitimacy, and positive attitudes improved attitudes toward the justice system. Importantly, when facility type and experiences were tested simultaneously in the same regression model, the effect for facility disappeared—providing support for the hypothesis that it is the types of experiences and not the facility per se that engender feelings of legitimacy.

Unfortunately, the available quantitative data did not allow Franke et al. (2010) to test explicitly concepts of procedural and distributive justice—concepts that have been linked to legitimacy. However, Franke et al. provided qualitative insights into the differences between the two environments that suggest these concepts are at play.

In one policy essay, Tyler (2010, this issue) focused on procedural justice in incarceration. He was concerned with why legitimacy did not increase in either of the institutions studied. He noted that research with police has shown that legitimacy can be increased—can this also be the case with incarceration? The findings of Franke et al. (2010) suggest that it can happen, but because procedural justice was not measured directly, their findings might reflect other differences, such as program offerings, staff conduct to create safe prison environments, and so on, which also can contribute to feelings of legitimacy. Taking a larger policy view, Tyler suggested that the policy of widespread prison use is, itself, questionable. However, he recognized that to the degree that prisons are needed, procedures should be in place that are fair and treat people with dignity, courtesy, and respect.

Kurlychek's (2010, this issue) policy essay "Transforming attitudinal change into behavioral change" asked what the characteristics of boot camps are that lead to positive attitude changes, and why it is that positive changes often do not translate into behavioral change once offenders return to the community? The positive experiences in boot camps might not result from anything to do with the military model but rather from the provision of services, a more attractive environment, lower risk offenders, and little idle time to share criminogenic attitudes and behavior. Although Franke et al. (2010) did not address the recidivism issue in their article, Kurlycheck predicted that the boot-camp sample would not show any differences from the prison sample (actually, MacKenzie, Bierie, and Mitchell, 2007, found marginal differences in recidivism, which favored the boot camp) because the skills learned in boot camp do not translate into success on the street, and offenders face structural barriers after release, such as poor employment opportunities and a lack of housing, access to health care, and so on, which hinder successful integration. Her point was that the inability to change offenders is not what causes rehabilitation programs to fail; they fail because we still do not do a good job in prisoner reentry. Successful reentry requires a better understanding of matching offenders to services as well as reforming social policies that block successful reentry.

09012-CrimJournal-Guts.indd 86 12/21/09 7:14:09 PM

Both policy essays placed boot camps in a larger policy context. For Tyler (2010), it is our overreliance on incarceration; for Kurlycheck (2010), it is impediments to successful offender reintegration. Each essay recommended major policy changes that would affect individual offenders as well as the larger community where offenders return after incarceration. Although I recognize that a focus on fairness and resulting legitimacy is necessary for all components of the legal system—from police, to courts, to corrections—I also must realize that reintegrating prisoners into society brings a myriad of challenges in addition to perceptual improvements.

References

Franke, Derrick, David Biere, and Doris L. MacKenzie. 2010. Legitimacy in corrections: A randomized experiment comparing a boot camp with a prison. *Criminology & Public Policy*. This issue.

Kurlycheck, Megan C. 2010. Transforming attitudinal change into behavioral change. *Criminology & Public Policy*. This issue.

MacKenzie, Doris L. 2006. *What Works in Corrections: Reducing the Criminal Activities of Offenders and Delinquents*. Cambridge, UK: Cambridge University Press.

MacKenzie, Doris L., David Bierie, and Ojmarrh Mitchell. 2007. An experimental study of a therapeutic boot camp: Impact on impulses, attitudes, and recidivism. *Journal of Experimental Criminology*, 3: 221–246.

Tyler, Tom R. 2010. Legitimacy in corrections: Policy implications. *Criminology & Public Policy*. This issue.

Wilson, David B., Doris L. MacKenzie, and Fawn Ngo Mitchell. 2003. *Effects of Correctional Boot Camps on Offending*. Oslo, Norway: The Campbell Collaboration.

Susan Turner is a professor in the Department of Criminology, Law, and Society and the Director of the Center for Evidence-Based Corrections at the University of California, Irvine (UCI). Turner led a variety of research projects while she was a Senior Behavioral Scientist at RAND, which included studies on racial disparity, field experiments of private-sector alternatives for serious juvenile offenders, work release, day fines, and a 14-site evaluation of intensive supervision probation. Her areas of expertise include the design and implementation of randomized field experiments as well as research collaborations with state and local justice agencies. At UCI, Turner currently is assisting the California Department of Corrections in the development and testing of a risk assessment tool as well as an evaluation of a parole violation decision-making instrument designed to provide an orderly decision-making process for responses to violations of parole.

09012-CrimJournal-Guts.indd 87 12/21/09 7:14:09 PM

RESEARCH ARTICLE

LEGITIMACY PERCEPTIONS IN CORRECTIONS

Legitimacy in corrections

A randomized experiment comparing a boot camp with a prison

Derrick Franke
University of Maryland, College Park

David Bierie
Federal Bureau of Prisons

Doris Layton MacKenzie
Pennsylvania State University

Research Summary

To extend research on legitimacy to the correctional system, we study a sample of 202 adult inmates randomly assigned to serve their 6-month sentence at one of two institutions—a traditional prison or a military-style correctional boot camp. Findings show that perceptions of justice system legitimacy changed during the course of incarceration, that the prison (but not the boot camp) proved delegitimizing, and that certain regime characteristics explained why.

Policy Implications

Across academic disciplines, studies continue to link compliance with perceived legitimacy. Compliance with the law, for instance, is related closely to the legitimacy of the justice

This project was supported, in part, by 2003-DB-BX-004, awarded by the U.S. Department of Justice. The Assistant Attorney General, Office of Justice Programs, coordinates the activities of the program offices and bureaus. Funding was provided by the State of Maryland, Governor's Office of Crime Control and Prevention, BYRN-2003-1063. Points of view or opinions contained within this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice or any state agency. The authors would like to thank Brian Johnson, Doug Smith, and the graduate student research group at the University of Maryland for their valuable comments on earlier drafts. Direct correspondence to Derrick Franke, 2220 LeFrak Hall, University of Maryland, College Park, College Park, MD 20742 (e-mail: dfranke@crim.umd.edu); Davide Berie, 14 Laurel Hill Rd. Unit H, Greenbelt, MD 20770 (e-mail: dbierie@bop.gov); Doris L. MacKenzie, Justice Center for Research, 327 Pond Building, The Pennsylvania State University, University Park, PA 16802 (e-mail: dlm69@psu.edu).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 89 12/21/09 7:14:09 PM

system and its actors. These findings suggest implementing legitimacy-building policies such as procedurally fair treatment and decision making by police officers and judges. This article, by finding legitimacy to be malleable even at the final stage of the justice process, proposes the efficacy of similar policies in the correctional system. As research from England and Wales has shown, legitimizing strategies in this context could increase compliance both during and after incarceration.

Keywords

legitimacy, prison, boot camp, procedural justice, randomized experiments

Perceptions of legitimacy can play an important role in the effectiveness of the criminal justice system. As confidence in the institution wanes and cynicism grows, practitioners find it harder to ensure the effective rule of law. Citizens are less likely to support, cooperate with, and obey authorities who lack legitimacy in their eyes (Tyler, 1990). Recent justice system initiatives have emerged to address this legitimacy problem. Community policing, for instance, encourages public support and cooperation by changing the role of police officers. When police work with community members to solve problems, it is thought that officers will be viewed as more trustworthy and their authority perceived as more legitimate. Similarly, restorative justice involving all stakeholders affected by crime has sought to provide a more legitimating alternative to the court process.

Consistently ignored in these discussions, however, is the role that legitimacy plays in the *correctional* system—or put differently, the role corrections plays in molding perceptions of legitimacy. With a few important exceptions comparing prison regimes in Great Britain (Bottoms, 1999; Liebling and Arnold, 2004; Sparks and Bottoms, 1995), little is known about legitimacy at the final stage of the justice system. For instance, we do not know whether the experience of being incarcerated influences perceptions of authority and whether these attitudes extend beyond the prison staff and regime to reflect the entire justice system. Several theoretical arguments assert that qualities of prison regimes, such as the nature of inmate–guard interactions, can change attitudes (Clemmer, 1940; DiIulio, 1987; Goffman, 1961; Sykes, 1958). Yet little empirical work has tested these claims, especially in relation to the concept of perceived legitimacy versus the more general concept of antisocial attitudes.

To address these gaps, this article focuses on two main questions. First, do inmates' attitudes toward the justice system change during the course of their confinement? Are there circumstances in which inmates can leave confinement with *more* respect for the law and legal authorities than when they entered? Or, if the correctional system is merely a depository for individuals who were acted on by the police and court system, are perceptions of legitimacy relatively fixed by the time an offender enters prison? Second, if perceptions do change, then what explains the change? Also examined is whether these outcomes vary by punishment regime. In particular,

09012-CrimJournal-Guts.indd 90 12/21/09 7:14:09 PM

what role might alternatives to traditional prison, such as correctional boot camps, play in affecting perceptions of legitimacy?

To answer these questions, we analyzed self-report data collected from a recent experiment in which inmates were randomly assigned to serve a 6-month sentence at either a traditional prison or a military-style correctional boot camp for adults. The purposes were (a) to show whether either (or both) facilities influenced perceptions of justice system legitimacy among inmates sent there, and (b) to explain why. As such, this study was not an attempt to show whether prisons or boot camps are inherently better than one another. Instead, the article aimed to show whether perceptions of legitimacy are malleable by the correctional system and what regime qualities (common to any form of incarceration) might account for such changes. We conclude with suggestions of policies that can be implemented in both traditional and alternative institutions.

What Is Legitimacy?

Drawing on Weber's (1968) view of authority, Tyler (2003: 308) defined legitimacy as "a quality possessed by an authority, a law, or an institution that leads others to feel obligated to obey its decisions and directives." Applying this definition to the justice system, legitimacy refers to the extent to which people feel they ought to obey the law and its authorities. The more legitimate people view the institution, the more obligated they will feel to obey it. That is, perceiving the justice system (and its actors) as legitimate leads people to "buy in" to the decisions of legal authorities (Thibaut and Walker, 1975). Tyler (1990: 21) noted that it is too simplistic to assume that people obey rules mainly out of the fear of consequences; if this were true, then social regulation would be a "much simpler and more easily calculable matter than in fact it is—one would simply need sufficiently powerful and certain forms of sanction." Instead, compliance with the law and cooperation with authorities hinges in large part on the voluntary acceptance of their rules and actions—in other words, on their perceived legitimacy.

In this sense, legitimacy has less to do with personal morality (e.g., issues of right and wrong) and more to do with perceptions of the authority, institution, or law as being deserving of compliance (Hoffman, 1977). As Tyler (2003: 309) explained, "this feeling of responsibility reflects a willingness to suspend personal considerations of self-interest and to ignore personal moral values because a person thinks that an authority or a rule is entitled to determine appropriate behavior within a given situation or situations."

Legitimacy and the Criminal Justice System

Mounting empirical literature suggests that how people view the criminal justice system has important behavioral consequences. In general, citizens are less likely to support the justice system as their faith in the institution declines. When law enforcement is perceived as illegitimate, the public becomes less likely to cooperate with the police by way of failing to report crimes, assist investigations, or serve as witnesses, and less likely to offer political support because of the belief that the police need less instead of more discretionary power (Tyler, 2003). Resistance

09012-CrimJournal-Guts.indd 91 12/21/09 7:14:09 PM

to "snitching" (i.e., exchanging incriminating information for reward) and often the retaliatory violence associated with the practice also has been linked to illegitimately perceived police practices (Rosenfeld et al., 2003).[1]

Consequences extend beyond general cooperation with legal authorities. More directly, when legitimacy is low, citizens are less likely to obey police orders (Skogan and Frydl, 2003) or accept court decisions (Tyler, 2000) and are more likely to react defiantly to sanctions (Sherman, 1993). Self-reported lawbreaking (Tyler, 1990), recidivism (Paternoster et al., 1997), and even national crime rates (LaFree, 1998) have been linked to confidence and trust in the justice system and its actors. In fact, Tyler (1990) found perceived legitimacy of the law to be *more* influential in determining offending behavior than the perceived risk of being caught and punished.

From studies examining interactions with the police and court system, we have learned that attributions of legitimacy often result from the way in which authorities perform their job. Did the police officer explain his decision and allow the suspect to tell his side of the story? Did the judge take all of the evidence into account before making her decision? These aspects—which reflect issues of fair treatment and the quality of decision making—represent what has been termed procedural justice. Tyler's (1990) theory claims that issues of procedure as opposed to outcome bear most heavily on citizens' perceptions of legitimacy.

In support of this perspective, Paternoster et al. (1997) found that domestic violence suspects who perceived their treatment by police to be fair were more likely to accept the officer's decision and less likely to commit a subsequent spousal assault, regardless of the outcome (arrest or warning). Similarly, defendants who reported being treated fairly by the courts tended to be more satisfied with and to accept the outcome more often than individuals who thought they were treated unfairly, even when the latter group receives more favorable sentences (Benesh and Howell, 2001; Tyler, 1984). In short, legitimacy has more to do with the way decisions are made and how people are treated, and less to do with the decision itself (Lind and Tyler, 1988; Thibaut and Walker, 1975; Tyler, 1990). As a result, when authorities and institutions behave in a manner that is perceived of as fair and enact policies that come across as just, people become more loyal to the authority, supportive of its directives, and respectful of the law (Tyler, 1990).

Legitimacy and Prisons

Unfortunately, our knowledge of citizen assessments of justice system legitimacy stems almost exclusively from studies examining interactions with the police and courts. Far less is known about whether the correctional system influences perceived legitimacy. Prison experiences—characterized by routine and prolonged interactions with authorities—might provide an even

1. These issues have been pronounced in Baltimore City particularly, the location from which most of our data derive. For example, after a video surfaced in 2004 featuring a prominent basketball star threatening to kill informants, a Baltimore official commented that "threatening snitches has become mainstream—so much that they made a DVD joking about it" (Brown, 2007: 172). Near the same time, the city recognized a growing problem with juries refusing to convict defendants, which prompted an advertising campaign in which the slogan "Convict the Guilty if you're on a Jury" was emblazoned on city busses (Linder and Associates, 2002: 16). Both trends seem tied to the declining legitimacy of the justice system.

09012-CrimJournal-Guts.indd 92 12/21/09 7:14:09 PM

richer context to study the effects of legitimacy than police and court experiences, which are by comparison less frequent and shorter.

We do know that confinement experiences can have lasting effects. Laub and Sampson (2003) discussed the role of positive and negative experiences in early juvenile reformatory schools. Among a sample of delinquent boys followed to age 70, some looked back on their lives and recalled positive experiences while incarcerated. One subject referred to spending time at the Lyman School for Boys as "great" because "they taught you a lot of things there.… I still go back to that whenever I think of a change in my life" (p. 123). Others, though, spoke of entirely different experiences remembering instead brutality and unfairness. Referring to his time spent at a different juvenile institution, another subject placed particular blame on staff mistreatment, commenting "If you looked at one of them cross-eyed, he'd give you a backhander … there was a lot of physical abuse" (p. 218). In the latter group, these negative experiences often led to "cynicism about the world, particularly the criminal justice system" (p. 219), and these attitudes persisted decades later.

According to some, prisons are generally delegitimizing (DiIulio, 1987) and often promote criminality (Bonta and Gendreau, 1990). In their review of inmate behavioral studies, Bukstel and Kilmann (1980: 472) found consistent evidence of "overwhelming positive reinforcement" by other inmates and even staff for antisocial behaviors and attitudes. Although the belief that prisons are in essence "schools for crime" has found support, more recent research has identified important circumstances whereby regimes can lead inmates to recognize the legitimate authority of staff and "buy in" to the program. Procedural fairness in addition to efficiency of services (e.g., programming), distributive fairness (e.g., consistent outcomes), and humane regime qualities (e.g., visitation and time out of cell) affects attitudes (Sparks and Bottoms, 1995; Sparks, Bottoms, and Hay, 1996). Aspects of the prison environment can matter as well, such as the level of situational control, safety, structural integrity of facilities, and the nature of inmate interactions (Bottoms, 1999; Sparks et al., 1996).

Attitudes also can change according to the type of offender, although results are mixed. Although some evidence indicates that high-risk inmates cope least well in prison (Zamble and Porporino, 1990), other research suggests the opposite (Leschied and Gendreau, 1994). Emotions also have been linked to the phase of incarceration. For instance, many inmates enter confinement with heightened emotional states and are eager to participate in rehabilitative programs (Henderson, 2000), but they become disillusioned, less involved, and often more hostile through time (Paulus and Dzindolet, 1993; Zamble and Porporino, 1988). Other inmates are initially anxious but eventually gain a sense of control and balance (Ruback, Carr, and Hopper, 1986).

Incarceration, therefore, involves a wider range of experiences, more diverse interactions, and a longer duration than contact with police or courts. As a result, more potential avenues are available through which attitudes are affected that include perceptions of justice system legitimacy. A procedurally just regime, for instance, still might fail to ensure order if other regime qualities lack legitimacy.

09012-CrimJournal-Guts.indd 93 12/21/09 7:14:10 PM

It is important, though, not to confuse legitimacy with coddling or pleasing inmates. Rather than providing luxuries or considering outlandish requests, legitimacy is about "maintaining the standards and expectations of fair and humane treatment now taken for granted in the world beyond the wall" (Sparks and Bottoms, 1995: 58). Illustrating this point, Sparks and Bottoms (1995: 60) argued:

> Every instance of brutality in prisons, every casual racist joke and demeaning remark, every ignored petition, every unwarranted bureaucratic delay, every inedible meal, every arbitrary decision to segregate or transfer without giving clear and well founded reasons, every petty miscarriage of justice, every futile and inactive period of time—is delegitimizing.

In terms of outcomes, research suggests that prison regimes that emphasize legitimacy-building characteristics can experience less disorder. The best evidence for this finding comes from research examining English prisons in the late 1980s. Tasked by the Home Office to explain the "nature of control problems [in long-term prisons] and the conditions leading to their emergence" (Sparks et al., 1996: 98), Sparks et al. (1996: 113) systematically evaluated two maximum-security prisons chosen from similar populations with "radically different regimes and control problems."[2] At the prison whose regime was characterized as less restrictive, less coercive, and more fair (procedurally and distributively), riots, interpersonal violence, inmate–staff violence, and violations in general were less prevalent (Bottoms, 1999; Sparks and Bottoms, 1995). Results indicated that prison administrations could effectively promote legitimacy and ultimately encourage rule compliance within their walls. Less is known about whether these perceptions extend beyond the prison to reflect global attitudes toward the justice system or whether attitudes change during the course of incarceration.

Legitimacy and Boot Camps

Similarly, we know little about the legitimacy of alternatives to traditional prison, such as correctional boot camps. These punishment regimes tend to combine high levels of situational control with harsh physical and psychological discipline. Military basic training provides the model for this environment, one that is characterized by both mentally and physically grueling tasks. In many respects, boot camps exert *more* situational control over day-to-day activities of inmates than prisons. Shameful remarks, heightened scrutiny of behavior, stressful situations, and summary punishment (i.e., on-the-spot sanctions with no institutional review or appeal process) are commonplace among correctional boot camps and are signs of enhanced control in an authoritarian setting. Because research has found these characteristics to be delegitimizing (Bottoms, 1999; Sparks et al., 1996), it might seem that boot camps adhere to some of the least

2. Although similar in most demographic respects, the authors noted two key differences between the facilities: A higher proportion of both younger prisoners and individuals convicted of an armed crime comprised the facility that experienced more disorder. The higher risk population at one facility, the authors admitted, might have affected their results.

09012-CrimJournal-Guts.indd 94 12/21/09 7:14:10 PM

legitimate aspects of prisons.

However, equally compelling arguments could predict the opposite outcome. For instance, the increased structure and constraints boot camps place on their inmates might not be negative at all. In fact, evidence suggests that boot camp inmates often form strong emotional bonds with staff as a result of the nature of their interaction. Despite the strict rules and discipline, drill instructors are commonly perceived by inmates as caring genuinely about their welfare and improvement (MacKenzie, Wilson, Armstrong, and Gover, 2001; Styve, MacKenzie, Gover, and Mitchell, 2000), which is a position not commonly attributed to prison guards. Despite heightened levels of physical and mental stress, boot camp inmates often perceive the stress to be constructive. Perhaps as a result, studies have documented that boot camps tend to create less oppositional and more prosocial attitudes among inmates during the course of the program, whereas inmates in traditional prisons typically stay the same or become worse in terms of these attitudes (MacKenzie and Souryal, 1995; MacKenzie et al., 2001).

Finally, completing difficult tasks in boot camps might create feelings of accomplishment. Improvements in self-esteem and self-discipline, in turn, often coincide with a new respect for authority (MacKenzie and Souryal, 1995), which suggests that boot camps might actually promote legitimacy. In short, enhanced situational control might not be delegitimizing if implemented fairly, consistently, with explanation, and—perhaps most important—by an official who seems to care about the future of the inmate. Thus, it might be less a matter of control or constraint itself and more a matter of the way in which that control is implemented that predicts legitimate views of authority.

Gaps in the Current Research

These predictions, however, have not been tested. In fact, several gaps exist in both prison and boot camp research. Regarding prison research, we do not know whether perceptions of legitimacy change during the course of incarceration, nor do we know the extent to which differences between prisons are caused by regime qualities or characteristics of the inmates themselves. A reliance on cross-sectional data and nonexperimental designs has limited the ability to address these issues.

Similarly, few boot camp studies have moved beyond evaluations based solely on success or failure. Among those studies that do examine intermediate outcomes, it remains unclear whether participants alter their overall assessment of the justice system. Finally, we know little about *why* incarceration might affect these attitudes, be it procedural justice or other experiences with the prison regime.

These gaps are important to fill, especially at a time when record numbers of inmates are leaving confinement and returning to the community. As the final stage before reentry, prisons should concern themselves with issues of legitimacy and guard against the possibility of exacerbating conditions. To address these gaps, this article presents results from a study that randomly assigned inmates to serve 6-month sentences at one of two different punishment regimes—different along lines thought to impact perceptions of legitimacy. By measuring

09012-CrimJournal-Guts.indd 95 12/21/09 7:14:10 PM

perceptions entering and exiting each facility, we isolate the causal impact of facility type on changes in perceived legitimacy and attempt to identify specific aspects of the incarceration experience that explain the attitudinal change. As such, this article advances current research on legitimacy by (a) analyzing a rare randomized experiment among an inmate sample, (b) examining change across time, and (c) assessing the mediating effects that confinement experiences exert on attitudinal change.

Hypotheses

1. Facility type will predict change in perceived legitimacy: Although attitudes will decline for both prisoners and boot camp participants, the decline will be significantly greater for prisoners. Scholars have argued that incarceration tends to be a delegitimizing experience (DiIulio, 1987; Scraton, Sim, and Skidmore, 1991; Sparks and Bottoms, 1995). Nonetheless, Sparks and Bottoms (1995) found numerous conditions under which correctional regimes actually promote legitimacy. Correctional boot camps likely emphasize many of these conditions (including elements of procedural fairness) more so than prisons, which possibly accounts for why studies tend to show more positive attitudinal changes at boot camps. Therefore, although legitimacy perceptions likely will worsen for each group (because inmates of any sort are not likely to *gain* respect for an institution that arrested, convicted, and incarcerated them), prisoners should experience a significantly steeper attitudinal decline than boot camp participants.

2. Positive experiences will be more prevalent at the boot camp, whereas negative experiences and environmental deprivation will be more common at the prison. Studies identifying attitudinal benefits associated with boot camps usually trace the positive outcomes to specific experiences associated with this correctional style (MacKenzie et al., 2001; MacKenzie and Souryal, 1995; Styve et al., 2000). Better inmate–guard relationships, less hostile interaction among inmates, more regularity and consistency of programming, more safety, and fairer procedures are among the qualities that separate modern-day correctional style boot camps from prisons. Therefore, the boot camp should fare better than the prison across procedural, distributive, environmental, and interpersonal lines.

3. Incarceration experiences will affect perceptions of legitimacy such that negative experiences and environmental deprivation will worsen attitudes and positive experiences will improve attitudes. Third, we examine specific aspects of incarceration that affect perceptions of legitimacy. Prior studies have outlined certain routines, environmental factors, and regime characteristics that influence whether inmates believe the staff are exercising their authority in a legitimate manner. Hypothesis 3 predicts that negative experiences (e.g., allowing fights between inmates) and environmental deprivation (e.g., lack of privacy) will be delegitimizing, whereas positive experiences (e.g., learning from academic classes) will improve legitimacy perceptions.

4. Incarceration experiences will mediate the effect of facility type on changes in perceived legitimacy. The final hypothesis predicts that incarceration experiences will explain why the facility caused changes in perceived legitimacy. Again, the basis for this prediction is evidence

09012-CrimJournal-Guts.indd 96 12/21/09 7:14:10 PM

suggesting that micro-level regime characteristics (not the overall correctional style itself) influence assessments of legitimacy (Bottoms, 1999; Sparks and Bottoms, 1995). In short, the facility variable should have no significant effect on legitimacy perceptions once experiential controls are included in the model.

Methods

Sample and Procedure

This research used data previously collected for the purpose of evaluating Maryland's Toulson Correctional Boot Camp for Adults (MacKenzie, Bierie, and Mitchell, 2007; MacKenzie et al., 2004). Offenders whose original sentences ranged from 1 to 5 years became eligible for the experiment by choosing to accept a contract with the state, which reduced their term to 6 months if they successfully completed the program (none refused). Between January 2002 and January 2004, a cohort of study participants arrived at the boot camp each month. Randomization occurred on a cohort-by-cohort basis before each group arrived.[3] While at the boot camp (separated from other inmates and staff), all participants completed the pretest survey. They were then informed which facility they would be serving their sentence and, on the same day, the control group was sent by bus to the prison.[4] Overall, 234 male inmates were randomly assigned—111 to the military-style correctional boot camp and 123 to the mixed-security traditional prison.

At each facility, subjects were given priority for access to life-skills training, education, and drug-treatment programming. In the event of misconduct, each facility allowed inmates to be removed from their program (i.e., to lose their 6-month contract). A total of 19 subjects at the boot camp and 16 at the prison lost their contract for disciplinary reasons.[5] In addition, 13 boot camp participants and 11 prison inmates were released early by either court order or successful appeal. Neither of these differences was statistically different between facilities. Regardless of

3. A random number generator was used to place participants into their respective facility with the possibility that an unequal number of inmates would end up at each.

4. Interestingly, most offenders (by show of hands) wished to be sent to the traditional prison. The general perception among inmates was that the boot camp would be harder. Also, the prison was geographically closer to most of the offenders' families.

5. Sanctions for these 35 inmates took two forms. Some participants were fully expelled (i.e., shipped to another traditional prison to serve the remainder of their original sentence). Others received minor sanctions (i.e., served an additional 30 or 60 days at the control or treatment facility). Although the prison was more likely to use the minor sanction option, overall time served was not significantly different between facilities (208 days at the boot camp, 196 at the prison). Additional analyses excluding these participants produced no substantive changes in terms of direction, magnitude, or result significance.

09012-CrimJournal-Guts.indd 97 12/21/09 7:14:10 PM

completion status, all inmates were kept in the study and analyzed as randomized.[6] All subjects who left (by way of removal via court order or appeal) and could later be located were asked to respond to questions in reference to the facility where they originally had been assigned. Overall, 202 (86%) subjects provided adequate time 1 and time 2 data to be included in the analysis—96 from the boot camp and 106 from the prison.[7]

Observational Data

While collecting data for the original study (MacKenzie et al., 2004), members of the research team conducted more than 100 visits to each facility. In addition to administering surveys, the researchers conducted unstructured interviews with inmates and staff as well as made systematic observations regarding facility structure, regime differences, and the quality of staff–inmate interactions. Although electronic recording devices were prohibited, the researchers' field notes illustrate some important distinctions between the two facilities being examined and, thus, provide some context for the results that follow (see MacKenzie et al., 2007; Bierie, 2009, for a more thorough facility description).

Boot camp. Located in a rural town outside of Baltimore, the boot camp averaged 325 inmates during the course of the study—well below its capacity of 450. Well-manicured lawns and flowerbeds, and a large military-type obstacle course, mark the grounds. Participants are housed in well-maintained dormitory-style units, but spend most of their time either outside or in one of the classroom buildings. Researchers noted a sense of quiet and calm within the housing wings and on the grounds, which is uncommon in correctional environments. Although inmates are allowed access to books, they are not allowed televisions, radios, or commissary and have only limited access to phones.

It is important to note that the particular boot camp in this study had been operating for slightly more than 10 years at the time of this research. Many boot camps initially began with a focus on discipline and retribution alone. As boot camps developed, staff and administrators realized that participants need additional programming if they are going to succeed after being released. These "second-generation" boot camps began to increase human services and rehabilitation programs within the daily schedule of activities. Falling into this category, the

6. In the course of the study, one assignment error occurred in which six inmates who should have remained at the boot camp were sent to the comparison prison after 1 month at the boot camp. These inmates were kept in the study and analyzed in the group they were assigned rather than the group they were sent to maintain the methodological integrity of the randomized design (Heckman and Smith, 1995). The six inmates, however, were prompted to respond to the questions in relation to their time at the boot camp (despite having only served 1 month there). Analyses excluding these six participants did not change the results.

7. Relevant demographic variables (i.e., age, race, and criminal history) did not predict absence from the study, which implies that the amount of error induced by missing data might be minor. Scales were considered "complete" if the inmate answered at least 75% of the items, whereby mean substitution was used for the remaining missing items. Of the 15 missing boot camp cases, 11 were because of incomplete surveys and 4 were early releasees who could not be located. Of the 17 missing prison cases, these numbers were 14 and 3, respectively.

09012-CrimJournal-Guts.indd 98 12/21/09 7:14:10 PM

boot camp examined here explicitly focuses on drug treatment, attitude change, and education. Along with exercise and drill, the daily schedule includes a significant amount of time for therapeutic activities.

The highly structured boot camp regime affords inmates little idle time for casual interaction with other inmates. From the day they arrive, inmates are enrolled in therapeutic treatment programs in which they must remain for their entire stay. Waking at dawn, platoons rotate among exercise, drill, and programming (e.g., education, substance abuse treatment, and life-skills training). At least one drill instructor is always with the group—even in classrooms—to ensure that inmates are compliant with facility rules. Meals also are highly controlled; inmates have limited time to eat their food, talking is prohibited, and faces must be forward at all times. Each day concludes with a platoon-run community meeting where inmates can voice complaints or congratulate one another on an achievement. Researchers noted a rare sense of prosocial comradery, particularly during these meetings, compared with traditional institutions. A successfully passed General Educational Development (GED) examination or completed program often was met with applause from fellow inmates.

An average of 100 staff members was employed at the boot camp during the study.[8] By design, staff maintained highly structured interactions with participants. Observers noted inmates instinctively standing at attention when drill instructors or staff members entered a room and shouting in unison "good morning sir [ma'am]." They did not speak unless spoken to by the drill instructor and stood at attention until dismissed. If an inmate failed to stand at attention or committed any other minor rule infractions, then drill instructors handed down an immediate summary punishment. Instructors had the option of correcting the inmate with a warning, a physical exercise such as pushups, or a one-page essay on the rule violation. A more severe punishment option was ordering the inmate to the "decision bag," which was a 20-lb bag of sand with the word "decision" written on it. The inmate was told to carry the bag around the facility jogging track (1 mile) and to think about his misbehavior, why it was wrong, and what he was going to do to ensure that it did not happen again. Many times, the entire platoon would be forced to suffer the punishment alongside the individual at fault.

The nature of staff–inmate interactions was particularly noteworthy because staff seemed to implement physically demanding and sometimes humiliating punishments in a procedurally fair manner. They treated the inmates equally and followed facility rules "by the book." Inmates certainly bemoaned the distributive qualities of punishment (i.e., 20 pushups for slouching in class seemed a bit harsh). Yet, inmates reported knowing what to expect from day to day. Researchers heard few complaints regarding staff bias (which was also rare in corrections), and they noted high levels of consistency in treatment for the length of the study and between drill instructors.

8. Boot camp staff tended to receive more training than staff at other Maryland Department of Corrections facilities. Along with prior military experience (which many had) and standard Department of Corrections training, drill instructors hired at the boot camp were sent to Marine Corps Base Quantico for specific training on how to run platoons.

09012-CrimJournal-Guts.indd 99 12/21/09 7:14:11 PM

However, one element of procedural fairness seemed lacking at the boot camp: Inmates had few avenues by which to have their voice heard before or after being sanctioned—a concept Tyler (1990) referred to as "representation." When a rule was broken at the boot camp, drill instructors effectively served as judge, jury, and executioner. Punishments were enforced swiftly with little regard to an inmate's explanation. Unsolicited explanations often, in fact, brought more sanctioning (recall the policy of not speaking unless spoken to).

However, this lack of voice did not seem to cause the level of frustration one might expect. As mentioned, inmates knew what to expect, but an equally important element surfaced as well: Drill instructors took time to explain to inmates what rule they violated and why they were being punished in this way, which ensured that the inmate understood exactly why he was being sanctioned. On one occasion, researchers observed an inmate performing pushups. Hovering above, the drill instructor explained, "Remember, I'm doing this to make you a better father, a better employee, and a better person. If I don't instill self-discipline in you, then I will have failed you—and I'm not going to fail you." These explanations were common at the boot camp and seemed to imply a genuine concern or stake in each inmate's future.

Traditional prison. Located in Baltimore City, the traditional prison is one of the oldest operating prisons in the United States (built in the mid 1800s). Although designed for 1,200 inmates, the prison was overcrowded during the study and averaged more than 1,500. Unlike the well-manicured landscaping at the boot camp, the research team observed no grass on prison grounds, as well as abandoned buildings, cars, and trash lined the adjacent streets. The facility itself seemed dilapidated, with dirt, dust, and walls in need of paint. Inmates complained of temperature extremes and a pigeon infestation (both of which were confirmed by researchers' observations). Subjects also complained of broken phones, lousy food, too few microwaves, and loud televisions or radios. These final complaints were particularly interesting as phones, televisions, and microwaves (luxuries not available at the boot camp) still caused frustration among prisoners.[9]

Unlike the boot camp, access to programming was limited at the prison. The school, for instance, had only 80 seats available and demand for education always exceeded supply. Despite an agreement stating that study participants would receive the same therapeutic programming at each facility, this opportunity did not always occur. Participation rates at the prison were 31% for academic education, 64% for drug treatment, and 43% for cognitive skills training, compared with 100% for all three at the boot camp (see O'Neill, MacKenzie, and Bierie, 2007, for more information on programming differences). Partly because of this discrepancy, inmates at the prison had much more unstructured time.

9. Another source of frustration discovered later concerned release dates. Some cohorts thought they were being held longer than the promised 6 months. Indeed, our data confirmed this claim. Approximately 20% of the prison sample was held 1–4 weeks past their release date, an error the administration attributed to a shortage of parole officials.

09012-CrimJournal-Guts.indd 100 12/21/09 7:14:11 PM

Importantly, the prison in this study serves as a pre-release center. Its population consists of inmates who already have served most of their sentence (i.e., all but the final 2 years) at another institution. Although it resembles a typical maximum-security facility, the technical security level for this prison is "all." Ultimately, the prison houses a more heterogeneous (and greater risk) mixture of offenders than the boot camp, but did not segregate the population by risk. Study participants, therefore, interacted with more serious offenders than subjects at the boot camp and might have developed more oppositional attitudes as a result (Wellford, 1967).

With nearly 500 members during the course of the study, staff at the traditional prison outnumbered boot camp staff 5:1. However, interactions with staff were less frequent. Some subjects reported having not even met their case manager 3 months into their incarceration. Grievance procedures were more formal at the prison (i.e., a written complaint followed by a hearing), but inmates reported that their complaints often went unanswered. Similar formal procedures dictated sanctions, yet inmates reported inconsistencies across guards. Some guards were described as "by the book," but others were called vindictive, cruel, lazy, and incompetent. At any rate, prisoners did not trust their correctional authority figures the same as the subjects did at the boot camp. Although the researchers could not confirm staff mistreatment, they did conclude that an apathetic culture existed among some prison guards. During one interview, a custody officer commented, "It doesn't really matter what we do to help these guys. Most of them will just end up back here anyway."

In sum, our observations and interviews with inmates led us to believe that important qualitative differences existed between the facilities. From the structure, environment, and regime style, to interactions with other inmates and staff, the boot camp seemed to operate with more legitimating qualities than the prison. The most important issue, according to inmates, seemed to be staff treatment. Although harsh and confrontational at times, boot camp drill instructors were trusted. In contrast, prison inmates seemed to distrust their captors and not know from day to day how they would be treated. The result of these differences led researchers to conclude that boot camp participants were more likely to "buy in" to the regime. As buy in relates to legitimacy, observational data suggest that attitudes toward the legitimacy of the correctional system might decline at a steeper rate for individuals sent to the prison.

Quantitative Data

The primary data sources for this article are a pretest and a posttest survey. Self-report questionnaires were administered to the inmates prior to random assignment (the week before inmates began the program) and again 6 months later (the week prior to release).[10] The questionnaires provide data on personal and criminal histories, perceptions of confinement, as well as interme-

10. The 1-week lag between the exit survey and release helped ensure that survey results were not influenced by the often-jubilant graduation ceremonies at the boot camp. Although prior research has not focused on graduation effects, we find it unlikely that anticipation of the graduation ceremony alone would have much influence on attitudinal change. Graduation reflects the positive perspective of staff and administrators as a whole. In context, it is simply the culmination of an entire program and reflects this general positive perspective promoted throughout the boot camp.

09012-CrimJournal-Guts.indd 101 12/21/09 7:14:11 PM

diate change across several criminogenic qualities, including attitudes toward the justice system. The men in our sample came mostly from Baltimore City, with most (91% boot camp, 90% prison) serving time for a felony drug offense.[11] The average participant was African American (87%), 23 years of age (*SD* 4.03) with slightly more than five prior arrests (*SD* 4.08). Less than 10% was married, and typical of most inmate samples, education and employment history was notably deficient, with nearly two thirds having dropped out of high school and more than 15% never having held a job. As expected from randomization, no significant demographic or criminal history differences were present between the two groups.

Key Variables

Perceived legitimacy of the criminal justice system. Social scientists commonly use one of four approaches to measure legitimacy perceptions. The most direct extension involves items tapping the perceived obligation of an individual to obey or defer to authority. For example, Tyler's (1990) early study used an index of items measuring the obligation of Chicago residents to obey the law by asking them whether they agree with such statements as "People should obey the law even if they think it is wrong" and "Disobeying the law is seldom justified." Similarly, researchers have asked under what (if any) circumstances respondents would accept the decision of an authority even if the respondent felt the decision was unfair. In either scenario, the argument is made that legitimacy is achieved when citizens voluntarily obey these rules, laws, or decisions regardless of their subjective attitudes toward them.

Some conceptualize legitimacy in terms of a more general affect. This approach commonly involves gauging an institution on a continuum of positive or negative feelings. For instance, respondents might be asked to think about a thermometer and rate how warmly they feel about the police from 1 to 10, with 1 being cold and 10 being warm (Tyler and Huo, 2002). Research examining attitudes toward the U.S. Supreme Court (Gibson, Caldeira, and Spence, 2003) and Congress (Kimball and Patterson, 1997) have used similar measures.

Others argue that legitimacy reflects a lack of cynicism about the law and legal authorities (Ewick and Silbey, 1998; Sampson and Jeglum Bartusch, 1998; Tyler and Huo, 2002). Here, the idea is to establish the extent to which people believe that the law and legal authorities operate to protect their interests. Individuals who question the motives of police, for example, are in turn more likely to be cynical of law enforcement. Similarly, the legitimacy of the legal system suffers when its laws are perceived as inconsistent, biased, or discriminatory. In short, legitimacy grows as cynicism declines.

11. Most inmates were serving time for selling cocaine or heroin. In fact, inmates were eligible for this experiment only if they had never been convicted of a violent crime. However, in the piloting of self-report criminal offending calendars at the boot camp for an analogous study, many inmates reported serious and violent prior offending—primarily armed robberies—for which they formally had never been caught or convicted. Several inmates also had been charged previously (but not convicted) of serious violent offenses, which included homicide. Still, the mainly nonviolent composition of offenders should be noted when generalizing findings.

09012-CrimJournal-Guts.indd 102 12/21/09 7:14:11 PM

The most common approach, however, defines legitimacy in terms of support, allegiance, trust, or confidence in an authoritative body. Traditionally, political scientists have used this type of measure to examine the perceived legitimacy of political groups or governments (Easton and Dennis, 1969). More recently, psychologists and criminologists have adapted this approach to study the justice system (Sunshine and Tyler, 2003; Tyler, 1990, 2001). Respondents typically were asked whether they agreed with statements, such as "I have a great deal of confidence in the criminal justice system"; "I feel that I should support the police"; or "I trust the court system." Several items, therefore, have been used often to measure support, trust, and confidence (i.e., perceived legitimacy) in the justice system.

The measure used in this article reflects a blend of these approaches but most closely mirrors the last conceptualization. From the data, we identified 13 true-or-false items tapping into issues of trust, respect, and confidence in numerous justice system domains (see Table 1).[12] The questions asked about attitudes toward the institutions themselves (e.g., the overall criminal justice system, policing, the court system, and corrections) as well as the practitioners who constitute these branches (e.g., police officers, judges, and correctional officers). Items were combined to form a global measure of perceived justice system legitimacy that ranged from 0 (no legitimacy) to 13 (complete legitimacy). Items in our scale correlated highly among each other (coefficient alpha = .78).

Although not asking directly "Do you trust the justice system," items such as "Police and judges say one thing and do another," and "I have been dealt with unjustly by the police, court or prison," reflect distrust of the justice system and a lack of confidence in it as an authoritative body. Other items refer to the attitudes of respondents toward legal authorities and reflect cynicism (or lack thereof) toward them (e.g., "Most police will try to help you" and "Police usually treat you dirty"). "The criminal justice system sucks" reflects a more general affect toward the institution. Admittedly, this measure has not been subjected to rigorous validity tests. However, a good deal of research has examined directly measures of legitimacy. Gibson et al. (2003), for instance, found empirical support for conceptualizations that reflect confidence and general affect. In particular, they suggested three core items that provide a "reasonable beginning point for developing a scale of institutional commitment" (p. 363): whether to do away with the institution, whether to reduce its jurisdiction, and whether it can be trusted. Based on these criteria, the measure used in this article seems to have face validity.

12. The original evaluation was designed in part to measure attitudinal change. Among the 200+ questions were items from four criminogenic scales that have been shown to predict offending (see MacKenzie et al., 2007, for more information on these instruments). Our legitimacy measure represents a small factor-analyzed subset of two of these scales—the Jesness Inventory (Jesness, 1996) and the Self-Appraisal Questionnaire (Loza and Loza-Fanous, 2000).

09012-CrimJournal-Guts.indd 103 12/21/09 7:14:11 PM

TABLE 1

Legitimacy Items at Exit (Time 2)

Items	Range[b]	Boot Camp Mean	Boot Camp SD	Prison Mean	Prison SD	Mean Difference	SE Difference
Most police will try to help you	0–1	.38	.49	.32	.47	.06	.07
Prisons offer bogus programs[a]	0–1	.73	.45	.53	.50	.20**	.07
Police stick their noses in a lot of things that aren't their business[a]	0–1	.16	.37	.06	.24	.09*	.04
It is better to stick together than to call the policea	0–1	.26	.44	.24	.43	.02	.06
Police usually treat you dirty[a]	0–1	.34	.48	.22	.41	.13*	.06
Police and judges will say one thing and do another[a]	0–1	.29	.46	.27	.45	.02	.06
Cops and guards are pigs[a]	0–1	.60	.49	.52	.50	.08	.07
The criminal justice system sucks[a]	0–1	.27	.45	.19	.39	.08	.06
I am doing more time than I really deserve[a]	0–1	.64	.48	.43	.50	.21*	.07
I am doing time for bogus charges[a]	0–1	.73	.45	.65	.48	.08	.07
I have been dealt with unjustly by police, court, or prison[a]	0–1	.46	.50	.35	.48	.11	.07
I would shake hands with a correctional officer	0–1	.70	.46	.55	.50	.15*	.07

[a] Reverse coded.
[b] 0 = False; 1 = True.
*$p < .05$. **$p < .01$.

The surveys were administered to all inmates before random assignment (time 1) and again 6 months later before release (time 2).[13] Four legitimacy items had means of .60 or higher at time 1, which indicates that at least 60% of inmates reported a "legitimate" response, whereas each of the other 8 was rated at .40 or worse. The items scoring higher on legitimacy tended to be those factors concerned with the correctional system (e.g., guards and programming), whereas the lowest ratings reflected opinions of the police and the justice system as a whole.[14]

Random assignment again created groups that would be expected to vary only randomly. No significant differences in any of the legitimacy items were present at the time of the entry survey. The combined mean score on the legitimacy scale was 5.4 (out of 13), suggesting that inmates entered confinement with generally negative attitudes toward the system; nonetheless, most inmates recognized at least some legitimacy. For the most part, attitudes followed a normal

13. If three or fewer responses were missing, then we used the mean of the items that were answered to generate the mean value of the scale. If more than three items were missing, then we excluded the case. Thirty-two cases were excluded for this reason, which left a total of 202. Chronbach reliability alphas for the scales were .78 at time 1 and .76 at time 2.

14. This trend may be from sample characteristics. Although our participants averaged more than five arrests, they had never been incarcerated (as adults) prior to the experiment.

09012-CrimJournal-Guts.indd 104 12/21/09 7:14:12 PM

distribution. Although many inmates expressed little confidence and trust in the justice system (32% scored 3 or less), a surprising number reported the opposite (11% scored 9 or higher). Most inmates (49%) fell somewhere in the middle between 4 and 8.

Confinement experiences. As mentioned, studies have suggested that a relatively broad range of experiences should be examined when predicting a prison regime's legitimacy. Interactions with other inmates, relationships with staff, environmental characteristics, and distributive and procedural fairness all should be taken into account. Questions in the exit survey (time 2) were used to capture the perceptions of such experiences of inmates while incarcerated. A total of 16 "experience" items were grouped into one of three categories.[15] The first category reflects positive aspects of confinement and includes items such as "My experience here will help me get a job" and "The staff here helped me change." The remaining dimensions, however, reflect experiences that would more likely be delegitimizing for inmates. The negative experiences scale includes statements such as "I worried about my safety here" or "Inmates fight with other inmates here." The final scale reflects environmental deprivation. This construct measures other aspects of the facility that might have been problematic for inmates such as a lack of privacy or noise. We predict that these experiences will vary between facilities and perhaps explain attitudinal change.

Analytic Strategy

To test the first hypothesis, we conducted numerous *t* tests that compared mean levels of perceived legitimacy, and we examined whether scores were the same at time 1, whether the attitude of boot camp participants and prisoners became significantly worse (or better) throughout their incarceration, and whether the difference at time 2 between facilities was significant.[16] The next stage employed *t* tests to determine whether perceptions of positive, negative, and environmentally depriving experiences varied by facility, and whether these experiences also predicted change in perceived legitimacy.

By adding the facility ("boot camp") and control variables, the final stage of the analysis tested whether the experience variables mediated the facility effect. Control variables account for other factors that might influence our independent and dependent variables. For example, many studies have found that Blacks perceive the justice system to be less legitimate than Whites (Hagan and Albonetti, 1982; Hagan, Shedd, and Payne, 2005; Henderson, Cullen, Cao, Browning, and Kopache, 1997). Perhaps African Americans are also more likely to perceive negative experiences while incarcerated and as a result leave confinement with worse attitudes than their White counterparts. A similar argument could be made for older inmates and individuals with

15. Principal components factor analysis confirmed our grouping of these items. Reliability alphas for each scale are reported in Table 2.

16. To calculate change scores, we used simple subtraction (time 2 – time 1), which means that someone who attributed 0 (no legitimacy) at time 1 and 3 (little legitimacy) at time 2 experienced the same change as someone who went from 9 to 12 on the scale. Although sometimes starting and ending at different absolute values, the relative change inmates experience across time is of more interest to this article. In this example, both inmates experienced the same relative improvement in attitudes.

09012-CrimJournal-Guts.indd 105 12/21/09 7:14:12 PM

more extensive criminal histories. By including these controls, we isolate the effect of incarceration experiences on change in legitimacy and then test whether these experiences explain why facility type predicts attitudinal change.[17]

Results

The first stage in the analysis examined the effect of facility assignment (boot camp or prison) on changes in perceived legitimacy.[18] Figure 1 shows a graphic representation of attitudinal change among each sample of inmates. Although similar at time 1, attitudes clearly diverged as inmates left the facilities 6 months later.

Supporting Hypothesis 1, the difference in attitudes at time 2 ($x_2 - y_2$) was statistically significant between facilities.[19] Facility type did predict change in legitimacy, which was a result driven by prison inmates getting much worse ($y_2 - y_1$). Contrary to the prediction that attitudes would worsen at both facilities, results indicate that the mean legitimacy score among boot camp participants did not decline at all.

In terms of the specific legitimacy items, inmates at the boot camp left with comparably higher attributions of legitimacy for each of the 13 items compared with inmates at the prison. Table 1 shows that five items in particular drive the effect. An interesting note is that these items span the entire scope of the justice system. From the police ("Police usually treat me dirty"), to courts ("I'm doing more time than I really deserve"), to corrections ("I would shake hands with a correctional officer"), attitudes were comparatively better at the boot camp.

Hypothesis 2 stated that specific confinement experiences would vary across facilities. Table 2 displays mean comparisons for each of the 16 experience items as well as the three scales. Both positive and negative experiences significantly differed between the boot camp and the prison. In each case, the direction of the relationship suggests that inmates sent to the boot

17. None of the control variables, however, can mediate the relationship between facility type and perceived legitimacy as neither race, age, nor criminal history vary across facilities. Controls only are included to test whether the effect of experiences on change in legitimacy remains significant after accounting for demographic differences.

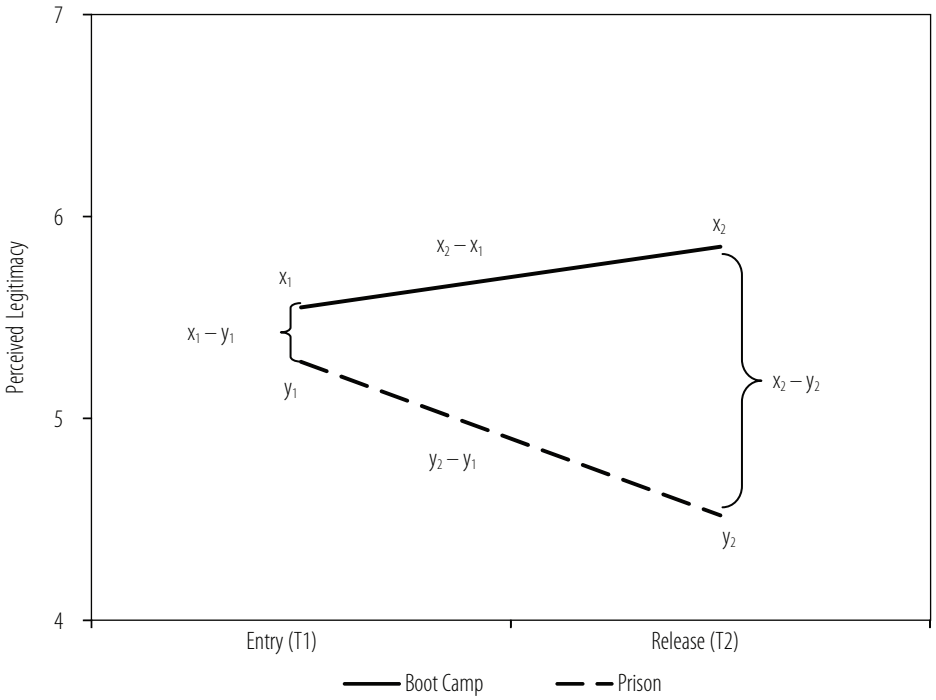
18. Where statistical modeling is concerned, however, the appropriate dependent variable for such analyses is debated. Some scholars argue that a simple change score (time 1 – time 2) is the most appropriate and straightforward dependent variable. Others, however, note that a change score is dependent on the initial time 1 score (Cohen and Cohen, 1983). Therefore, if the independent variable (in this case, "experiences") is correlated with the time 1 measure of legitimacy, it will have a stronger impact on the change score because time 1 legitimacy is part of the explained variance. Cohen and Cohen (1983) recommended using the time 2 measure as the dependent variable and control for time 1 where this is the case, which ensures that the relationship between the independent variables and the regressed change is net of their relationship to the time 1 measure. In this article, however, time 1 legitimacy is not correlated with any independent "experience" variables. As a result, a simple change score seems the most appropriate dependent variable for our analyses. We did, however, run an additional analysis with time 2 legitimacy as the dependent variable, controlling for time 1 legitimacy. Results were unchanged.

19. Because demographic variables, criminal history, and perceived legitimacy ($x1 - y1$) were statistically identical between facilities at time 1, a simple *t* test that compared the mean for perceived legitimacy at time 2 between facilities was sufficient to test the first hypothesis. However, bivariate regression results of change in legitimacy on facility type also are reported in the first model of Table 3.

09012-CrimJournal-Guts.indd 106 12/21/09 7:14:12 PM

FIGURE 1

Change in Perceived Legitimacy



	B	SE
Within-group difference		
Boot camp ($x_2 - x_1$)	0.30	.31
Prison ($y_2 - y_1$)	–0.76*	.29
Between-group difference		
Time 1 ($x_1 - y_1$)	0.27	.43
Time 2 ($x_2 - y_2$)	1.33*	.41

*$p < .05$.

09012-CrimJournal-Guts.indd 107 12/21/09 7:14:12 PM

camp perceived more positive experiences and less negative experiences. In particular, boot camp inmates were far more likely to think the staff helped them. In terms of environmental deprivation, however, issues like the lack of privacy, boredom, and noise were equally problematic at the boot camp and at the prison.

TABLE 2

Differences in Experiences Between Boot Camp and Prison

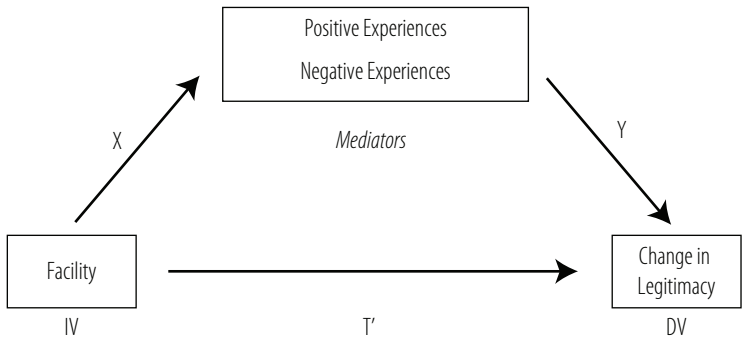
		Boot Camp		Prison		Mean	SE
	Range	Mean	SD	Mean	SD	Difference	Difference
Positive experiences[a]							
Substance abuse treatment helped me change for the better	0–4	2.64	1.13	2.49	1.29	.15	.17
The experience here has been good for me	0–4	2.92	1.14	1.76	1.46	1.16**	.19
My experiences here will help me get a job	0–4	2.90	1.14	2.11	1.53	.78**	.19
Things here keep me focused on my future	0–4	3.03	1.04	2.54	1.34	.49**	.17
I learned a lot from academic classes here	0–4	2.67	1.17	1.95	1.50	.71**	.19
The programs I was in helped me change for the better	0–4	2.83	1.09	2.39	1.27	.44*	.17
The staff here helped me change for the better	0–4	2.28	1.20	1.11	1.33	1.17**	.18
Negative experiences[a]							
I worried about my safety here	0–4	1.38	1.33	2.32	1.38	−.94**	.19
Drugs are easy to get here	0–4	1.32	1.04	2.70	1.32	−1.38**	.17
Guards ignore conflicts among inmates here	0–4	1.27	1.04	2.29	1.22	−1.03**	.16
Many accidents happen here	0–4	1.74	1.13	2.67	1.17	−.93**	.16
Weapons are easy to get here	0–4	1.09	1.05	2.74	1.30	−1.64**	.17
Inmates fight with other inmates here	0–4	2.11	1.12	3.25	0.87	−1.13**	.14
Environmental deprivation[b]							
Boredom	0–3	1.64	1.18	1.85	1.08	−.21	.17
Lack of privacy	0–3	2.12	1.00	2.19	1.02	−.06	.15
Loud environment	0–3	1.63	1.06	1.74	1.09	−.12	.16
Scales							
Positive experiences (alpha = .87)	0–28	19.40	5.92	14.33	6.69	5.07**	.90
Negative experiences (alpha = .83)	0–24	8.95	4.27	15.90	4.83	−6.95**	.64
Environmental deprivation (alpha = .63)	0–9	5.31	2.53	5.82	2.33	−.51	.34

[a] 0 (Strongly disagree) through 4 (Strongly agree).
[b] 0 (Not a problem for me) through 3 (Major problem for me).
*$p < .05$. **$p < .01$.

09012-CrimJournal-Guts.indd 108 12/21/09 7:14:12 PM

FIGURE 2

Mediation Conceptual Model



Notes. DV = dependent variable; IV = independent variable

The next question asked whether these experiences also predicted change in perceived legitimacy. Model 2 in Table 3 shows support for this hypothesis because all three experience measures affect attitudinal change in the predicted direction. Negative experiences and environmental deprivation worsened perceptions of legitimacy, whereas positive experiences improved attitudes toward the justice system. Results suggest that inmates who gained something from their incarceration developed more favorable attitudes toward the justice system in general.[20] In short, the experiences inmates had while serving their 6-month sentences played a significant role in how they came to view the entire justice system at their release.

To this point, results at least partially support Hypotheses 1, 2, and 3. Facility type predicts change in legitimacy, with inmates sent to prison developing more negative attitudes toward the justice system compared with individuals sent to boot camp. Two of the three experience scales varied by facility, and all three predicted change in attitudes toward the justice system. Not only did negative events and environmental deprivation plunge attitudes even lower, positive experiences actually *promoted* legitimacy.

The final question asked whether incarceration experiences mediated the facility effect. In other words, was the boot camp effect explained by these experiences? Figure 2 illustrates the predicted relationship.[21] $X \times Y$ represents the mediated effect and $(X \times Y) + \tau'$ represents the total effect. If the direct effect (τ') of facility on change in legitimacy is 0, then Hypothesis 4 will be supported.

20. For example, scoring one standard deviation above the mean on the positive experience scale corresponded with a 14% improvement in perceived legitimacy. Together, the model predicts a 37% difference in attitudinal change between individuals who scored one standard deviation better across the three experience scales versus those who scored one standard deviation worse on all three.

21. Because "environmental deprivation" did not vary across facilities, it is similar to our control variables in that it cannot mediate the relationship between facility type and attitudinal change. Therefore, positive and negative experiences are the only factors in the model that can explain why facility type predicts change in perceived legitimacy.

09012-CrimJournal-Guts.indd 109 12/21/09 7:14:13 PM

Model 3 in Table 3 highlights two important findings. First, the control variables have little effect on any of the three experience scales.[22] Regardless of age, race, or criminal history, the experiences of inmates while incarcerated influences whether they left confinement with better or worse attitudes toward the justice system. Second, the facility effect disappeared. In fact, the magnitude of the facility coefficient dropped by more than 200%, implying that aspects of confinement such as treatment by staff, quality of programming, and interactions with other inmates (or at least the perception of these things) were what caused prisoners to develop less respect, faith, and confidence in the justice system.

TABLE 3

Ordinary Least-Squares Regression Models Predicting Perceived Legitimacy of Justice System

	Model 1		Model 2		Model 3	
	B	SE	B	SE	B	SE
Facility						
Boot camp	1.06*	(.42)			0.23	(.57)
Confinement experiences						
Positive experiences			0.08*	(.04)	0.07*	(.04)
Negative experiences			–0.10*	(.04)	–0.09	(.05)
Environmental deprivation			–0.27**	(.09)	–0.28**	(.09)
Controls						
Age					–0.13*	(.06)
White					0.80	(.69)
Prior arrests					0.02	(.06)
Intercept	–0.76**	(.29)	1.28	(1.14)	3.92*	(1.87)
R^2	.03		.15		.18	

*$p < .05$. **$p < .01$.

In sum, each hypothesis was supported by the data. Particularly noteworthy is the finding that serving a 6-month sentence can substantially alter inmates' views of the entire justice system. Especially *de*legitimizing was the experience of serving time in a prison where attitudes toward the justice system worsened significantly. The military-style correctional boot camp, however, was not a delegitimizing experience. Most inmates there left confinement with similar or even better attitudes toward authority. Finally, certain experiences—both positive and negative—explained this facility effect.

22. "Positive experiences" remained marginally significant in the final model ($p = .08$).

09012-CrimJournal-Guts.indd 110 12/21/09 7:14:13 PM

Discussion

By using a randomized design, we attempted to answer the following three questions: whether perceptions of justice system legitimacy change during the course of incarceration; whether the type of correctional facility matters; and if so, why? Findings suggest that perceptions do change and that facility type does matter, but only to the extent that the facility employs specific legitimacy-building characteristics. By increasing positive experiences, decreasing negative experiences, and limiting environmental deprivation, correctional facilities can substantially improve inmates' perceptions of justice system legitimacy.

As such, it would be a mistake to interpret these results as either a blanket recommendation for correctional boot camps or a condemnation of traditional prisons. Rather, in accordance with Sparks and Bottoms' (1995) recommendations to the British penal system, our results suggest that *all* correctional options can implement legitimacy-building approaches. Of particular importance are issues related to staff (including—but not limited to—fair treatment), program effectiveness, ensuring safety, limiting environmental stressors, and preventing contraband. These issues are not about coddling inmates. The inmates in our sample did not seem to want special treatment. They wanted equal treatment; they wanted to gain something from the incarceration experience, and they benefited when staff members followed suit. In this study, the prison performed worse across all dimensions. Yet we would not be surprised to see a traditional prison outperform a boot camp in another setting.

Furthermore, we agree with Bottoms' (1999) notion that, when examining predictors of legitimacy within a correctional framework, it is best to be inclusive. More than procedural or distributive fairness, broader aspects of the regime itself bear heavily on inmates' attitudes toward authority; factors such as staff intentions, daily routines, environmental structure, humaneness, and interactions with other inmates can influence attitudes substantially. In fact, most of these aspects change not only perceptions of the correctional system but also attitudes toward the police, courts, and the justice system as a whole.

Particularly encouraging is the finding that certain experiences actually can improve (and, in some cases, create) perceptions of justice system legitimacy. Too often we think of prisons as inherently illegitimate. Indeed, some scholars have argued that it is impossible for prisons to create any sense of legitimacy and, therefore, should rule strictly based on coercion, fear, and control and pay little or no attention to due-process rights (Scraton et al., 1991). This article presents evidence contrary to that view. Even at the prison, the presence of positive experiences and the lack of negative ones promoted legitimacy. Inmates who gained something from their experience by virtue of meaningful interactions with staff or programming, for instance, were more likely to recognize the legitimacy of the justice system at release. The problem, it seems, is not that prisons are inherently illegitimate, but that some regimes fail to address the structural, behavioral, procedural, distributive, and environmental problems that can be delegitimizing. At the prison in this sample, delegitimizing experiences such as violence and perceived staff incompetence seemed to be the norm.

09012-CrimJournal-Guts.indd 111 12/21/09 7:14:13 PM

The correctional boot camp, however, painted an entirely different picture. Here, positive experiences dwarfed negative ones. Inmates were far more likely to feel that they benefited from their time at the boot camp—despite being yelled at, having to carry rocks and do pushups, having little free time, and being denied television. Regardless of these less favorable experiences, boot camp inmates remained more likely than prisoners to report that the staff and programming helped them, and that the entire experience was good for them. As a result, these experiences translated to a higher average legitimacy score at release.

Unfortunately, our survey items did not include measures of specific theoretical constructs that have been linked to legitimacy, such as procedural and distributive justice. We can, however, return to our observational data to view the findings in terms of process and outcome fairness. For example, we observed at the boot camp a rigidly structured, grueling regime, and we noted many instances of distributively unfair punishments. Yet—despite high levels of situational control and harsh punishments—inmates bought in to the regime. This cooperation seemed in part because of procedural safeguards: Treatment was consistent, impartial, and ethical, and sanctioning decisions were meticulously explained. But it was also a function of the trust drill instructors attained. Boot camp inmates truly believed that the staff shared their main interests—namely, to become better human beings. This "motive-based trust," which Tyler (1990) distinguished from procedural and distributive fairness, was in fact one of the most discernable differences between the two facilities. According to many inmate accounts (and occasionally the direct observations of researchers), staff at the traditional prison simply did not care. In the end, boot camp inmates might have left their confinement with a greater stake in the justice system mainly because staff took a greater stake in them.

Without this qualitative information, discussions of these theoretical concepts might prove problematic. Indeed, just as the rich landscape of authority interactions in corrections provides new avenues for theoretical development, it also introduces some measurement complexity. For instance, many of our quantitative "experience" items could be interpreted as either procedural or distributive. Boot camp participants were more likely to report "The staff here helped me change for the better," but was this self-improvement because drill instructors were more respectful than prison guards (procedural) or because therapeutic programming was more easily accessible at the boot camp (distributive)? Even program availability itself (on its face, decidedly distributive) could become an issue of procedural fairness if some inmates were admitted to classes instead of others because of a perceived bias.

Considering this complexity, future endeavors into legitimacy in corrections should take a nuanced approach to examine these constructs. We show that legitimacy remains malleable at the final stage of the justice system (i.e., it can improve or worsen) and that certain incarceration experiences explain this attitudinal change. Future research should test directly the theoretical constructs discussed and examine the relative importance of procedural fairness, distributive fairness, and motive-based trust. Providing even more context through inmate narratives would be another useful research agenda.

09012-CrimJournal-Guts.indd 112 12/21/09 7:14:13 PM

Likewise, the extent to which perceptions of legitimacy were because of tangible aspects of the prison regime versus the diffusion of beliefs from other inmates is unknown. Recall that participants sent to the prison were housed in general population, whereas each incoming boot camp platoon was isolated from the rest of the population. Therefore, a more thorough examination of the mechanisms linking the incarceration experience to perceptions of legitimacy also would be a useful direction for future research.

Furthermore, it would prove useful to explore whether confinement experiences directly affect perceptions of legitimacy or whether they influence legitimacy through other pathways, such as programming. It is possible that positive experiences lead inmates to buy in to educational classes, therapeutic treatment, life-skills training, and so on, and it could be the effectiveness of these programs (not the positive or negative experiences themselves) that caused perceptions of the justice system to change.

Another direction for future research would be to examine the antecedents of these incarceration experiences. In other words, what characteristics of the overall prison regime predict whether inmates will perceive positive, negative, or environmentally depriving conditions? Why, in our case, did the boot camp employ more legitimacy-fostering characteristics than the prison? Finally, it would be useful to know the extent to which perceptions of legitimacy affect reoffending outcomes after release. To our knowledge, this issue has not been tested with an inmate sample.

Those questions aside, this article has addressed several gaps in the current research. In particular, the use of a randomized design—rare in correctional research—provided solid ground for drawing conclusions about the effects of two different punishment regimes. An advantage to this study is that neither the type of offender nor the phase of incarceration confounded the results, because the design created groups that would be expected to vary only randomly and surveys were administered at the same phase for all participants. Second, we have attempted (despite being unable to record interviews) to provide some observational context to the quantitative findings. Third, a repeated-measures design allowed for examining change across time. To date, most legitimacy studies have measured perceptions at only one point in time and, thus, could not determine whether experiences with the justice system improve or worsen such attitudes.

But perhaps the largest gap this article addresses—and one that remains wide—is the scarcity of research that explores legitimacy in the correctional system. To date, studies have tended to stop at experiences with the police and the court system, thus ignoring the influence incarceration has on perceptions of offenders. As our results indicate, the incarceration experience can change perceptions of legitimacy. From a policy standpoint, correctional regimes should recognize two important points. First, inmates can leave confinement with much worse or even better attitudes toward the justice system than when they entered. Second, the nature of that outcome is in no small measure influenced by the way each facility is administered. By providing an environment in which inmates can benefit from their confinement effectively, correctional institutions can create among their inmates a greater sense that the law and legal authorities should be obeyed—that these institutions are, in fact, legitimate.

09012-CrimJournal-Guts.indd 113 12/21/09 7:14:13 PM

References

Benesh, Sarah and Susan Howell. 2001. Confidence in the courts: A comparison of users and non-users. *Behavioral Sciences & the Law*, 19: 199–214.

Bierie, David. 2009. Cost matters: A randomized experiment comparing recidivism between two styles of prison. *Journal of Experimental Criminology*, 5: 371-397.

Bonta, James A. and Paul Gendreau. 1990. Reexamining the cruel and unusual punishment of prison life. *Law and Human Behavior*, 14: 347–366.

Bottoms, Anthony. 1999. Interpersonal violence and social order in prisons. In (Michael H. Tonry and Joan Petersilia, eds.), *Prisons, Crime, and Justice: An Annual Review*. Chicago, IL: University of Chicago Press.

Brown, Ethan. 2007. *Snitch: Informants, Cooperators, and the Corruption of Justice*. Philadelphia, PA: Public Affairs Books.

Bukstel, Lee and Peter Kilmann. 1980. Psychological effects of imprisonment on confined individuals. *Psychological Bulletin*, 88: 469–493.

Clemmer, Donald. 1940. *The Prison Community*. Boston, MA: Holt, Rinehart, Winston.

Cohen, Jacob and Patricia Cohen. 1983. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Hillsdale, NJ: L. Erlbaum Associates.

DiIulio, John. 1987. *Governing Prisons: A Comparative Study in Correctional Management*. New York: Free Press.

Easton, David and Jack Dennis. 1969. *Children in the Political System*. Chicago, IL: University of Chicago Press.

Ewick, Patricia and Susan Silbey. 1998. *The Common Place of Law: Stories from Everyday Life*. Chicago, IL: University of Chicago Press.

Gibson, James L., Gregory A. Caldeira, and Lester Kenyatta Spence. 2003. Measuring attitudes toward the United States Supreme Court. *American Journal of Political Science*, 47: 254–367.

Goffman, Erving. 1961. *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates*. New York: Doubleday Anchor.

Hagan, John and Celesta Albonetti. 1982. Race, class, and the perception of criminal injustice in America. *American Journal of Sociology*, 88: 329–355.

Hagan, John, Carla Shedd, and Monique R. Payne. 2005. Race, ethnicity, and youth perceptions of criminal injustice. *American Sociological Review*, 70: 381–407.

Heckman, James J. and Jeffrey A. Smith. 1995. Assessing the case for social experiments. *Journal of Economic Perspectives*, 9: 85–110.

Henderson, Martha. 2000. *The Meaning of Punishment: What Do Inmates Expect from Prison?* Unpublished doctoral dissertation, University of Cincinnati.

Henderson, Martha, Francis T. Cullen, Linqun Cao, Sandra Lee Browning, and Renee Kopache. 1997. The impact of race on perceptions of criminal injustice. *Journal of Criminal Justice*, 25: 447–462.

Hoffman, Martin. 1977. Moral internalization: Current theory and research. *Advances in Experimental Social Psychology*, 10: 85–133.

09012-CrimJournal-Guts.indd 114 12/21/09 7:14:14 PM

Jesness, Carl. 1996. *The Jesness Inventory: Manual*. North Tanawanda, NY: Multi-Health Systems.

Kimball, David C. and Samuel C. Patterson. 1997. Living up to expectations: Public attitudes toward Congress. *Journal of Politics*, 59: 701–728.

LaFree, Gary. 1998. *Losing Legitimacy: Street Crime and the Decline of Social Institutions in America*. Boulder, CO: Westview Press.

Laub, John H. and Robert J. Sampson. 2003. *Shared Beginnings, Divergent Lives: Delinquent Boys to Age 70*. Cambridge, MA: Harvard University Press.

Lind, Allen and Tom R. Tyler. 1988. *The Social Psychology of Procedural Justice*. New York, NY: Plenum Press.

Linder and Associates, Inc. 2002. *Baltimore Believe: Progress Report*. Retrieved 6/4/2008 from rebuildingmadison.info/BelieveReport.pdf.

Leschied, Alan W. and Paul Gendreau. 1994. Doing justice in Canada: YOA policies that can promote community safety. *Canadian Journal of Criminology*, 36: 291–303.

Liebling, Alison and Helen Arnold. 2004. *Prisons and Their Moral Performance: A Study of Values, Quality, and Prison Life*. Oxford, UK: Oxford University Press.

Loza, Wagdy and Amel Loza-Fanous. 2000. Predictive validity of the self-appraisal questionnaire (SAQ). *Journal of Interpersonal Violence*, 15: 1183–1191.

MacKenzie, Doris L., David Bierie, and Ojmarrh Mitchell. 2007. An experimental study of a therapeutic boot camp: Impact on impulses, attitudes, and recidivism. *Journal of Experimental Criminology*, 3: 221–246.

MacKenzie, Doris L., Ojmarrh Mitchell, David Bierie, Mischelle Van Brakle, Lauren O'Neill, Derrick Franke, and Fawn Ngo Mitchell. 2004. *A Randomized Study of the Maryland Correctional Boot Camp for Adults: Final Report*. Towson: Maryland Governor's Office of Crime Control and Prevention.

MacKenzie, Doris L. and Claire Souryal. 1995. Inmates' attitude change during incarceration: A comparison of boot camp with traditional prison. *Justice Quarterly*, 12: 325–353.

MacKenzie, Doris L., David B. Wilson, Gaylene S. Armstrong, and Angela R. Gover. 2001. The impact of boot camps and traditional institutions on juvenile residents: Perceptions, adjustment, and change. *Journal of Research in Crime and Delinquency*, 38: 279–313.

O'Neill, Lauren, Doris L. MacKenzie, and David Bierie. 2007. Educational opportunities within correctional institutions: Does facility type matter? *The Prison Journal*, 87: 311–327.

Paternoster, Raymond, Robert Brame, Ronet Bachman, and Larry Sherman. 1997. Do fair procedures matter? The effect of procedural justice on spouse assault. *Law & Society Review*, 31: 163–204.

Paulus, Paul B. and Mary T. Dzindolet. 1993. Reactions of male and female inmates to prison confinement: Further evidence for a two-component model. *Criminal Justice and Behavior*, 20: 149–166.

Rosenfeld, Richard, Bruce Jacobs and Richard Wright. 2003. Snitching and the code of the street. *British Journal of Criminology*, 43: 291–309.

09012-CrimJournal-Guts.indd 115 12/21/09 7:14:14 PM

Ruback, R. Barry, Timothy S. Carr, and Charles H. Hopper. 1986. Perceived control in prison: Its relation to reported crowding, stress, and symptoms. *Journal of Applied Social Psychology*, 16: 375–386.

Sampson, Robert J. and Dawn Jeglum Bartusch. 1998. Legal cynicism and (subcultural?) tolerance of deviance: The neighborhood context of racial differences. *Law & Society Review*, 32: 777–804.

Scraton, Phil, Joe Sim, and Paula Skidmore. 1991. *Prisons Under Protest*. Philadelphia, PA: Open University Press.

Sherman, Larry. 1993. Defiance, deterrence, and irrelevance: A theory of the criminal sanction. *Journal of Research in Crime and Delinquency*, 30: 445–473.

Skogan, Wesley G. and Kathleen Frydl. 2003. *Fairness and Effectiveness in Policing: The Evidence*. Washington, DC: National Academies Press.

Sparks, Richard and Anthony Bottoms. 1995. Legitimacy and order in prisons. *British Journal of Sociology*, 46: 45–62.

Sparks, Richard, Anthony Bottoms, and Will Hay. 1996. *Prisons and the Problem of Order*. Oxford, UK: Clarendon Press.

Styve, Gaylene, Doris L. MacKenzie, Angela R. Gover, and Ojmarrh Mitchell. 2000. Perceived conditions of confinement: A national evaluation of juvenile boot camps and traditional facilities. *International Journal of Offender Therapy and Comparative Criminology*, 44: 450–467.

Sunshine, Jason and Tom R. Tyler. 2003. The role of procedural justice and legitimacy in shaping public support to policing. *Law & Society Review*, 37: 513–547.

Sykes, Gresham M. 1958. *The Society of Captives: A Study of a Maximum-Security Prison*. Princeton, NJ: Princeton University Press.

Thibaut, John W. and Laurens Walker. 1975. *Procedural Justice: A Psychological Analysis*. Hillsdale, NJ: Lawrence Erlbaum.

Tyler, Tom R. 1984. The role of perceived injustice in defendants' evaluations of their courtroom experience. *Law & Society Review*, 18: 59–71.

Tyler, Tom R. 1990. *Why People Obey the Law*. New Haven, CT: Yale University Press.

Tyler, Tom R. 2000. Social justice: Outcome and procedure. *International Journal of Psychology*, 35: 117–125.

Tyler, Tom R. 2001. Public trust and confidence in legal authorities: What do majority and minority group members want from the law and legal authorities? *Behavioral Science & the Law*, 19: 215–235.

Tyler, Tom R. 2003. Process-based regulation: Procedural justice, legitimacy, and the effective rule of law. In (Michael H. Tonry, ed.), *Crime & Justice*. Chicago IL: University of Chicago Press.

Tyler, Tom R. and Yuen Huo. 2002. *Trust in the Law: Encouraging Public Cooperation with the Police and Courts*. New York: Russell Sage Foundation.

Weber, Max. 1968. The types of legitimate domination. In (Guenther Roth and Claus Wittich, eds.), *Economy and Society*. New York: Bedminster Press.

09012-CrimJournal-Guts.indd 116 12/21/09 7:14:14 PM

Wellford, Charles. 1967. Factors associated with adoption of the inmate code. *Journal of Criminal Law, Criminology, and Political Science*, 58: 197–203.

Zamble, Edward and Frank Porporino. 1988. *Coping, Behavior, and Adaptation in Prison Inmates*. New York: Springer-Verlag.

Zamble, Edward and Frank Porporino. 1990. *Coping with imprisonment. Criminal Justice and Behavior,* 17: 53–70.

Derrick Franke is a doctoral student at the University of Maryland, College Park. His research interests include legitimacy, procedural justice, restorative justice, and reentry. Recent publications appear in *Western Criminology Review* and *International Studies Review*.

David Bierie is a research analyst with the Federal Bureau of Prisons who received his Ph.D. from the Department of Criminology and Criminal Justice at the University of Maryland, College Park, in 2007. His work has appeared in *The Journal of Experimental Criminology*, *Crime & Delinquency*, *Western Criminology Review*, and *The Criminal Law Bulletin*. His primary research interests include quantitative methods, program evaluation, cost–benefit methodology, and incarceration.

Doris L. MacKenzie is director of the Justice Center for Research – College of the Liberal Arts and University Outreach, and Professor, Crime, Law and Justice Program, Department of Sociology at the Pennsylvania State University. Her most recent book is titled *What Works in Corrections: Examining the Criminal Activities of Offenders and Delinquents* (2006, Cambridge University Press). This work summarizes systematic reviews and meta-analyses of evaluations of correctional management strategies and rehabilitation programs. She has co-edited the following books: *Correctional Boot Camps: Military Basic Training or a Model for Corrections?* (2004, Sage) and *Different Crimes, Different Criminals* (2006, Lexis/Nexis/Anderson).

09012-CrimJournal-Guts.indd 117 12/21/09 7:14:14 PM

POLICY ESSAY

LEGITIMACY PERCEPTIONS IN CORRECTIONS

Transforming attitudinal change into behavioral change

The missing link

Megan C. Kurlychek
University at Albany

When correctional boot camps first emerged as an alternative to traditional prison in 1983, they intrigued the public, correctional policy makers, and researchers alike. In fact, one review identified 152 independent studies of correctional boot camps (Wilson, MacKenzie, and Mitchell, 2003). Despite the popularity of these programs, however, evaluations have remained far from glowing. A careful meta-analysis of 44 different experimental and quasi-experimental evaluation studies concluded that correctional boot camps do not provide either a significant increase or decrease in the odds of recidivism compared with other incarceration options (odds ratio = 1.02; Wilson et al., 2003).

One might then reasonably inquire why researchers (including the guilty author of this essay) continue to explore this topic relentlessly? I propose several key reasons for the continued love–hate relationship with correctional boot camps. First, the boot camp concept simply has curb appeal. The idea of taking offenders and subjecting them to grueling physical routines seems to be the best of both worlds because it surely exudes punishment for actions committed while simultaneously offering hope of rehabilitation by instilling self-discipline. However, those who oppose boot camps continue to search for their "fatal flaw" because this treatment method is viewed as too harsh. According to this line of thinking, the degrading comments and physical punishments handed down by drill sergeants are, at best, detrimental to the rehabilitation process and, at worst, simply inhumane (Lutze and Brody, 1999; Morash and Rucker, 1990).

Perhaps, though, the most intriguing findings that keep this research and correctional boot camps alive and well are the findings that inmates who serve time in correctional boot camps *do* fair better than their prison counterparts on many measurable characteristics. Specifically, they seem to experience positive attitudinal change, routinely report feeling safer than inmates

Direct correspondence to Megan C. Kurlychek, School of Criminal Justice, University at Albany, 135 Western Avenue, Albany, NY 12222 (e-mail: mkurlychek@albany.edu).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 119 12/21/09 7:14:14 PM

who serve time in traditional prisons, and report forming strong bonds and attachments with staff and counselors (Kempinen and Kurlychek, 2003a; MacKenzie and Shaw, 1990; MacKenzie and Souryal, 1994, 1995). Yet despite these findings, study after study reports that these changes have no long-term impact on offending behavior (Burns, Anderson, and Dyson, 1997; Florida Department of Corrections, 1990; Kempinen and Kurlychek, 2003b; Kurlychek and Kempinen, 2006; MacKenzie, Shaw, and Gowdy, 1990; MacKenzie, 1991; MacKenzie and Souryal, 1994).

"To help, or at least, to do no harm"
—Hippocrates
(Epidemics: Section XI)

Franke et al. (2010) add yet one more piece to this puzzle. Based on a sample of 234 inmates randomly assigned to serve their time in either prison or boot camp, Franke et al. report that the boot camp has significant advantages compared with the prison regarding both the experiences of offenders while incarcerated and the influence these experiences had on the attitudes of offenders toward the legitimacy of the system. That is, although those serving time in prison reported a decrease in perceptions of legitimacy, the boot camp sample did not. These attitudes of the offender toward legitimacy remained basically unchanged and, in some measures, even improved. As Franke et al. propose, attitudes toward system legitimacy are thought to affect illegal behavior because one is more likely to obey the law, regardless of personal morals and values, if one respects the authority of the system (Tyler, 1990, 2003). Yet, although Franke et al. conclude that the boot camp environment, at a minimum, "did no harm" to offender attitudes and that the prison had a deleterious effect, the study falls short of evaluating the later law-abiding behavior of the two groups.

Perhaps, such an evaluation is not really needed. I probably can tell before conducting such an experiment what I would find. Similar to so many studies before, it is most likely that the boot camp sample would recidivate at approximately the same rate as the prison sample. The compelling yet unanswered question then becomes "Why?" Why does the boot camp experience ignite the positive attitudinal changes in offenders that prison does not? Perhaps more importantly, why do these changes not then translate into better post-release outcomes? In the following sections, I provide some practical clues toward answering these questions.

Differences: The Prison and Boot Camp Experience

To address the first question, let us begin with a simple consideration of the physical and structural differences between the prison and the boot camp experience. First of all, boot camps are usually smaller than prisons and serve a defined, rather homogeneous, population. Indeed, this experience was the case for Franke et al. (2010), in which the overcrowded prison built for 1,200 offenders housed 1,500 and the smaller boot camp built for 450 offenders averaged a daily population of only 325. In addition, although the prison housed offenders of all types and risk levels, the boot camp was designed for relatively low-risk offenders.

09012-CrimJournal-Guts.indd 120 12/21/09 7:14:15 PM

In addition to size and population composition, Franke et al. (2010) draw our attention to the actual appearance of the facilities and referred to the boot camp as having "well-manicured lawns" and "well-maintained dormitory-style units," whereas the prison was described as "dilapidated" and suffering from a pigeon infestation. Given a choice of a facility at which to serve time, one would almost undoubtedly choose the boot camp—regardless of the programming and treatment offered.

These differences lead naturally to the consideration of treatment options and availability. Here, the study finds that 100% of the boot camp sample reported receiving education and treatment services compared with 34% of the prison sample receiving education, 43% receiving cognitive skills training, and 64% receiving drug treatment. This programming is typical of "new generation" boot camps in which education and treatment are built into the daily regimen. Beyond the simple fact that much of the cognitive skills training is aimed at improving prosocial attitudes and decision-making skills, this difference has a second implication. Specifically, Franke et al. (2010) note that although the "highly structured boot camp regime affords inmates little idle time," the inmates in the prison sample had "much more unstructured time." This unstructured time then allowed for the prison sample to interact with the heterogeneous inmate population, which continued to share criminogenic attitudes and behaviors in a casual manner.

However, none of the boot camp characteristics noted so far as "improvements" compared with the traditional prison environment really have had anything to do with the military experience for which such boot camps are touted. Instead, they have been associated with basic comfort and safety, the availability of programming, and the structuring of time. Perhaps then, it is the provision of more positive experiences while incarcerated and greater access to treatment that influences attitudinal change and not the military-style disciplinary experience. Franke et al. (2010) acknowledge this possibility in their discussion when they note that even in the prison environment, positive experiences increased attitudes of legitimacy, whereas negative experiences decreased attitudes of legitimacy. Yet we are still left with the pesky question of why observed attitudinal changes found in prior boot camp studies have failed to correspond with improved outcomes after release. I address this question in the following section.

Similarities: The Prison and Boot Camp Post-Release Experiences

Despite these obvious differences in the incarceration experience between a traditional boot camp and prison, why then have studies routinely found little to no difference in post-release recidivism? In the simplest terms, the problem seems to be that whatever changes we can invoke in offenders during their confinement are met with stark contrast when they are released.

Take, for example, the daily routine of the boot camp inmate—awake at the crack of dawn, perform a period of intense physical activity, and follow a day highly structured between education and therapeutic treatment. Throughout the program the inmates work as a "platoon" and share in the successes and failures of one another. Throughout time, the inmates learn how to adapt to the program to avoid the harsh physical discipline and might even buy into the program and develop a sense of pride in their accomplishments.

09012-CrimJournal-Guts.indd 121 12/21/09 7:14:15 PM

However, unlike the military recruit who attends boot camp to learn the skills needed to become a better soldier and then is assigned to his or her next post in the military, the correctional boot camp inmate is simply released back to the streets from where he or she came. For the would-be soldier, the boot camp is but the beginning of training. For the inmate, the correctional boot camp is the entire treatment experience. Moreover, many skills learned within the boot camp do not translate to success on the street (e.g., working as a group or responding to a highly regimented routine). As importantly, when leaving the military, the Marine has a life-long social status as a Marine, but the correctional boot camp inmate leaves the experience with the life-long stigma of "convicted offender" or "ex-con."

Compounding the situation more are the harsh realities of blocked opportunities. For example, many programs focus on helping an offender obtain a GED or job-skills training. Boot camp inmates in particular seem to have hope for the future, which indicates that they intend to get a job when released. Yet, employers are hesitant to hire convicted offenders, and when offenders do find employment, it is often in low-paying jobs (Stoll, Raphael, and Holzer, 2006). For example, the Pennsylvania Motivational Boot Camp Report to the Legislature (2003) indicates that, prior to going to boot camp, only 40% of their offender sample reported having been employed full time, but 66% reported earning more than $1,000 per month, with 19% of these reporting incomes of more than $4,000 per month, much from illegal earnings. After graduating from boot camp, 64% of the sample was able to obtain full-time employment. However, only 34% reported earning more than $1,000 per month and 0% earned more than $4,000 per month (Kempinen and Kurlychek, 2003b: 44–45, 56–57).

Thus, although most boot camp graduates could find some type of employment when released, their overall incomes dropped significantly when shifting from illegitimate to legitimate earnings, and most still were living well below the poverty line. In addition to financial difficulties, much research has noted that ex-convicts also have difficulties finding housing and medical care, reforming familial bonds, and refraining from prior associations with delinquent peer networks to name a few issues. Although a full discussion of the patterns of recidivism and theoretical explanations for such patterns is clearly beyond the scope of this essay, the primary point I wish to impart is that the perceived failure of so many rehabilitation programs might not be because they fail to invoke change in offenders. In the case of boot camps especially, we have evidence to the contrary. What we are apparently sorely lacking is a better understanding of the programs and services necessary to help offenders overcome the obstacles to reentry and to capitalize on the gains made while in treatment.

Aftercare Programming and Research: The Missing Link?

Although expanded efforts have been made to offer aftercare programming to help institutionalized offenders in this process, to date, findings on aftercare maintain mixed results, with some showing reduced recidivism (Hiller, Knight, and Simpson, 1999; Inciardi, Martin, and Butzin, 2004; Kurlychek and Kempinen, 2006), others showing no change (Brown, O'Brady, Battjes, Farrell, Smith, and Nurco, 2001; Greenwood, Deschenes, and Adams, 1993), and still

09012-CrimJournal-Guts.indd 122 12/21/09 7:14:15 PM

others reporting increased recidivism (Sealock, Gottfredson, and Gallagher, 1997). Clearly, no consensus can be made on the state of aftercare.

So it is here that I offer some critical directions for future research and policy. First, I propose that reentry is simply not a "one size fits all" process. All too often, "aftercare" serves as a catchall term for any services offered after release instead of representing a theoretically based, well-planned transition program. Future research is needed to determine which aftercare programs truly work and why. In particular, a better understanding is needed of which services work, for whom, and in what dosages. Second, I propose that to customize this process we first must develop a deeper understanding of the pathways to both failure and success when inmates are released from institutional settings. This process involves addressing personal problems, such as addictive behaviors, and social processes, such as the draw of prior delinquent peer networks. If this can be achieved, then perhaps we can better develop and deliver services that help offenders maintain the positive gains achieved during treatment.

Finally, we also must develop methods to help offenders bridge the gap between simply maintaining attitudinal improvement and transforming such attitudes into real behavioral change. Toward this latter end, I suggest that we expand our attention beyond merely reforming the offender. I believe the solution ultimately will require that we also "reform" those social policies that limit and block opportunities for success. Although these obstacles are too many to address in one short essay, I will call to attention just a few that require immediate attention—education, employment, housing, health care, and civil disenfranchisement. As a profession, criminologists routinely study the social causes of crime. However, as a system, criminal justice too often remains focused strictly on the offender and fails to acknowledge or address those forces beyond the control of the offenders that ultimately influence criminal behavior.

In Conclusion: Legitimacy and Behavior

On a positive note, we have learned from Franke et al. (2010) and others that we can make significant changes in offender attitudes—even in the latest stages of criminal justice processing. The incarceration experience then is more than a period of incapacitation; it is an opportunity for reform. Moreover, as demonstrated by Franke et al., if we can improve offender attitudes, then we also have the potential to harm them as well. Thus, equally important, I propose that we have a moral responsibility and vested social interest to ensure that, at a minimum, we do no more harm to the offender while incarcerated.

The missing link, however, seems to be finding the appropriate services to help offenders maintain positive changes after their release and to translate these attitudinal changes into behavioral changes. Although I have suggested some research strategies to advance our knowledge about aftercare programs and effective transitional services, I do not naively assume that the task will be either quick or easy; real solutions seldom are. However, with an estimated 700,000 offenders returning to society from institutions each year, and upward of 60% of these failing, and failing quickly, it is a task that must be undertaken both for the well being of the offender and of our society.

09012-CrimJournal-Guts.indd 123 12/21/09 7:14:15 PM

In conclusion then, although the Franke et al. (2010) study addresses an interesting question about the influence of incarceration on offender attitudes toward legitimacy, I do not hold much enthusiasm that the difference in attitudes toward the system will lead to differences in long-term offending behavior. Indeed, all existing evidence predicts this outcome to be unlikely. Rather, it seems that unless something is done to remove impediments to success on the outside and to help offenders navigate this difficult period, all we will have done is provide them with more positive attitudes toward the system into which they will soon be returning.

"When can I come back and visit?"
—Boot Camp Graduate
(Pennsylvania Motivational Boot Camp Survey, 2003)

References

Brown, Barry S., Kevin E. O'Grady, Robert J. Battjes, Eugene V. Farrell, Nikeba P. Smith, and David N. Nurco. 2001. Effectiveness of a stand-alone aftercare program for drug-involved offenders. *Journal of Substance Abuse Treatment*, 21: 185–192.

Burns, Jerald, James F. Anderson, and Laronistine Dyson. 1997. What disciplinary rehabilitation unit participants are saying about shock incarceration. *Journal of Contemporary Criminal Justice*, 13: 172–183.

Florida Department of Corrections. 1990. *Boot Camp: A Twenty-Five Month Review*. Tallahassee, FL: Author.

Franke, Derrick, David Bierie, and Doris L. MacKenzie. 2010. Legitimacy in corrections: A randomized experiment comparing a boot camp with a prison. *Criminology & Public Policy*. This issue.

Greenwood, Peter W., Elizabeth Piper Deschenes, and John Adams. 1993. *Chronic Juvenile Offenders: Final Results from the Skillman Aftercare Experiment*. Santa Monica, CA: RAND.

Hiller, Matthew L., Kevin Knight, and D. Dwayne Simpson. 1999. Prison-based substance abuse treatment, residential aftercare, and recidivism. *Addiction*, 94: 833–842.

Inciardi, James A., Steven S. Martin, and Clifford T. Butzin. 2004. Five-year outcomes of therapeutic community treatment of drug-involved offenders after release from prison. *Crime & Delinquency*, 50: 88–107.

Kempinen, Cynthia and Megan C. Kurlychek. 2003a. An outcome evaluation of Pennsylvania's boot camp: Does rehabilitative programming within a disciplinary setting reduce recidivism? *Crime & Delinquency*, 49: 581–602.

Kempinen, Cynthia and Megan C. Kurlychek. 2003b. *Pennsylvania's Motivational Boot Camp: 2003 Report to the Legislature*. State College: Pennsylvania Commission on Sentencing.

Kurlychek, Megan C. and Cynthia Kempinen. 2006. Beyond boot camp: The impact of aftercare on offender reentry. *Criminology & Public Policy*, 5: 363–388.

09012-CrimJournal-Guts.indd 124 12/21/09 7:14:15 PM

Lutze, Faith E. and David C. Brody. 1999. Mental abuse as cruel and unusual punishment: Do boot damp prisons violate the Eighth Amendment? *Crime & Delinquency*, 45: 242–255.

MacKenzie, Doris L. 1991. The parole performance of offenders released from shock incarceration (boot camp prisons): A survival analysis. *Journal of Quantitative Criminology*, 7: 213–236.

MacKenzie, Doris L. and James W. Shaw. 1990. Inmate adjustment and change during shock incarceration: The impact of correctional boot camp programs. *Justice Quarterly*, 7: 125–150.

MacKenzie, Doris L., James W. Shaw, and Voncile B. Gowdy. 1990. *Evaluation of a Shock Incarceration in Louisiana: Executive Summary*. Washington, DC: National Institute of Justice.

MacKenzie, Doris L. and Claire Souryal, 1994. *Multisite Evaluation of Shock Incarceration*. Washington DC: U.S. Department of Justice, National Institute of Justice.

MacKenzie, Doris L. and Claire Souryal. 1995. Inmates' attitude change during incarceration: A comparison of boot camp and traditional prison. *Justice Quarterly*, 12: 325–353.

Morash, Merry and Lila Rucker. 1990. A critical look at the idea of boot camp as a correctional reform. *Crime & Delinquency*, 36: 204–222.

Sealock, Miriam D., Denise C. Gottfredson, and Christine A. Gallagher. 1997. Drug treatment for juvenile offenders: Some good and bad news. *Journal of Research in Crime and Delinquency*, 34: 210–236.

Stoll, Michael, Steven Raphael, and Harry Holzer. 2006. Will employers hire ex-offenders? Employer preferences, background checks, and their determinants. In (Mary Patillo-McCoy, David Weiman, and Bruce Western, eds.), *The Consequences of Mass Incarceration on Families and Communities*. New York: Russell Sage Foundation.

Tyler, Tom R. 1990. *Why People Obey the Law*. New Haven, CT: Yale University Press.

Tyler, Tom R. 2003. Process-based regulation: Procedural justice, legitimacy, and the effective rule of law. In (Michael H. Tonry, ed.), *Crime and Justice*. Chicago, IL: University of Chicago Press.

Wilson, David B., Doris L. MacKenzie, and Fawn Ngo Mitchell. 2003. *Effects of Correctional Boot Camps on Offending*. Oslo, Norway: The Campbell Collaboration.

Megan C. Kurlychek is an assistant professor in the School of Criminal Justice at the University at Albany, SUNY. Her primary research interests include juvenile justice and delinquency, corrections and sentencing, and offender reentry. Kurlychek has published several pieces focusing on the effect of rehabilitative programming offered in correctional boot camps and aftercare programs. Most recently, her research has focused on the many obstacles these offenders face as they attempt to reenter society, and the use of criminal history records in limiting legitimate opportunities. She worked previously as a research analyst for the Pennsylvania Commission on Sentencing, the National Center for Juvenile Justice, and the Pennsylvania State Senate, and continues to retain a major focus on public policy and evaluation research.

09012-CrimJournal-Guts.indd 125 12/21/09 7:14:15 PM

"Legitimacy in corrections"

Policy implications

Tom R. Tyler
New York University

The legal system is increasingly recognizing the importance of legitimacy to the effectiveness of law (Tyler, 2006a, 2006b). Legitimacy is the widespread belief among members of the public that the police, the courts, and the legal system are authorities entitled to make decisions and who should be deferred to concerning matters of criminal justice. Recent discussions of the police and policing, for example, emphasize the need for the police to create and maintain legitimacy (Skogan and Frydl, 2004; Skogan and Meares, 2004).

Police research suggests that the police cannot secure public compliance and cooperation without some sense within the public that they are entitled to make decisions and issue directives (Sunshine and Tyler, 2003; Tyler and Fagan, 2008). Similarly, court administrators have focused their efforts on designing court systems to gain and retain the trust and confidence of the public (Burke and Leben, 2008; Rottman, 2005, 2007, 2008). Therefore, the importance of understanding how individuals who deal with legal authorities experience their encounters is being more widely recognized, particularly how those experiences shape their judgments about the legitimacy of the police and the courts.

These efforts to study legitimacy consistently suggest that the central factor that shapes people's evaluations of the police and the courts is the fairness of the manner in which the authority was exercised—procedural justice (Tyler, 2007, 2008, in press). Legal authorities legitimate their decisions by making them through fair procedures, which leads people to be more willing to voluntarily accept those decisions (Tyler and Huo, 2002) and to follow them across time irrespective of whether their behavior is being monitored. Furthermore, when authorities act fairly, they create legitimacy and encourage general rule-following behavior in the everyday lives of people.

Direct correspondence to Tom R. Tyler, Department of Psychology, New York University, 6 Washington Place, Room 578, New York, NY 10003 (e-mail: tom.tyler@nyu.edu).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 127 12/21/09 7:14:16 PM

The same studies suggest that general judgments about the fairness of the authorities shape people's everyday compliance with the law as well as their willingness to cooperate with efforts to maintain social order in their communities. When the police and the courts are viewed as acting fairly, they are viewed as legitimate and enjoy public trust and confidence. These public judgments motivate supportive public behavior, which includes compliance and cooperation.

The question addressed by the Franke, Bierie, and MacKenzie (2010, this issue) article is whether these findings extend to the arena of corrections. The question is a crucial one, because America is a highly punitive society in which many members, particularly those who are poor or members of minority groups, have experiences with some type of correctional authority. In general, experiences in jail and prison have been viewed as undermining support for the law and legal authorities and as leading to heightened subsequent criminal activity (Bonta and Gendreau, 1990; Bukstel and Kilmann, 1980), which suggests that high levels of incarceration are antithetical to a legitimacy-based model of legal authority.

Franke et al.'s (2010) findings are hopeful because they suggest that experiences in incarceration can be positive or at least neutral if those incarcerated experience fairness from the authorities. If this suggestion is true, then the policy implication is clear; the design of institutions of incarceration needs to focus on the influence of inmate experiences on their judgments about the fairness of their experience. Jails and prisons need to develop policies designed to produce procedural fairness for inmates.

The argument that procedural justice matters in incarceration is not new. In a classic study of prisons, Sparks, Bottoms, and Hay (1996) argued that the justice experienced in prisons shapes inmate evaluations of the law and their behavior in prison. As they noted, "prisoners can hardly be expected to take the state's pretentions to moral authority and concern seriously unless the representations of that authority and those concerns, which they receive from the activities of prison administrators and staff, give them reason to do so" (1996: 308). In other words, the legitimacy of authorities is demonstrated by their use of fair procedures. This argument has been supported empirically in subsequent research on prisons (Gray, 2007), on plea bargaining hearings (Beloof, 2007; O'Hear, 2007), on drug courts (Gottfredson, Kearley, Najaka, and Rocha, 2007; McIvor, 2009), and on policing in general (Engel, 2005; Gau and Brunson, 2009; McCluskey, 2003; Reisig and Chandek, 2001; Shute, Hood, and Seemungal, 2005; Wells, 2007; Wiley and Hudik, 1974).

Franke et al.'s (2010) findings provide uneven empirical support for this argument. They suggest that the declines in legitimacy associated with prison do not occur in a boot camp environment where legitimacy remains constant. Unfortunately, they do not suggest that legitimacy increases in either of the programs they studied. This distinction between decline and increase in legitimacy is important because it touches on a recent controversy within the policing literature. Skogan (2006) agued that the police gain little by providing positive treatment, but are hurt by acting unfairly. Tyler and Fagan (2008) disagreed and demonstrated that both fair and unfair experiences significantly change views about legitimacy, although in different directions. So, the police can gain legitimacy by being fair and lose it by being unfair. Their conclusion recently

09012-CrimJournal-Guts.indd 128 12/21/09 7:14:16 PM

received independent confirmation in a study of personal experiences with the police conducted in London (Bradford, Jackson, and Stanko, 2009: 21), which concluded "positive contacts do contain the potential to improve aspects of trust and confidence." Therefore, legitimacy can increase through contacts with legal authorities, which leads us to wonder whether this also might happen in a correctional setting.

Two questions that both are relevant to corrections policy can be distinguished. The first is whether any negative imprisonment costs can be mitigated by procedural fairness during the experience. If prisons are breeding grounds for delegitimation and crime, then evidence that this effect can be mitigated is important. The findings of this study suggest that such mitigation can occur. It would be more valuable to demonstrate that fair treatment in prison can build legitimacy. One core argument of Tyler and Fagan (2008) is that the police can increase their legitimacy even while delivering negative outcomes if they deliver those outcomes through fair procedures—an argument untested in correctional settings. In either case, however, the implication is the same; emphasize fair procedures in correctional settings the same as in policing and in the courts. Doing so should minimize the negative implications of incarceration and maximize the likelihood of positive gains.

Enacting Fair Procedures in a Correctional Setting

Of course, the subsequent question is how to create procedural fairness. This study is not particularly helpful in this respect because it does not measure procedural fairness. Instead, Franke et al. (2010) report the opinion of observers that the inmates in prisons and boot camps experienced a different quality of "staff treatment," which led to different levels of trust and variations in "buy in to the regime." Although completely consistent with a procedural justice argument, the findings of this study do not distinguish staff efforts to provide valuable training or to restrict access to drugs or weapons from the fairness of staff treatment of inmates. Although all of these aspects of experiences with staff might potentially be important, efforts to design public policy need a framework within which to understand them.

A procedural justice framework for corrections would define the following four issues as the key to institutional design: voice, neutrality, treatment with respect and dignity, and trust in authorities.

The first, voice, refers to providing opportunities for inmates to participate in decision making, which means both providing formal mechanisms that allow inmates to have a meaningful say in the process of making policy decisions about their daily lives as well as providing opportunities in the context of everyday conflicts and decisions for inmates to state their case before decisions are made by the staff. This desire to participate in decision making by being able to voice concerns is one reason that informal dispute resolution mechanisms are popular. In such settings, people can state what they think the issues involved are and make suggestions for how they should be handled. Such opportunities for voice need not involve a formal or elaborate mechanism. Studies of street stops, for example, indicate that when officers provide

09012-CrimJournal-Guts.indd 129 12/21/09 7:14:16 PM

people an opportunity to tell their side of the story before they take action, people are much more likely to feel fairly treated (Tyler and Huo, 2002).

Neutrality refers to making decisions based on the consistent application of rules rather than on personal opinions or prejudices of authorities. A prison environment provides considerable opportunity for the capricious and arbitrary exercise of authority and for authorities to act based on personal prejudice. By acting based on rules and by applying those rules across people and time, authorities are viewed as acting fairly. Because rules typically are specified in correctional settings, the authorities have considerable capacity to shape and explain their actions by reference to the rules.

Treatment with respect and dignity consistently is one of the most important issues that people are concerned with when dealing with authorities. When people feel demeaned or subjected to negative stereotypes, they view themselves as diminished as people and disrespected beyond what is appropriate when dealing with the law. Conversely, acknowledging people's rights and acting with courtesy leads them to feel fairly treated.

Finally, people are influenced by their inferences about the motivations of the authorities with whom they deal. If people feel that those authorities are acting out of a sincere desire to do what is right, then they view the authorities as acting more fairly. Conversely, if people think that an authority is not concerned about their well-being, then they react negatively to their actions. How can authorities communicate trust? They can give people a chance to explain their concerns, show that what people say is being considered, and explain why and how decisions are made.

The exact manner in which the elements of fair treatment are enacted depends on the setting and will vary depending whether that setting involves the courts, the police, or corrections. For example, in the courts, judges have been encouraged to avoid actions that communicate disrespect, such as reading or signing orders when litigants are speaking and explaining the basis for their actions. For the police, model procedures involve presenting people with a written statement that specifies their rights and tells them how they can complain if they feel that unfair treatment has occurred, and which explains why actions such as street stops are occurring. General courtesy toward the people with whom they are dealing is another method for creating legitimacy.

In a correctional setting, where contact between authorities and inmates is more involved and longer term, several types of policy can be enacted to create more procedurally just correctional practices. A core contribution of the Franke et al. (2010) article is its suggestion that a larger set of issues might be involved in reactions of inmates to their experience. Beyond the quality of interpersonal treatment and the fairness of decision making, inmates might react to the degree that authorities help them learn meaningful skills and develop opportunities to enter the post-prison world with viable possibilities for a noncriminal life. They also might react to whether the guards create a safe and less dangerous world for them to live in as prisoners. Recognizing the importance of inmate feelings about their treatment to the legitimacy of authorities suggests a general need to examine the sources of those feelings. This point clearly includes traditional

09012-CrimJournal-Guts.indd 130 12/21/09 7:14:16 PM

procedural justice issues, but it also might extend beyond those issues to other sources of legitimacy that can be tapped to enhance the positive consequences of corrections.

Legitimacy as a General Concern

Underlying the arguments in Franke et al.'s (2010) article is the general suggestion that legitimacy needs to be a central focus when designing criminal justice institutions (for a similar argument, see Tyler, 2007, 2008). This focus on legitimacy highlights a broader change in how individuals in criminal justice view the relationship between people and legal authorities. Rather than focusing primarily on compliance and seeking to motivate compliance by threatening or by actually punishing people for wrongdoing, the focus is increasingly on gaining willing and voluntary cooperation with the police, courts, and law. This willingness includes both voluntary deference to the law and the decisions of legal authorities, and active and voluntary cooperation with efforts to manage communities. Viewing legal authorities as legitimate motivates such willing cooperation. When the authorities are viewed as legitimate, they are considered to be entitled to make decisions and to dictate people's behavior across different settings. Furthermore, they have the right to call on people for help to manage crime and social order within different communities—help in the form of joint voluntary actions such as joining a neighborhood watch or attending community meetings.

From this perspective, a larger policy question is the advisability of having a large incarceration system. The United States has one of the world's largest incarceration systems in proportion of the population in prison (Pew Center for the States, 2008), and prison is generally a delegitimating experience. Although one implication of the findings noted is to make prison consequences as positive as possible, a larger suggestion is that the policy of widespread prison use is questionable, which is the case because it is easiest to build legitimacy prior to experiences in prison (i.e., prevention is more effective than rehabilitation, although rehabilitation can be effective [Lipsey and Cullen, 2007]). Therefore, resources should be directed at changes in the justice system to develop effective alternatives to prison rather than toward changing the nature of the prison experience or toward post-prison rehabilitation. Of course, because some form of prison is necessary, working toward improving the quality of incarceration is clearly important.

Another policy implication is to highlight questionable policing practices. From a legitimacy perspective, practices that lead to increases in the number of people who have an incarceration experience should be avoided. One current practice of concern is zero-tolerance policing—the practice of stopping, citing, and sometimes incarcerating people for minor lifestyle offenses. This practice has expanded greatly in cities such as New York and Baltimore (Fagan, Geller, Davies, and West, 2009; Howell, 2009). From a legitimacy perspective, it is a questionable policy, which is likely to have the consequence of widely delegitimating the police, courts, and law among the general population.

09012-CrimJournal-Guts.indd 131 12/21/09 7:14:16 PM

If widespread street stops are to be a policing practice, then the question of how to engage in "smart stops" in which the police are fully trained in the importance of using fair procedures is central to policy. As Tyler and Fagan (2008) demonstrated, it is possible to engage in stops while building legitimacy if attention is paid to what happens during those stops. The broader point is that authorities should be focusing on policies that build positive experiences with the public, not on delivering sanctions and facilitating widespread incarceration.

Conclusion

Franke et al.'s (2010) study extends the arguments in favor of focusing on legitimacy to the corrections setting. Given the large number of people incarcerated in the United States and the negative long-term effects of their experiences in jail or prison, this focus is long overdue. Approximately 2.3 million people are currently incarcerated in the United States, with approximately 700,000 reentering our communities each year (not counting juveniles). Having a set of policies that address how to prevent prison from creating and releasing alienated prisoners is critical. The focus should be on policies that produce prisoners who will emerge as a positive benefit to themselves and their communities (Petersilia, 2003; Travis, 2000).

The procedural justice focus through which Franke et al. (2010) frame their findings and their implications fits well with their general suggestion that clear policies can be enacted to address the crippling effects of imprisonment. These policies need to emphasize the creation of fair procedures to manage people who are incarcerated and emphasize both fair decision making when developing and implementing these policies behind bars, as well as the need to treat everyone in such settings with dignity, courtesy, and respect. Every experience that people have with legal authorities (whether those authorities are the police, courts, or those in corrections) socializes both the people involved and the individuals who witness or hear about the actions of the authorities about the legitimacy of the law. A focus on policies that promote procedural fairness most likely would maximize that legitimacy and minimize violence and misconduct within prisons and subsequent reoffending.

References

Beloof, Douglas E. 2007. Dignity, equality, and public interest for defendants and crime victims in plea bargaining. *Marquette Law Review*, 91: 349–355.

Bonta, James A. and Paul Gendreau. 1990. Reexamining the cruel and unusual punishment of prison life. *Law and Human Behavior*, 14: 347–366.

Bradford, Ben, Jonathan Jackson, and Elizabeth A. Stanko. 2009. Contact and confidence: Revisiting the impact of public encounters with the police. *Policing and Society*, 19: 20–46.

Bukstel, Lee and Peter Kilmann. 1980. Psychological effects of imprisonment on confined individuals. *Psychological Bulletin*, 88: 469–493.

Burke, Kevin and Steve Leben. 2008. Procedural fairness: A key ingredient in public satisfaction. *Court Review*, 44: 4–25.

09012-CrimJournal-Guts.indd 132 12/21/09 7:14:16 PM

Engel, Robin S. 2005. Citizens' perceptions of distributive and procedural injustice during traffic stops with police. *Journal of Research in Crime and Delinquency*, 42: 445–481.

Fagan, Jeffrey, Amanda Geller, Garth Davies, and Valerie West. 2009. Street stops and broken windows revisited: The demography and logic of proactive policing in a safe and changing city. In (Stephen K. Rice and Michael D. White, eds.), *Race, Ethnicity, and Policing*. New York: New York University Press.

Franke, Derrick, David Bierie, and Doris L. MacKenzie. 2010. Legitimacy in corrections: A randomized experiment comparing a boot camp with a prison. *Criminology & Public Policy*. This issue.

Gau, Jacinta M. and Rod K. Brunson. 2009. Procedural justice and order maintenance policing: A study of inner-city young men's perceptions of police legitimacy. *Justice Quarterly*, 2009: 1–25.

Gottfredson, Denise C., Brook W. Kearley, Stacy S. Najaka, and Carlos M. Rocha. 2007. How drug treatment courts work. *Journal of Research in Crime and Delinquency*, 44: 3–35.

Gray, Jamael. 2007. *Fair Decision Making and Legitimacy: The Effects of Treatment on Compliance and Outcome Satisfaction in a Sample of Incarcerated Males*. Unpublished doctoral dissertation, University of Southern Illinois.

Howell, K. Babe. 2009. Broken lives from broken windows: The hidden costs of aggressive order-maintenance policing. *New York University Review of Law and Social Change*, 33: 271–326.

Lipsey, Mark and Francis T. Cullen. 2007. The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science*, 3: 297–320.

McCluskey, John D. 2003. *Police Requests for Compliance: Coercive and Procedural Just Tactics*. New York: LFB Scholarly Publishing.

McIvor, Gill. 2009. Therapeutic jurisprudence and procedural justice in Scottish drug courts. *Criminology and Criminal Justice,* 9: 29–49.

O'Hear, Michael M. 2007. Plea bargaining and victims. *Marquette Law Review*, 91: 323–347.

Petersilia, Joan. 2003. *When Prisoners Come Home: Parole and Prisoner Reentry*. Oxford, UK: Oxford University Press.

Pew Center for the States. 2008. *One in 100: Behind Bars in America*.

Reisig, Michael D. and Meghan S. Chandek. 2001. The effects of expectancy disconfirmation on outcome satisfaction in police–citizen encounters. *Policing*, 24: 88–99.

Rottman, David B. 2005. *Trust and Confidence in the California Courts*. San Francisco: Office of the California Courts.

Rottman, David B. 2007. Adhere to procedural fairness in the justice system. *Criminology & Public Policy*, 6: 835–842.

Rottman, David B. 2008. Procedural fairness as a court reform agenda. *Court Review*, 44: 32–35.

09012-CrimJournal-Guts.indd 133 12/21/09 7:14:17 PM

Shute, Stephen, Roger Hood, and Florence Seemungal. 2005. *A Fair Hearing? Ethnic Minorities in the Criminal Courts*? Devon, UK: Willan.

Skogan, Wesley G. 2006. Asymmetry in the impact of encounters with the police. *Policing and Society*, 16: 99–126.

Skogan, Wesley G. and Kathleen Frydl. 2004. *Fairness and Effectiveness in Policing: The Evidence*. Washington, DC: National Academy Press.

Skogan, Wesley G. and Tracey Meares. 2004. Lawful policing. *The Annals of the American Academy of Political and Social Science,* 593: 66–83.

Sparks, Richard, Anthony Bottoms, and Will Hay. 1996. *Prisons and the Problem of Order*. Oxford, UK: Clarendon.

Sunshine, Jason and Tom R. Tyler. 2003. The role of procedural justice and legitimacy in shaping public support for policing. *Law & Society Review*, 37: 555–589.

Travis, Jeremy. 2000. *But They All Come Home: Rethinking Prisoner Reentry*. Washington, DC: Urban Institute.

Tyler, Tom R. 2006a. Legitimacy and legitimation. *Annual Review of Psychology,* 57: 375–400.

Tyler, Tom R. 2006b. *Why People Obey the Law*. Princeton, NJ: Princeton University Press.

Tyler, Tom R. 2007. *Psychology and the Design of Legal Institutions*. Nijmegen, The Netherlands: Wolf.

Tyler, Tom R. 2008. Psychology and institutional design. *Review of Law and Economics,* 4: Article 6.

Tyler, Tom R. In press. Legitimacy and criminal justice: The benefits of self-regulation. *Ohio State Journal of Criminal Justice*.

Tyler, Tom R. and Jeffrey Fagan. 2008. Legitimacy and cooperation: Why do people help the police fight crime in their communities? *Ohio State Journal of Criminal Law*, 6: 231–275.

Tyler, Tom R. and Yuen J. Huo. 2002. *Trust in the Law*. New York: Russell-Sage.

Wells, William. 2007. Type of contact and evaluations of police officers: The effects of procedural justice across three types of police–citizen contacts. *Journal of Criminal Justice*, 35: 612–621.

Wiley, Mary Glenn and Terry L. Hudik. 1974. Police–citizen encounters. *Social Problems*, 22: 119–127.

Tom R. Tyler is a university professor at New York University who teaches in both the Department of Psychology and the School of Law. His research explores the dynamics of authority in groups, organizations, and societies. In particular, he examines the role of judgments about the justice of group procedures in shaping legitimacy, compliance, and cooperation. He is the author of several books, which include *The Social Psychology of Procedural Justice* (1988), *Social Justice in a Diverse Society* (1997), *Cooperation in Groups* (2000), *Trust in the Law* (2002), and *Why People Obey the Law* (2006).

09012-CrimJournal-Guts.indd 134 12/21/09 7:14:17 PM

EDITORIAL INTRODUCTION

PROBLEM-ORIENTED POLICING

Evaluating the effectiveness of problem-oriented policing

Michael S. Scott, *Senior Editor*
University of Wisconsin Law School

What works in policing? David Weisburd, Cody W. Telep, Joshua C. Hinkle, and John E. Eck (2010, this issue) sought to answer this core question in their review of problem-oriented policing. Herman Goldstein proposed the problem-oriented approach in 1979 as the means by which the police could achieve their objectives more effectively and, thereby, improve the overall police institution. Weisburd et al. set out to put the problem-oriented approach to the test some 30 years later. Given the ambitiousness of Goldstein's proposal, it follows that Weisburd et al.'s study is equally ambitious and important. If problem-oriented policing is proven to work, then it would stand to reason that police and the local governments should commit to the approach more fully. If it is proven not to work, then reconsideration of the approach—either to better understand why it is not working or to pursue alternative approaches to policing—would be warranted.

Weisburd et al.'s (2010) overall conclusion was that reliable evidence, albeit modest in strength and limited in quantity, indicates that problem-oriented policing is an effective approach to reducing crime and disorder. Accordingly, they recommend that police agencies and government funding agencies alike maintain or strengthen their commitment to the problem-oriented policing approach.

As important and interesting as Weisburd et al.'s (2010) conclusions are, how they reached them is nearly equally so. Their study—a meta-analysis of problem-oriented studies and interventions—was conducted under the auspices and guidelines of the Campbell Collaboration (an international research network dedicated to reviewing the knowledge basis of the efficacy of social interventions in the criminal justice, education, and social welfare realms). The research guidelines of the Campbell Collaboration strongly favor experimental and quasi-experimental studies that include comparison groups and research designs that generally are regarded as yielding more reliable inferences about causal relationships than alternative research designs

Direct correspondence to Michael S. Scott, University of Wisconsin Law School, 975 Bascom Mall, Madison, WI 53706 (e-mail: mscott@wisc.edu).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 135 12/21/09 7:14:17 PM

do. In the context of Weisburd et al.'s study, this process means proving that problem-oriented policing does or does not cause crime and disorder to decline. The research guidelines of the Campbell Collaboration permit the consideration of other research designs—which Weisburd et al. did—but such research designs are deemed less reliable and, therefore, inferior.

Given the rising interest in evidence-based practice among many professional fields—criminal justice included—the question of how we know what does and does not work is hugely interesting and important. Indeed, in a broader sense, few topics are more important or more interesting than what constitutes the basis for human knowledge. Weisburd et al.'s (2010) study stands then as a contribution to the age-old and perpetual exploration of the matter, albeit in the narrower context of modern policing.

As one reads Weisburd et al.'s (2010) article, the policy essays by Nick Tilley (2010, this issue) and Anthony A. Braga (2010, this issue), as well as this introduction, it is worth noting that whatever critiques are made of problem-oriented policing and of this study, all the writers in this volume undoubtedly would consider themselves among the leading proponents of the problem-oriented approach to policing. Accordingly, this volume constitutes less vehement debate among scholars with adversarial positions and viewpoints than a spirited exchange among scholars committed to problem-oriented policing. The theory and core principles of the problem-oriented approach to policing have been endorsed nearly universally by scholars; no scholar has argued that the approach is fundamentally flawed, which is remarkable in the scholarly world where debunking theories is the norm. Why problem-oriented policing has been spared such harsh criticism turns out to be important to understanding the discussion and critiques of this study.

Weisburd et al.'s (2010) study posited that it is logical and appropriate to evaluate the problem-oriented policing approach by determining the extent to which it can be proven that the approach has reduced crime and disorder. On this count, Weisburd et al. concluded in the (qualified) affirmative. To those of us heavily invested in problem-oriented policing, this result is indeed welcome. However, it is worth considering the possible implications if the conclusion of the study was negative—insufficient evidence that problem-oriented policing reduced crime and disorder. Would such a conclusion prove that the theory and principles of problem-oriented policing are wrong? Or would it prove that the application of the theory and principles in the cases studied was inadequate? Or would it prove that the metrics employed to detect an effect were flawed?

If one understands the problem-oriented approach to policing to be, in essence, a process theory about how police should seek to be effective rather than a substantive theory about what policing methods reduce crime and disorder, then it is easier to understand why the problem-oriented approach has received little criticism; it does not depend on its substantive outcomes to justify its processes, just as the scientific method is not judged on the specific outcomes of scientific experimentation. Indeed, the failure of the process to achieve specific outcomes reasonably could be understood as confirmation of the process theory's validity; it

09012-CrimJournal-Guts.indd 136 12/21/09 7:14:17 PM

merely proved what did not work—a finding that might prove useful in the continuing search for what does work.

My own view is that the problem-oriented approach to policing should be measured not on the results of each application but rather on the degree to which its process and principles improve the prospects for more effective policing, particularly as compared with alternative policing approaches. As to applying the theory and principles to each discrete problem the police seek to address, I concur with Weisburd et al. (2010), Tilley (2010), and Braga (2010) that it is logical and appropriate to subject each theory application to as rigorous an outcome evaluation as possible or as is practical.

Weisburd et al.'s (2010), Tilley's (2010), and Braga's (2010) arguments and critiques regarding the relative merits of applied research methodology in the policing context speak well for themselves, and I encourage the reader to give them careful consideration. As one reflects on the particulars of this study and its critiques, one must also bear in mind that the police institution and the police research institution alike remain in the relatively early stages of developing a robust and rigorous evaluation tradition. Much work remains to be done to improve the capacity of police to undertake problem-oriented policing and of evaluators (internal and external) to assess it properly. It is hoped that Weisburd et al.'s study as well as Tilley and Braga's policy essays will contribute to that goal.

References

Braga, Anthony A. 2010. Setting a higher standard for the evaluation of problem-oriented policing standards. *Criminology & Public Policy*. This issue.

Tilley, Nick. 2010. Whither problem-oriented policing. *Criminology & Public Policy*. This issue.

Weisburd, David, Cody W. Telep, Joshua C. Hinkle, and John E. Eck. 2010. Is problem-oriented policing effective in reducing crime and disorder? Findings from a Campbell systematic review. *Criminology & Public Policy*. This issue.

Michael S. Scott is a clinical associate professor at the University of Wisconsin Law School and the director of the Center for Problem-Oriented Policing. He currently chairs the judging committee for the Herman Goldstein Award for Excellence in Problem-Oriented Policing. Scott was formerly chief of police in Lauderhill, FL. He has served in various civilian administrative positions in the St. Louis, MI, Metropolitan; Ft. Pierce, FL; and New York City police departments and was a police officer in the Madison, WI, Police Department. He was a senior researcher at the Police Executive Research Forum (PERF) in Washington, DC. He received PERF's 1996 Gary P. Hayes Award for innovation and leadership in policing. Scott holds a law degree from Harvard Law School and a bachelor's degree from the University of Wisconsin—Madison.

09012-CrimJournal-Guts.indd 137 12/21/09 7:14:17 PM

RESEARCH ARTICLE

PROBLEM-ORIENTED POLICING

Is problem-oriented policing effective in reducing crime and disorder?

Findings from a Campbell systematic review

David Weisburd
Hebrew University
George Mason University

Cody W. Telep
George Mason University

Joshua C. Hinkle
Georgia State University

John E. Eck
University of Cincinnati

Research Summary

We conducted a Campbell systematic review to examine the effectiveness of problem-oriented policing (POP) in reducing crime and disorder. After an exhaustive search strategy that identified more than 5,500 articles and reports, we found only ten methodologically rigorous evaluations that met our inclusion criteria. Using meta-analytic techniques, we found an overall modest but statistically significant impact of POP on crime and disorder. We also report on our analysis of pre/post comparison studies. Although these studies are less

This project was supported by Award 2007-IJ-CX-0045, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice and the Nordic Campbell Centre. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the Department of Justice or the Nordic Campbell Centre. We would like to thank David B. Wilson for his assistance with our effect size calculations and his comments on an earlier version of this article, Lorraine Green Mazerolle and Anthony A. Braga for data from their systematic reviews, and Charlotte Gill and the anonymous reviewers from the Campbell Collaboration and those from *Criminology & Public Policy* for their helpful comments. Direct correspondence to David Weisburd, Administration of Justice Department, George Mason University, 10900 University Boulevard, MS 4F4, Manassas, VA 20110 (e-mail: dweisbur@gmu.edu); Cody W. Telep, Administration of Justice Department, George Mason University, 10900 University Boulevard, MS 4F4, Manassas, VA 20100 (e-mail: ctelep@gmu.edu); Joshua C. Hinkle, Georgia State University, Department of Criminal Justice, PO Box 4818, Atlanta, GA 32302 (e-mail: jhinkle@gsu.edu); John E. Eck, University of Cincinnati, Department of Criminal Justice, PO Box 210389, Cincinnati, OH 45221 (e-mail: john.eck@uc.edu).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 139 12/21/09 7:14:17 PM

methodologically rigorous, they are more numerous. The results of these studies indicate an overwhelmingly positive impact from POP.

Policy Implications

POP has been adopted widely across police agencies and has been identified as effective by many policing scholars. Our study supports the overall commitment of police to POP but suggests that we should not necessarily expect large crime and disorder control benefits from this approach. Moreover, funders and the police need to invest much greater effort and resources to identify the specific approaches and tactics that work best in combating specific types of crime problems. We conclude that the evidence base in this area is deficient given the strong investment in POP being made by the government and police agencies.

Keywords

problem-oriented policing, Campbell systematic review, police effectiveness, meta-analysis

In a *Crime & Delinquency* article from 1979, Herman Goldstein critiqued police practices of the time by noting that they were more focused on the "means" of policing than its "ends." His critique drew from a series of recently completed studies, which suggested that such standard policing practices as "preventive patrol" (Kelling, Pate, Dieckman, and Brown, 1974) or "rapid patrol car response to calls for service" (Kansas City Police Department, 1977) had little impact on crime. Goldstein suggested that the research evidence was not idiosyncratic but reflected a more serious crisis in policing. To illustrate his concern, he referred to a newspaper article in the United Kingdom that reported on bus drivers in a small city who were driving by bus stops waving and smiling but failing to pick up passengers. When questioned by a reporter, a representative for the bus company responded that "it is impossible for the drivers to keep their timetable if they have to stop for passengers" (Goldstein, 1979: 236). Goldstein argued that, like these bus drivers, the police had become so focused on issues such as staffing and management that they had begun to ignore the problems that policing was meant to solve. Goldstein saw this dysfunction as the heart of the inability of police to be effective.

Goldstein (1979) called for a paradigm shift in policing that would replace the primarily reactive, incident-driven "standard model of policing" (National Research Council [NRC], 2004; Weisburd and Eck, 2004) with a model that required the police to be proactive in identifying underlying problems that could be targeted to alleviate crime and disorder at their roots. He termed this new approach "problem-oriented policing" (POP) to accentuate its call for police to focus on problems and not on the everyday management of police agencies. Goldstein also expanded the traditional mandate of policing beyond crime and law enforcement. He argued that the police had to deal with an array of problems in the community, which not only includes crime but also social and physical disorders. He also called for police to expand the tactics of policing beyond the law enforcement powers that were perceived as the predominant tools of

09012-CrimJournal-Guts.indd 140 12/21/09 7:14:18 PM

the standard model of policing. In Goldstein's view, the police needed to draw on not only the criminal law but also on civil statutes and rely on other municipal and community resources if they were to ameliorate crime and disorder problems successfully.

Goldstein's (1979) model was elaborated and extended by John E. Eck and William Spelman (1987) who drew on Goldstein's idea to create a straightforward model for implementing POP, which has become widely accepted. In an application of problem solving in Newport News in which Goldstein acted as a consultant, they developed the SARA model for problem solving, which is an acronym that represents four steps that they suggest police should follow when implementing POP. "Scanning" is the first step and involves the police identifying and prioritizing potential problems in their jurisdiction. After the potential problems have been identified, the next step is "analysis," which involves the police thoroughly analyzing the identified problem(s) using several data sources so that appropriate responses can be developed. The third step, "response," has the police developing and implementing interventions designed to solve the problem(s). Finally, once the response has been administered, the final step is "assessment," which involves evaluating the impact of the response.

POP has emerged as one of the most widely accepted and used strategies in U.S. policing. This popularity is indicated by the adoption of POP by major federal agencies and national policing groups, the creation of national awards for effective POP programs, and the widespread adoption of the approach in U.S and international policing. For example, the U.S. federal agency, the Office of Community-Oriented Policing Services (COPS), adopted POP as a key strategy and funded the Center for Problem-Oriented Policing (popcenter.org) and developed more than 50 problem-specific guides for police. The Police Executive Research Forum adopted POP as a "powerful tool in the policing arsenal" in the 1980s and began to run a yearly national conference to promulgate and advance POP strategies (Solé Brito and Allan, 1999: xiii). In 1993, the Herman Goldstein Award was created for "problem solving excellence," and since its inception, more than 800 submissions have been sent in from around the world. In the United Kingdom, the Tilley Award for POP was created in 1999 and has since received almost 600 submissions. Reflecting the wide-scale adoption of POP by U.S. police agencies, the 2003 Law Enforcement Management and Administrative Statistics (LEMAS) survey reported that 66% of local police agencies (more than 100 officers) claimed to be using POP tactics (Bureau of Justice Statistics, 2006).

Despite this widespread adoption of POP, no effort has been made to review the research on POP systematically and to assess whether its wide adoption is merited by the scientific evidence available. This delay is not to say that scholars have not reviewed the research evidence regarding POP as part of broader reviews of the effectiveness of policing strategies. The NRC, for example, in its review of police practices and policies, noted that "[t]here is a growing body of research evidence that problem-oriented policing is an effective approach" (NRC, 2004: 243; see also Weisburd and Eck, 2004). Nonetheless, the narrative review by the NRC and more general reviews of policing effectiveness have not tried drawing studies together to gain an overall statistical portrait of what is known about the effectiveness of this approach.

09012-CrimJournal-Guts.indd 141 12/21/09 7:14:18 PM

In this article, we synthesize the extant empirical evidence (published and unpublished) on the effects of POP on crime and disorder. We seek to go beyond prior studies in two ways. First, our review takes a much more comprehensive approach to identifying POP studies than prior narrative reviews. Second, we summarize knowledge about prior studies using meta-analytic statistical approaches (Lipsey and Wilson, 2001) and do not simply rely on counting the number of studies that reach a specific threshold of evidence, which has been common in prior reviews (the "vote counting approach"). Our findings are both surprising and challenging given the widespread popularity of POP. Despite the wide adoption of POP, relatively few high-quality studies allow us to comment on the effectiveness of this strategy in reducing crime and disorder. In turn, although the overall summary effect sizes we identify suggest that POP is effective, the effect sizes reached in our analyses are relatively modest. We conclude that the evidence base in this area is deficient given the strong POP investment that is being made by the government and police agencies.

The Evidence Base for Problem-Oriented Policing

Several studies going back to the mid-1980s demonstrate that problem solving can reduce fear of crime (Cordner, 1986), violent and property crime (Eck and Spelman, 1987), firearm-related youth homicide (Kennedy, Braga, Piehl, and Waring, 2001), and various forms of disorder, which include prostitution and drug dealing (Capowich and Roehl, 1994; Eck and Spelman, 1987; Hope, 1994). For example, a study of Jersey City, New Jersey public housing complexes concluded that police problem-solving strategies caused measurable declines in reported violent and property crime, although the results varied across the six housing complexes studied (Mazerolle, Ready, Terrill, and Waring, 2000). In another example, Clarke and Goldstein (2002) reported a reduction in thefts of appliances from new home construction sites after careful analysis of this problem by the Charlotte–Mecklenburg Police Department and changes in building practices by construction firms were implemented.

Two experimental evaluations of problem-solving applications in crime hot spots (Braga, Weisburd, Waring, Mazerolle, Spelman, and Gajewski, 1999; Weisburd and Green, 1995) have been cited often in support of POP approaches (e.g., see NRC, 2004).[1] In a randomized trial involving Jersey City, New Jersey violent crime hot spots, Braga et al. (1999) reported reductions in property and violent crime in the treatment locations. Although this study tested problem-solving approaches, it is important to note that focused police attention was brought

1. A systematic review of "hot spots policing" has been conducted by Anthony A. Braga (2001, 2007). Hot spots policing focuses on small geographic areas and concentrations of crime. Hot spots policing per se does not demand detailed analysis of the problem identified and often relies on a law enforcement response. POP can focus on small geographic areas (hot spots); however, more analysis is undertaken to determine the causes of crime in the hot spots, and responses are tailored to the needs of each hot spot. Furthermore, POP also examines nongeographic concentrations of crime—repeat offenders, repeat victims, hot products, and so forth. In short, although problem-solving at hot spots sometimes can be a type of POP, many hot spots policing programs do not use the more systematic methods associated with this strategy.

09012-CrimJournal-Guts.indd 142 12/21/09 7:14:18 PM

only to the experimental locations. Accordingly, it is difficult to distinguish between the effects of bringing focused attention to hot spots and of such focused efforts being developed with a problem-oriented approach. The Jersey City Drug Market Analysis Experiment (Weisburd and Green, 1995) provides more direct support for the added benefit of the application of problem-solving approaches in hot spots policing. In that study, a similar number of narcotics detectives were assigned to treatment and control hot spots. Weisburd and Green (1995) compared the effectiveness of unsystematic arrest-oriented enforcement based on ad hoc target selection (the control group) with a treatment strategy that involved an analysis of assigned drug hot spots followed by site-specific enforcement and collaboration with landlords and local government regulatory agencies. The study concluded with monitoring and maintenance for up to a week after the intervention. Compared with the control drug hot spots, the treatment drug hot spots fared better regarding disorder and disorder-related crimes.

As noted, past narrative reviews have concluded that research is supportive of the capability of problem solving to reduce crime and disorder (e.g., NRC, 2004; Weisburd and Eck, 2004). In turn, evidence of the effectiveness of situational and opportunity-blocking strategies, although not necessarily police-based, provides indirect support for the effectiveness of problem solving in reducing crime and disorder. POP has been linked to routine activity theory, rational choice perspectives, and situational crime prevention (Clarke, 1992a, 1992b; Eck and Spelman, 1987). Recent reviews of prevention programs designed to block crime and disorder opportunities in small places find that most studies report reductions in target crime and disorder events (Eck, 2002a; Poyner, 1981; Weisburd, 1997). Furthermore, many of these efforts were the result of police problem-solving strategies. At the same time, it is important to note that many of the studies reviewed employed relatively weak evaluation designs (Clarke, 1997; Eck, 2002a; Weisburd, 1997).

Methods

We set out to develop a Campbell systematic review of POP. Campbell reviews require a transparent and systematic search-and-analysis strategy that involves a methodological and substantive review of the project at both the proposal stage and before final reports are completed.[2] Our main research question is whether POP is effective in reducing crime and disorder. Originally, we hoped to use meta-analysis to examine additional questions that would have shed important light on the nature of problem solving, which included a review of whether different types of problem solving had differential effects on crime and disorder and whether specific types of crime or disorder seem more amenable to problem-solving approaches. Unfortunately, the number of studies that met our inclusion criteria was not large enough to examine these questions statistically.[3]

2. See campbellcollaboration.org/artman2/uploads/1/Review_Steps.pdf for a description of the steps included in a Campbell review.

3. We also wanted to examine questions of cost effectiveness in our review. However, none of the studies we examined provided data on cost-effectiveness issues.

09012-CrimJournal-Guts.indd 143 12/21/09 7:14:18 PM

As our review of the literature makes clear, departments using POP have applied a diverse group of tactics to ameliorate several problems. As such, it is important to note that we are examining the effectiveness of a process used by the police to develop tactics, not a particular police tactic. For our purposes, the method used to develop the intervention is the treatment. Bradford Hill (1962: 11), a pioneer in experimental medical research, recognized 50 years ago that it is often necessary to evaluate treatment approaches rather than specific regimens. He commented when explaining that such evaluations were appropriate that "one man's meat is another man's poison," and that it was often necessary to individualize treatment based on the specific characteristics of a patient and his or her illness. The studies we examined differ greatly in the problems addressed and the solutions implemented. Nonetheless, they share the common thread of using a problem-oriented approach. Our review provides an assessment of whether that approach is effective in reducing crime and disorder problems.

We recognize that crime reduction is just one of an array of outcomes potentially associated with POP that range from increasing departmental accountability to enhancing citizen satisfaction, but it is an outcome of great interest for scholars and practitioners. As Scott (2000: 131) noted, one way to answer the question "how will we know if problem-oriented policing works?" is to "search for proof that the problem-solving methodology reduces crime and disorder." Bullock and Tilley (2003) edited a volume titled *Crime Reduction and Problem-Oriented Policing*, and Braga (2002) details the impact of several POP interventions in his book, *Problem-Oriented Policing and Crime Prevention*. POP is a strategy designed, in part, to address crime and disorder, and as a result, we think it valid and necessary to assess the prospects of using this strategy to reduce such problems.

Criteria for Inclusion and Exclusion of Studies in the Review

There is no hard rule for determining when studies provide more reliable or valid results, or any clear line to indicate when there is enough evidence to come to an unambiguous conclusion regarding the effectiveness of an intervention (e.g., see Braga, 2002; Scott, 2000; Weisburd and Eck, 2004). Following standard protocols in the Campbell Collaboration, the scope of our main review is experimental and quasi-experimental studies that include comparison groups. Such research designs are assumed "to minimize the likelihood that the effects of interventions will be confused with the effects of biases and chance" (Chalmers, 2003: 22). In technical terms, the designs are assumed to have high internal validity (Farrington and Petrosino, 2001).

Randomized experiments are assumed to have the highest internal validity and allow for the strongest causal statements about the effects of interventions (see Boruch, Snyder, and DeMoya 2000; Cook and Campbell, 1979; Farrington, 1983, 2003; Shadish, Cook, and Campbell, 2002; Weisburd, 2003; Weisburd, Lum, and Petrosino, 2001). This is the case because subjects are randomized into treatment and control conditions, and thus, their inclusion in one group or another is determined simply by chance. Absent randomization of treatments, it generally is assumed that the identification of well-matched comparison groups provides a high level of internal validity for a study (Farrington, Gottfredson, Sherman, and Welsh, 2002). Quasi-

09012-CrimJournal-Guts.indd 144 12/21/09 7:14:18 PM

experimental studies with comparison groups could be developed by identifying similar areas to those that are subject to interventions or by matching subjects on known characteristics such as gender, age, race, and prior record. Such studies generally are considered more likely to rule out preexisting differences between treatment and control conditions than quasi-experimental studies without comparison groups (Farrington et al., 2002).

Although in our review we rely strongly on these general assessments of the ability of studies to make causal statements with high internal validity, we also recognize that other criteria are important to assess the strength of research. Several scholars have suggested that the results of randomized field experiments can be compromised by the difficulty of implementing such designs (Clarke and Cornish, 1972; Eck, 2002b; Pawson and Tilley, 1997). Accordingly, in assessing the evidence, we also take into account the quality of the implementation of the research design and recognize that a reliance on "internal validity" should not preclude a concern with "external validity"—the ability of the findings of a study to be generalized beyond the sample studied (Farrington, 2003; Shadish et al., 2002). Indeed, the identified studies have strong limitations because they are carried out in specific jurisdictions using specific tactics.

We also think it important to note that our reliance on Campbell Collaboration criteria to assess study quality does not mean that other methods for evaluating POP are not useful. Indeed, we think that qualitative and ethnographic studies are particularly important if we are to understand the processes that underlie POP and "why" it could impact crime and disorder problems (e.g., see Bazemore and Cole, 1994; Bullock, Erol, and Tilley, 2006; Cordner and Biebel, 2005; Rubenser, 2005). The fact that such studies are not included in our review does not mean that we cannot learn more generally from knowledge generated using such approaches. Rather, we follow an analytic framework that allows us to draw a specific set of conclusions regarding what is known about the effectiveness of POP based on experimental and quasi-experimental evaluations.

We recognize that experimental studies, and even quasi-experimental studies with comparison groups, could be difficult to implement in many POP interventions (Eck, 2002b). In particular, specific problems addressed by the police might be unique, and thus, it might be difficult to identify a reasonable comparison condition. Several POP experts who were contacted in the study identification stage of our research (see below) suggested that a review that ignores pre–post studies without control groups would miss many POP evaluations. Although we have strong concerns regarding the methodological rigor of such studies, we did identify and analyze them separately from our main analyses.[4]

4. Designs that simply examine crime trends across time are assumed to be highly vulnerable to mistaking historical trends with program impacts (Campbell and Russo, 1999; Cook and Campbell, 1979). For example, during a crime decline in a city, a research design that tracks crime across time will conclude mistakenly that a program is successful when the data examined simply reflect secular trends. Studies that rely only on statistical controls—generally termed nonexperimental or observational designs—often are perceived to lead to the weakest level of internal validity (Cook and Campbell, 1979; Sherman, Gottfredson, MacKenzie, Eck, Reuter, and Bushway, 1997). Although we recognize that observational research designs that reflect strong theoretical modeling also might achieve high levels of internal validity for assessing outcomes (see Heckman and Smith, 1995), we did not identify any such studies in our review of POP programs.

09012-CrimJournal-Guts.indd 145 12/21/09 7:14:18 PM

The eligibility criteria for our main analyses were as follows:

1. The study must be an evaluation of a POP intervention. For this review, only police interventions following the basic tenets of the SARA model will be eligible for inclusion. Such interventions must involve the identification of a problem believed to be related to crime and disorder outcomes, the development and administration of a response specifically tailored to this problem, as well as an assessment of the response effects on a crime or disorder outcome.[5]
2. The study must include a comparison group that did not receive the treatment condition (POP).[6]
3. The study must report on at least one crime/disorder outcome including sufficient quantitative data to calculate an effect size.
4. The study must deal with problem areas or problem people.

Search Strategy for Identification of Relevant Studies

Several strategies were used to perform an exhaustive search for literature fitting the eligibility criteria. First, a keyword search was performed in an array of online abstract databases.[7] Second, we reviewed the bibliographies of past reviews of POP.[8] Third, we performed forward searches for works that have cited seminal POP studies.[9] Fourth, we performed hand searches of leading journals in the field.[10] Fifth, we searched the publications of several research and professional

5. We did not require that a study specifically note that it used the SARA model, but rather that it followed these steps more generally.
6. We only included studies that had a comparison group with some demonstration of equivalence to the treatment group.
7. The databases searched were as follows: Criminal Justice Periodical Index, Criminal Justice Abstracts, National Criminal Justice Reference Services Abstracts, Sociological Abstracts, Social Science Abstracts, Social Science Citation Index, Dissertation Abstracts, Government Publications Office Monthly Catalog, Police Executive Research Forum database of problem-oriented policing examples (POPNet), C2 SPECTR (The Campbell Collaboration Social, Psychological, Educational and Criminological Trials Register), Australian Criminology Database (CINCH), and Centrex (Central Police Training and Development Authority) U.K. National Police Library. We identified 5,282 publications using our set of keywords on the 12 online databases. We narrowed this list by reviewing titles and abstracts and by removing any studies not related to policing, not in English, duplicates, and book reviews. In an effort to ensure we were not missing any key studies published in other languages, we did examine non-English studies that cited Goldstein (1979) or Goldstein (1990) on Google Scholar. After translating titles and/or abstracts, we determined that none of these studies met our inclusion criteria.
8. These bibliographies were as follows: Braga, 2002; Mazerolle and Ransley, 2005; Mazerolle, Soole, and Rombouts, 2005; NRC, 2004.
9. The seminal pieces used were as follows: Braga et al., 1999; Eck and Spelman, 1987; Goldstein, 1979, 1990; and Spelman and Eck, 1987.
10. These journals were as follows: Criminology, Criminology & Public Policy, Justice Quarterly, Journal of Research in Crime and Delinquency, Journal of Criminal Justice, Police Quarterly, Policing, Police Practice and Research, British Journal of Criminology, Journal of Quantitative Criminology, Crime & Delinquency, Journal of Criminal Law and Criminology, and Policing and Society. Hand searches covered from 1979 to 2006.

09012-CrimJournal-Guts.indd 146 12/21/09 7:14:19 PM

agencies.[11] Sixth, after finishing these searches and reviewing the studies, we e-mailed the list of studies meeting our eligibility criteria in June 2007 to 62 leading policing scholars and practitioners knowledgeable in the area of POP and asked them whether they were aware of additional studies not on our list (see Weisburd, Telep, Hinkle, and Eck, 2008: Appendix B). Finally, we consulted with an information specialist at the outset of our review and at points along the way to ensure that we used appropriate search strategies. Our initial searches of electronic databases were conducted in the fall of 2006. By the completion of our systematic search in the fall of 2007, we had identified ten studies that met all of our eligibility criteria. We coded each of these studies using a standard data collection instrument that identified both substantive and methodological variables (see Weisburd et al., 2008: Appendix A).

Although it is not uncommon in Campbell reviews to find only a few studies regarding a specific practice, the absence of a wide body of evidence for POP is particularly concerning. POP represents a broad array of strategies applied to a broad array of problems. The development of systematic knowledge for policing accordingly requires that an equally broad array of studies is available that would allow us to assess what types of strategies are effective, in what types of circumstances, and for what types of crime. Additionally, this omission of systematic studies using rigorous research methods is particularly troubling given the widespread adoption of POP in the United States and elsewhere.

One explanation for the relatively few studies that met the methodological criteria of our review might be that much evaluation of POP has used weaker research designs. Accordingly, we also identified 45 before–after evaluations of POP in our review. These studies will be analyzed separately and discussed in greater detail.

Characteristics of Studies in our Main Analyses

Detailed information that compares the studies on the problems addressed, use of the SARA model, responses, and evaluation design is provided in Table 1. The ten eligible studies come from eight different U.S. cities (Jersey City was the site for two studies) and six wards in the United Kingdom. Four of the eligible studies were randomized experiments, and six were quasi-experiments with a comparison group. The randomized experiments were all place-based interventions as were four of the six quasi-experiments. The two person-based interventions focused on probationers and parolees in Knoxville and San Diego.

The interventions covered various problems and demonstrated the wide applicability of POP. Two interventions dealt with reducing probationer–parolee recidivism, two targeted drug

11. The publications of the following groups and agencies were searched: Center for Problem-Oriented Policing (Tilley Award and Herman Goldstein Award submissions, Problem-Specific Guides for Police), Institute for Law and Justice, Community Policing Consortium (electronic library), Vera Institute for Justice (policing publications), RAND (public safety publications), Police Foundation, Home Office (United Kingdom), Australian Institute of Criminology, Swedish Police Service, Norwegian Ministry of Justice and the Police, Royal Canadian Mounted Police, Finnish Police (Polsi), Danish National Police (Politi), The Netherlands Police (Politie), and New Zealand Police.

09012-CrimJournal-Guts.indd 147 12/21/09 7:14:19 PM

09012-CrimJournal-Guts.indd 148

TABLE 1

SARA Characteristics and Research Design for Eligible Studies

Study	Problem	Scanning and Analysis	Treatment/Response	Research Design and Units
Baker and Wolfer (2003)	Park with alcohol use, drug use, and vandalism	Physical survey of the park, crime prevention surveys, and crime mapping	Target hardening, proactive patrol, curfew law, removed pay phone used for drug deals, and crime newsletter	Quasi-experiment—survey of 250 residents living near the park compared with a sample of 670 town residents
Braga et al. (1999)	Hot spots of violent crime (e.g., street fighting, robbery, and assault)	Computerized mapping used to create hot spots; officers completed report on problems	A tailored solution to meet the problems observed during analysis; responses varied, but all included aggressive order maintenance	Randomized experiment—12 hot spots receiving POP compared with 12 matched hot spots receiving normal patrol
Knoxville P.D. (2002)	Probationers frequently rearrested	Review of crime and probation revocation data with Tennessee Board of Probation & Parole	Collaboration of police, parole, and service providers to develop team supervision and treatment plan	Quasi-experiment—265 probationers in the program compared with a historical sample of 261 probationers
Mazerolle et al. (2000)	Drugs and disorder at nuisance locations	Beat Health team visited site, conducted physical survey, and worked with place managers	Tried to develop working relationship with property owners and could use team of city inspectors and civil law	Randomized experiment—50 Beat Health hot spots compared with 50 referred sites that received normal patrol
Sherman et al. (1989)	High numbers of calls at commercial and residential addresses	Call logs used to generate highest call addresses; officers diagnosed the problem and developed an action plan	Wide variation in strategies used by RECAP team; residential strategies often focused on helping landlords with problem tenants	Randomized experiment—comparing commercial (119 pairs) and residential (107) addresses that received POP with control addresses
Stokes et al. (1996)	Student violent victimization occurring on the way to school	Student focus groups and initial victimization survey used to map student-identified problem areas	Creation of a Safe Corridor—7–9 police officers patrolled a 10 × 3 block area from 8 to 9 a.m. and 2:30 to 4 p.m. with bikes, cars, and on foot	Quasi-experiment victim survey—414 target school students compared with 1,681 students at nearby schools
Stone (1993)	Drugs in public housing projects	Management team of police and housing authority conducted resident survey and meetings with police officers and investigators	Focused on improving lighting, abandoned cars, trash/litter, playground equipment, and poorly placed clotheslines to address problems associated with drugs	Quasi-experiment victim survey—149 residents of 2 target housing projects compared with 135 residents of 2 similar housing projects

12/21/09 7:14:19 PM

Thomas (1998)	High rearrest rates of juvenile probationers	Recognition that juvenile supervision was inadequate; examined crime and arrest data	Police/probation collaboration to increase community-based supervision, mentoring, and program referral	Quasi-experiment—80 program probationers compared with a historical sample of 80 probationers
Tuffin et al. (2006)	Varied by ward—all included antisocial behavior	Planning stages: Research, engage, public preferences, investigation and analysis, and public choices	Varied by site, but included increasing police presence and developing a targeted response with community stakeholders	Quasi-experiment—6 sites (neighborhoods in the United Kingdom) matched with comparison areas
Weisburd and Green (1995)	Drug and drug-related disorder	Step-wise process "planning stage" collecting data on the characteristics of the place using crime maps and community meetings	"Implementation stage" coordinated crackdown and use of government resources; "maintenance stage" ensured drug activity remained under control	Randomized experiment—28 hot spots receiving treatment compared with 28 hot spots receiving normal drug area patrol

09012-CrimJournal-Guts.indd 149 12/21/09 7:14:19 PM

markets, one responded to vandalism and drinking in a park, one combated crime in hot spots of violence, one addressed school victimization, two tackled problem addresses, and one targeted overall crime. These interventions also used diverse approaches to address crime and disorder. Of the ten eligible studies, eight reported findings in favor of POP, although those effects varied widely. In Table 2, we provide a summary of results for each eligible study.

Results: Meta-Analysis of the Impact of POP on Crime and Disorder

We completed a meta-analysis of the ten eligible studies to examine the standardized effect size for each study and to calculate an overall random effect for the impact of POP on crime and disorder.[12] Computation of effect sizes in the studies was not always direct. The goal was to convert all observed effects into a standardized mean-difference effect size metric. None of the studies we examined calculated standardized effect sizes, and indeed, it was sometimes difficult to develop precise effect size metrics from published materials.[13] This difficulty reflects a more general problem in crime and justice with "reporting validity" (Farrington, 2006; Lösel and Köferl, 1989; Perry and Johnson, 2008; Perry, Weisburd, and Hewitt, in press).

One problem in conducting meta-analyses in crime and justice is that investigators often do not prioritize the outcomes examined. This trend is common in social science studies in which authors view good practice as reporting all relevant outcomes. However, the lack of outcome prioritization in a study raises the question of how to derive an overall effect of treatment. For example, reporting one significant result might reflect a type of "creaming" in which the authors focus on one significant finding and ignore the less positive results of other outcomes. But authors commonly view the presentation of multiple findings as a method for identifying the specific contexts in which the treatment is effective. When the number of such comparisons is small and, therefore, unlikely to affect the error rates for specific comparisons, such an approach is often valid.

A primary outcome is defined in our review as a major focus of the POP intervention. The police needed to be targeting the crime or call type specifically for us to identify an outcome as primary. For example, in the Mazerolle et al. (2000: 220) study, the authors note that the

12. We used Biostat's Comprehensive Meta Analysis Program for our analyses to create the forest plots we will present later (see Borenstein, Hedges, Higgins, and Rothstein, 2009).

13. For the two probation studies (Knoxville Police Department, 2002; Thomas, 1998) and the Stokes, Donahue, Caron, and Greene (1996) study, we used the proportion of successes (or failures) to calculate an effect size. These calculations used the odds ratio method. For the Stone (1993) study, we used the difference in pre to post mean change between the treatment and comparison sites (see Weisburd et al., 2008, for details regarding computation of effect sizes), sample size, and the t statistic value from a paired group t test examining factor scores on a victimization survey. In the case of Weisburd and Green (1995), we calculated effect sizes from exact p values from the F tests used in the two-way analysis of variance calculations for service data calls. For Sherman, Buerger, and Gartin (1989), we used the chi-square values comparing the difference in calls for service at RECAP wth control targets before and after the intervention. We could find no satisfactory method for conversion of data from Braga et al. (1999) and, therefore, converted the estimates to an odds ratio following the method outlined in the Appendix of Farrington, Gill, Waples, and Argomaniz (2007). We think it important to note that this method is clearly conservative for estimating effect sizes in this case. We also used the odds ratio method for the Baker and Wolfer (2003) study, the Mazerolle, Price, and Roehl (2000) article, as well as the Tuffin, Morris, and Poole (2006) report.

09012-CrimJournal-Guts.indd 150 12/21/09 7:14:19 PM

Beat Health program "uses a variety of tactics to resolve drug and disorder issues." The authors present data on calls for service for disorder, drug crime, property crime, and violent crime. Because of this description of the intervention, we chose to include only drug and disorder calls as primary outcomes.

When several studies use similar outcome measures, it is possible to make comparisons across studies of outcomes for specific measures (e.g., specific types of crimes). In our review such an approach is not possible because the types of interventions and the types of crimes vary widely as noted. Accordingly, we analyze the studies using two approaches. The first is conservative because it combines all primary outcomes reported into an overall average effect size statistic. The second represents the largest effect reported in the studies and gives an upper bound to our findings. It is important to note that in some studies with more than one outcome reported, the largest outcome reflected what authors thought would be the most direct program effect. This point was true for the Jersey City Drug Market Analysis Experiment, which examined violent and property crimes but assumed that the largest program effects given the intervention would be found in the case of calls for disorder (Weisburd and Green, 1995).

TABLE 2

Crime/Disorder Outcomes for Eligible Studies

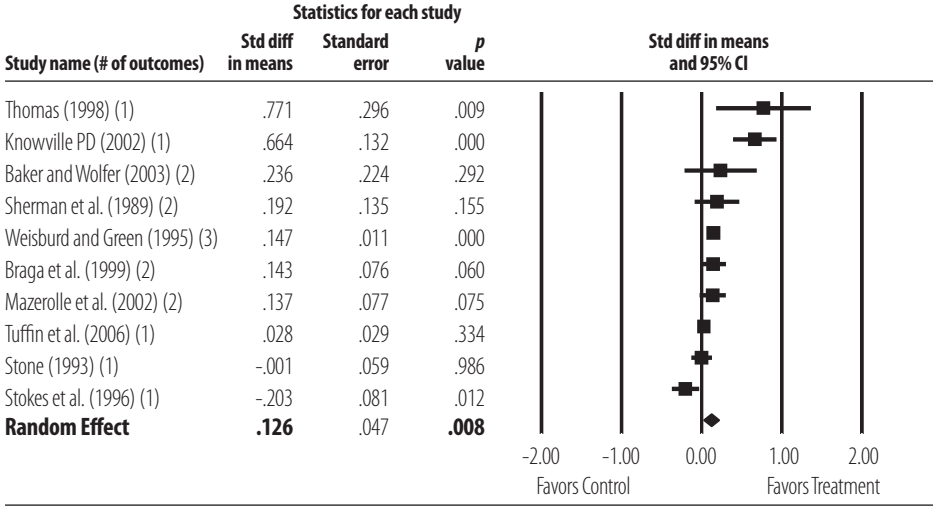
Study	Crime/Disorder Outcomes	Other Outcomes
Baker and Wolfer (2003)	Reduction in perceptions of crime problem in target group compared with comparison area	Target group more likely to see officers on patrol and report a fear reduction
Braga et al. (1999)	Significant decline in total criminal incidents and calls for service in treatment compared with control hot spots	Social and physical disorder declined in 10 of the 11 treatment hot spots
Knoxville P.D. (2002)	29% in program succeeded (completed parole without revocation) compared with only 11% success in comparison group	None
Mazerolle et al. (2000)	Significant decrease in experimental group drug calls compared with control group but no difference for disorder, violence, or property calls	None
Sherman et al. (1989)	Small decrease in calls in treatment residential addresses compared with control, but no difference in commercial addresses	None
Stokes et al. (1996)	Victimization rate in the target school increased, whereas significantly decreasing at the control schools	Percentage of students afraid of an attack increased at the test school and decreased at the control schools
Stone (1993)	Rate of being asked to buy or sell drugs increases more in the intervention than in the comparison area	None
Thomas (1998)	Those in the C.A.N. program had .25 the recidivism rate of a random group of those not selected for the program	Individuals in C.A.N. were more likely to complete probation conditions
Tuffin et al. (2006)	Only two of six sites have a larger crime decline than the comparison area	Target sites had increased confidence in the police
Weisburd and Green (1995)	Experimental group has significantly smaller increases in disorder calls compared with control group but no impact on violence or property calls	None

09012-CrimJournal-Guts.indd 151 12/21/09 7:14:20 PM

In Figure 1, we present the mean effect sizes for all eligible studies.[14] Five studies had only one outcome, so the mean effect size would be equal to the largest effect size (discussed later). For the Thomas (1998) and Knoxville Police Department (2002) studies, the outcome was probation/parole success (recidivism rate). For Tuffin et al. (2006), total crime incidents were reported. In Stone (1993), a victimization survey question was reported that asked residents whether they had been asked to buy or sell drugs, and in Stokes et al. (1996), a victimization survey question was reported that asked students whether they had been attacked or bothered on the way to or from school. For the other five studies, we combined multiple primary outcomes. In Baker and Wolfer (2003), we took the mean effect for reports of seeing vandalism and drinking. For Braga et al. (1999), we combined the total crime calls and the total crime incidents. For Mazerolle et al. (2000), we averaged the calls for service regarding drugs and disorder. In Sherman et al. (1989), the two coded outcomes were commercial calls for service and residential calls for service. For Weisburd and Green (1995), property, violence, and disorder calls for service were all combined.

FIGURE 1

Mean Effect Sizes for All Eligible Studies



Positive effect sizes indicate an effect that favors POP leading to a reduction in crime and disorder. The forest plots in Figure 1 show the standardized difference in means between the treatment and control or the comparison group (effect size) with a 95% confidence interval plotted around this point for all eligible studies. Points plotted to the right of 0 indicate a treatment effect; in this case, the study showed a reduction in crime or disorder. Points to the

14. The combined effects were computed using the Comprehensive Meta Analysis Program, which averaged effects and variances. This process is the same as assuming a correlation of 1.0 among the outcomes, which yields the largest possible standard error. Thus, the mean effect size is a conservative approach.

09012-CrimJournal-Guts.indd 152 12/21/09 7:14:20 PM

left of 0 indicate a backfire effect in which crime or disorder actually increased after a POP intervention. We used a random effects model because, as noted, POP interventions are a heterogeneous treatment that can vary considerably between studies. The common factor is the process used by the police to select an intervention strategy. Heterogeneity also is found in the types of problems addressed and in the outcomes examined. Our assumption regarding the large degree of heterogeneity in our review is confirmed when we examine the *Q* statistic, which was significant at the $p < .05$ level ($Q = 58.240$, $df = 9$).

Averaging all outcome measures in each study, we find a significant effect in favor of POP strategies. The size of the effect is relatively modest, however, with a standardized mean difference (Cohen's *d*) of .126. This result means that on average the POP intervention led to a .13 standard deviation unit decline in the metric examined. This magnitude of effect is defined by Lipsey (1990) as small but meaningful and could "easily be of practical significance" (Lipsey, 2000: 109). Cohen (1988), however, defines a small effect as having a *d* value of at least .20. Importantly, if we had used a simple "vote counting" approach to these data, then relying only on statistically significant results ($p < .05$), we would have concluded that POP was not effective, which is the case because the average of all effects in only four of the ten studies met the traditional significance criterion (and one of these results was a significant backfire).

In examining the average effect sizes for all outcome measures for specific studies, the two person-based studies have the largest overall effects. Both the probationer/parolee studies have a moderate-to-large positive impact on probation success. The Baker and Wolfer (2003) as well as the Sherman et al. (1989) studies both have a modest impact on crime, but each fails to reach statistical significance because of large standard errors. Braga et al. (1999), Mazerolle et al. (2000), and Weisburd and Green (1995) all also show a modest impact on crime and disorder. The Weisburd and Green (1995) study is highly statistically significant, and the Braga et al. (1999) and Mazerolle et al. (2000) studies have *p* values below .10 using this effect size metric.[15] The other three studies all failed to show a positive impact of POP on crime and disorder. In the Tuffin et al. (2006) and Stone (1993) studies, essentially no impact of POP on crime was observed. The Stokes et al. (1996) study had a highly significant backfire effect. Limitations of these studies, which might have led to these null and negative findings, are discussed in the next section.

Given the important distinction in methodological quality between quasi-experimental and randomized experimental studies, we also report the results separately by method. In Figure 2, we examine the mean effect sizes across all outcomes for only the four randomized experiments. The overall random effect becomes slightly larger (.147) and remains highly statistically significant ($p < .001$). In Figure 3, we look at only the quasi-experiments. The overall random effect (.158) is larger than in Figures 1 and 2 primarily because of the large effects in the two probationer/parolee studies, but it fails to reach statistical significance ($p = .108$).

15. Our effect size estimates for Weisburd and Green (1995) differ from previous systematic reviews that included these studies (Braga, 2007; Mazerolle et al., 2008) because our use of the original ANOVA data from the study allowed us to compute more exact effect sizes from the p values of the F tests.

09012-CrimJournal-Guts.indd 153 12/21/09 7:14:21 PM

FIGURE 2

Mean Effect Sizes for Randomized Experiments

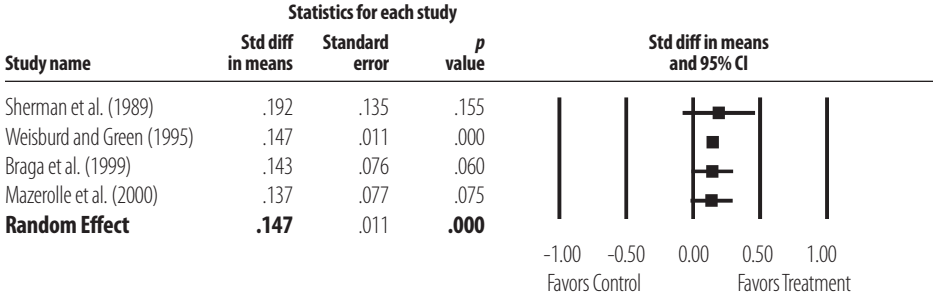
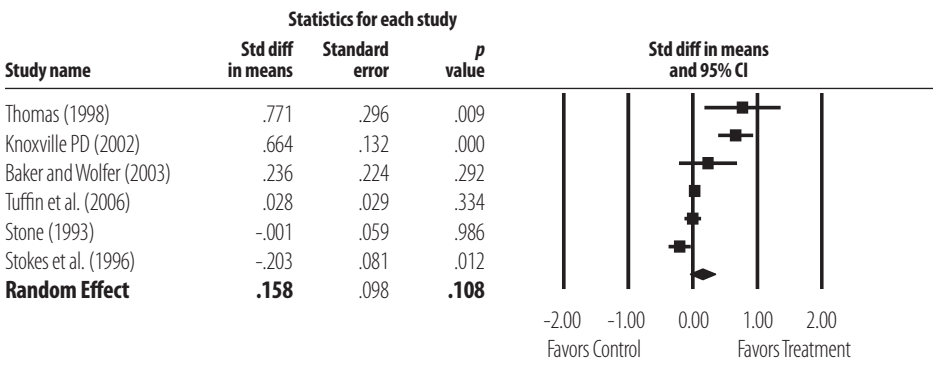


FIGURE 3

Mean Effect Sizes for Quasi-Experiments



In Figure 4, we present the meta-analytic results for the largest effect size for each study. As noted, this result can be viewed as an upper limit for the effects of POP based on existing studies. This can also been seen as where problem-oriented policing programs that examined multiple outcomes were most effective. For studies with a single outcome, the estimates were identical to those in Figure 1. As expected, the overall random effect was substantially larger (.297) than the mean combined effect size, and this effect remained statistically significant (p = .0397).[16] Among the five studies with more than one coded outcome, several of the largest effect sizes were substantially larger than the mean. For the Jersey City Drug Market Analysis Program (Weisburd and Green, 1995), the largest effect (disorder calls for service) was more than four times the size of the mean effect (.696 vs. .147) For RECAP (Sherman et al., 1989), the largest effect (residential calls for service) of .369 was nearly double the mean effect and was

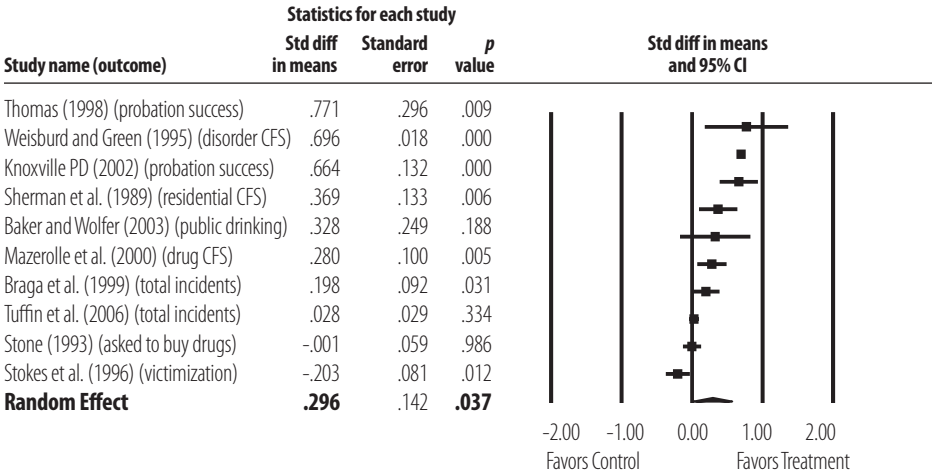
16. The p value for the random effect combining largest effects is greater than the p value for the mean effects because the standard errors for the largest effects tended to be larger than the standard errors for the smaller effects.

09012-CrimJournal-Guts.indd 154 12/21/09 7:14:21 PM

highly statistically significant. The largest effect for the Beat Health Project (Mazerolle et al. 2000) (drugs calls for service) was more than double the mean effect. In the Jersey City POP in violent places study (Braga et al., 1999), the largest effect (total incidents) was not substantially larger than the mean, but it did reach statistical significance in this analysis. The public drinking estimate for Baker and Wolfer (2003) was about .10 larger than the mean effect, but it still failed to reach statistical significance.

FIGURE 4

Largest Effect Sizes and the Outcomes for All Eligible Studies



The largest effects for just the randomized experiments are shown in Figure 5. All four randomized studies reach statistical significance when examining just the largest effect, and the overall random effect of .394 (p value = .011) indicates a moderate impact of POP on crime and disorder. In Figure 6, we present the largest effect sizes for quasi-experiments. The random effect of .167 is substantially smaller than for randomized experiments and fails to reach statistical significance at the $p < .05$ level.

Publication Bias

Publication bias presents a strong challenge to any review of evaluation studies (Rothstein, 2008). Campbell reviews (such as ours) take several steps to reduce publication bias, which is represented by the fact that six of the ten eligible studies in our review came from unpublished sources (one dissertation, two government reports, and three unpublished reports or award submissions). Wilson (2009) has argued, moreover, that often little difference in methodological quality exists between published and unpublished studies, which suggests the importance of searching the "grey literature." For our review, an upward bias also might be present in unpublished studies because two of these studies were identified through the Goldstein Award competition. The San Diego C.A.N. project (Thomas, 1998) and the Knoxville Public Safety

09012-CrimJournal-Guts.indd 155 12/21/09 7:14:21 PM

Collaborative (Knoxville Police Department, 2002) were both Goldstein Award submissions. These two studies also reported the largest overall effect sizes, both of which were highly statistically significant. Although these studies were submitted for an award, and so, they are biased toward success (because, as we will discuss, we would not expect police departments to submit unsuccessful interventions to a POP competition), both studies made strong efforts to identify reasonable and statistically valid comparison groups.

We compared mean effect sizes for unpublished versus published studies. The mean effect size for published studies is .147 ($p < .001$) compared with .153 ($p = .102$) for unpublished studies. The similarity between the mean effect sizes within the published and unpublished

FIGURE 5

Largest Effect Sizes for Randomized Experiments

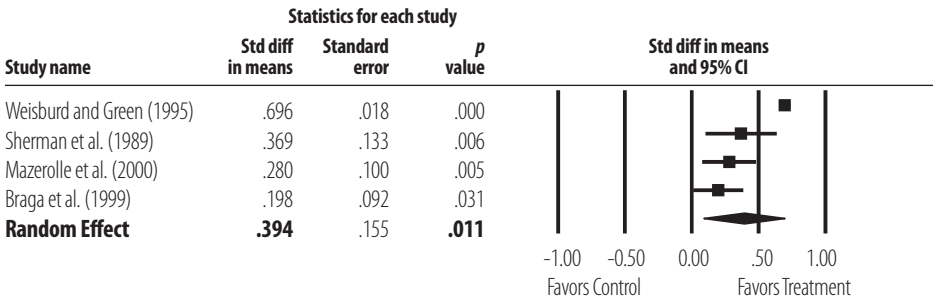
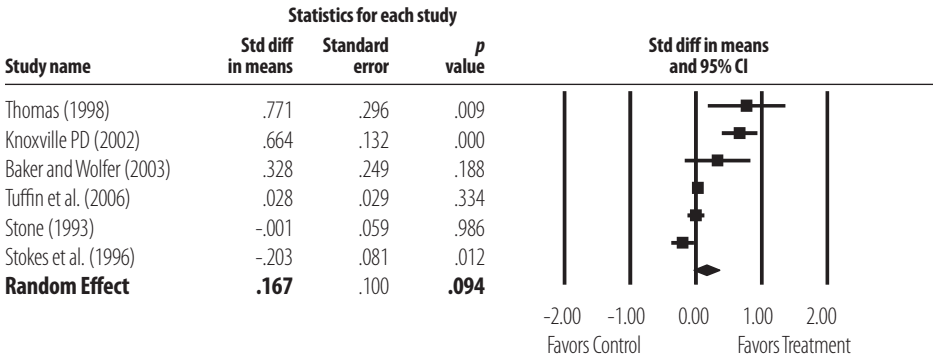


FIGURE 6

Largest Effect Sizes for Quasi-Experiments



09012-CrimJournal-Guts.indd 156 12/21/09 7:14:22 PM

literature suggests that publication bias might not have a major impact on the outcomes of this review.[17]

Study Implementation

Overall, most studies report at least a moderate level of success in implementing treatment. Nonetheless, specific implementation problems occurred in some studies, which provided a context for understanding differences in impacts across the programs. Of the experimental studies, only Mazerolle et al. (2000) reported full implementation without any significant problems. The Braga et al. (1999) study originally intended for officers to focus on 56 problem hot spots (in 28 matched pairs), but because of organizational changes in the Jersey City Police Department caused by massive retirements and extensive non-POP work, the final project included only 12 pairs of hot spots (Braga, 1997). After limited progress in the first 9 months of the experiment, Weisburd and Green (1995) extended the intervention period to achieve fuller implementation. The experiment achieved full implementation during the last 5 months of the intervention.

The Sherman et al. (1989) RECAP study presented more serious intervention problems (see Buerger, 1993). Multiple issues developed with the selection of hot spots for the intervention. Even after extensive efforts to remove duplicate calls from the computer logs, the researchers estimated that up to 15% of calls were "mirrors"—duplicates created because of multiple people calling 911 for the same incident. In addition, certain high call addresses showed remarkable instability in examining year-to-year call trends, which affected the precision of estimates. Certain addresses reviewed by police and thought to correspond with separate places were actually different entrances for the same location, which lead to problems when what was initially one location could be both in the treatment and control group. In implementing the project, the team of five officers assigned to the intervention was overwhelmed by the number of hot spot locations. In turn, the 226 addresses with a multitude of different problems were difficult to respond to adequately in a year. The absence of calls for service reductions in the second half of the experiment might be a result of officer fatigue with the intervention and an inability of officers to stay motivated during the entire year.

The two programs to reduce probationer/parolee recidivism faced no major implementation difficulties and simultaneously achieved the largest effect sizes in the study. In turn, although

17. We also generated a funnel plot to examine possible selection bias in our results using the trim-and-fill procedure developed by Duval and Tweedie (2000; see Weisburd et al., 2008). This approach suggested that an upward bias might be present in our review. However, the trim-and-fill results likely are to be misleading for our data. As Rothstein (2008) points out, this method assumes publication bias when asymmetry occurs toward the bottom of the funnel plot, as in our analyses. These studies are smaller and have a larger effect size (see Weisburd [1993] for a similar finding regarding randomized experiments in crime and justice). We think that the result is particularly understandable in POP. As we review, when POP projects endeavor to tackle too much at one time, they often face serious implementation issues. A second issue with trim-and-fit methods pointed out by Rothstein (2008) is an assumption of a relatively homogenous population of studies. As noted, the studies we review are not at all homogeneous.

09012-CrimJournal-Guts.indd 157 12/21/09 7:14:22 PM

these studies could not rely on the strong assumptions of a randomized experiment, they put significant effort toward trying to identify valid comparison conditions. The Knoxville Police Department study (2002) made a particular effort to choose a comparable historical sample of parolees, and the University of Tennessee assisted with statistical analyses to offer evidence of compatibility. The San Diego C.A.N. project (Thomas, 1998) also took strides to use a well-chosen comparison group by comparing the 80 project participants with a random sample of 80 juveniles who were on probation but not chosen for program participation.

The Baker and Wolfer (2003) study did not present significant implementation failures, but the evaluation method was potentially problematic. The comparison group of borough residents not living near the park still could have included residents who used the park and were aware of the police intervention. The survey sample sizes were also fairly small, which helps explain the large standard errors for the effect size estimates as well as the lack of a statistically significant program outcome.

The other three quasi-experiments had more substantial problems, which might explain the weak or negative study outcomes observed. Stone (1993) reported that the Atlanta Police Department did not seem entirely interested in properly implementing the POP project. Many officers did not view problem solving as "real" police work, so effort was often limited. A lack of administrative support was present from top officials in the department, and the POP training was delivered poorly and limited. In addition, Atlanta hosted the Democratic National Convention prior to the intervention, which forced officers to delay vacations because of high staffing demands. Finally, as the intervention began in the summer, officers frequently took time off, which left the POP program chronically understaffed.

Stokes et al. (1996), who produced the only backfire effect in our review, evidenced serious implementation difficulties with their school safety corridor. The largest problem seemed to be that despite an awareness campaign, two thirds of students at the target school reported that they were unaware of the existence of the corridor. In addition, even though violence was more likely in the afterschool afternoon hours, the corridor was more poorly staffed during this time period because of police shift changes and more limited police resources. Also, the victimization survey used by the researchers was not ideal for a middle-school population, and many students had difficulty answering the questions.

Tuffin et al. (2006) reported several problems with the full implementation of reassurance policing. Only two of the six target sites fully implemented the program. The other four sites had difficulties in effectively partnering with the community and using targeted problem solving. The sites that fully implemented the response showed the strongest results in favor of POP.

09012-CrimJournal-Guts.indd 158 12/21/09 7:14:22 PM

Results: Pre/Post Studies

As noted, we also collected pre/post studies that did not have a control or comparison condition. Typically, these studies examined official crime data before and after a POP intervention to determine how the POP project affected crime. These studies rarely took statistical steps to account for "history"—the idea that crime rates might be rising or falling independent of the specific POP project.

We should note that these studies vary somewhat in methodological quality and not all can be categorized as "simple pre-post."[18] These studies also covered various problems that ranged from neighborhood disorder to homicide. As with our main analyses, responses also varied greatly but frequently included a combination of increased community involvement, targeted enforcement, and situational/environmental improvements (see Weisburd et al., 2008, for more information on each study).

Thirty two of the 45 studies come from Goldstein or Tilley Award submissions. The fact that more than 70% of the studies are submissions for an award leads to a potential publication bias (Rothstein, 2008) or, rather, to a "nonpublication" bias. In our case, these nonpublished award submissions would be expected to be more positive than the published literature. We will address this issue later.

In Figure 7, we use a bar graph to display the percent change in crime and disorder reported in each study. When more than one primary outcome was present in a study, we averaged to create a single outcome. The results are overwhelmingly in favor of POP effectiveness. Of our 45 pre/post studies, 43 report a decline in crime or disorder after the POP intervention. Thus, even though 32 of our studies were award submissions, and 31 of these showed a positive impact, 12 of our 13 other studies also reported a beneficial impact of POP. Only one study (Maguire and Nettleton, 2003) reported an increase in crime after using POP. The average percent change in crime across all studies was a sizeable 44.45% decrease.

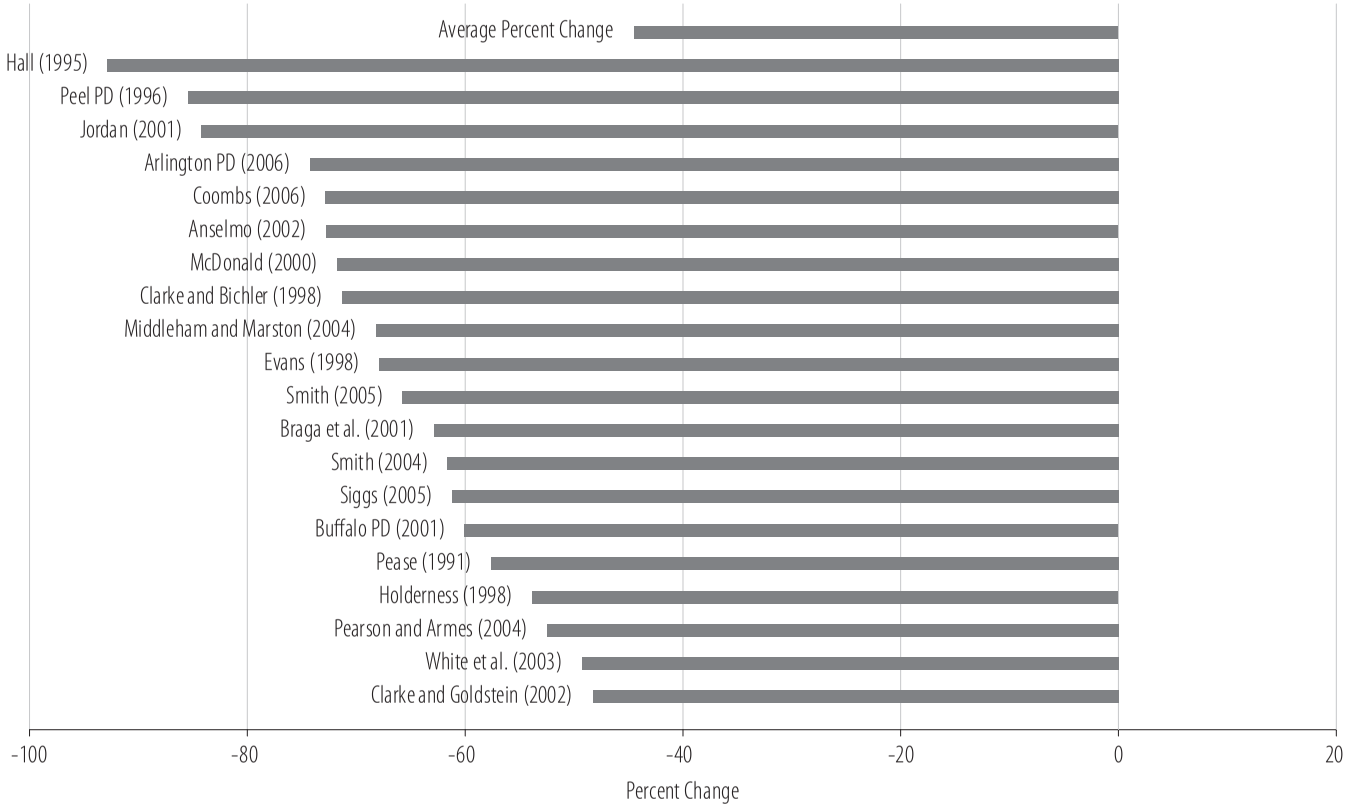
To account for variation in sample size (i.e., crime incidents or calls for service) between studies, we calculated a weighted average percent change by weighting each study by the in-

18. For example, Braga et al.'s (2001) evaluation of the Boston Gun Project used a time-series analysis and a comparison with similar-sized cities to assess the impact of Operation Ceasefire on youth homicide rates. We chose to include the Braga et al. (2001) study in this section (rather than in the main analysis) because we found this comparison with other cities insufficient for meeting our inclusion criterion that required a comparison group strongly matched to the treatment group (see footnote 6). Cities chosen as comparisons for Boston were matched only on population or geographic proximity. We do not include other "pulling levers policing" programs (see Kennedy, 2006), although such interventions sometimes are defined as POP programs. We question whether the SARA model is followed adequately in situations in which interventions are adopted from a different city (i.e., Boston). Pulling levers projects also typically involve a multiagency working group, and the police might not take a dominant enough role in this working group for the project to qualify as problem-oriented policing. In the Tita et al. (2003: xv) study, for example, the researchers lament that no group took a leadership role in the intervention, and the police viewed the project as the "RAND study." We think it important to note as well that Anthony A. Braga and David Weisburd have begun a Campbell systematic review of pulling levers policing. Accordingly, the excluded studies will be assessed in a future review.

09012-CrimJournal-Guts.indd 159 12/21/09 7:14:23 PM

FIGURE 7

Percent Change for Pre/Post Studies
(Top Bar Is Average Percent Change)



09012-CrimJournal-Guts.indd 160 12/21/09 7:14:23 PM

09012-CrimJournal-Guts.indd 161 12/21/09 7:14:24 PM

verse of its variance and assuming crime follows a Poisson distribution.[19] With this sampling variance, we constructed a confidence interval around the percentage change for each study. A plot of proportion change with confidence intervals is presented in Figure 8. After weighting each study by the inverse of its variance, we recalculated the average percent change. Even with weighting, the average decrease in crime was still 32.49%. Accordingly, although these before and after studies did not employ the methodological rigor of a randomized experiment, they did show consistently a substantial impact of POP on crime and disorder in both the award submissions and the published journal articles.

To address our publication bias concern, we compared the percent change for all studies as well as for the published and unpublished studies separately. As noted, the overall percent change for all studies was a 44.45% decrease in crime and disorder. When we examine only award submissions, a larger percent decrease of 47.79% occurs. For the non-award submissions, the percent decrease is smaller but still substantial (35.55%). For the six published studies, the average percent decrease is similar to the award submissions (47.42%). Thus, although publication type varies, the results are not substantively different. Across the publication medium POP is associated with a sizable crime decline in before–after studies that do not employ comparison groups. It is important to note, however, that prior reviews of crime and justice studies suggest that weaker quasi-experimental designs are more likely to overestimate program effects (Lipsey and Wilson, 1993; Weisburd et al., 2001).

Conclusions and Policy Implications

We began our study with a main research question regarding the effectiveness of POP in reducing crime and disorder. Overall, our main review of experimental and quasi-experimental studies with comparison groups reinforces prior findings based on narrative reviews (NRC, 2004; Sherman and Eck, 2002; Weisburd and Eck, 2004) and more general assumptions regarding the crime and disorder prevention benefits of POP approaches (Bullock and Tilley, 2003; Eck and Spelman, 1987; Goldstein, 1990; Scott, 2000). Whether we used a more conservative mean effect size approach or examined the largest effects on crime and disorder reported, we found that POP approaches have a statistically significant effect on the outcomes examined. Importantly, the results are similar whether we look at experimental or nonexperimental studies.

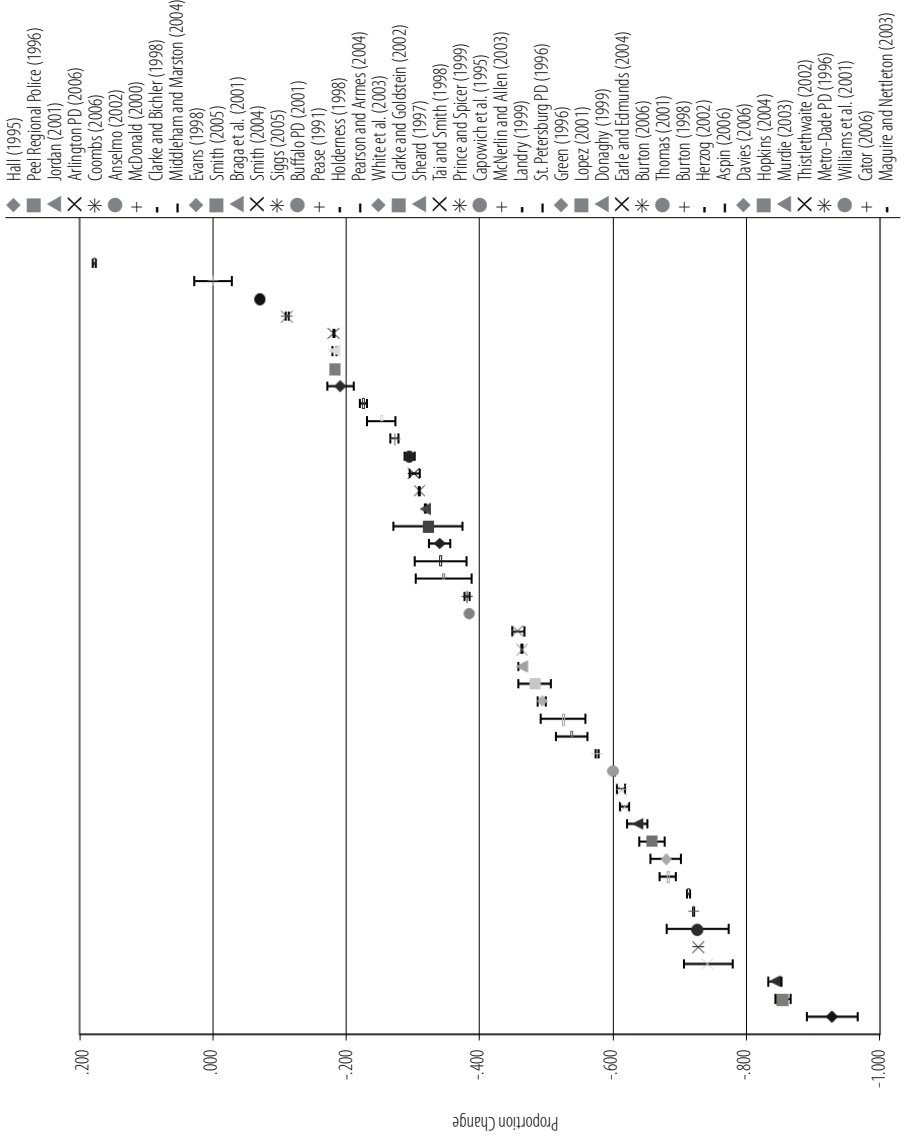
One surprise in our analyses given prior discussion of POP is the relatively modest effects observed in the studies that met our main inclusion criteria. The average mean effect size of between .10 and .20 for POP interventions, although meaningful and statistically significant, might be smaller than some scholars have assumed. One explanation for this trend is suggested by our identification of implementation problems in some studies we reviewed. We find that weaker program effects are often the result of a failure to implement fully POP interventions. This finding is consistent with other reviews in criminology that have identified treatment

19. An approximation to the variance based on the Poisson distribution is as follows: Variance = (pre count + post count) / [pre count2 × (pre count / post count)].

09012-CrimJournal-Guts.indd 162 12/21/09 7:14:24 PM

FIGURE 8

Proportion Change with Confidence Intervals for Pre/Post Studies



09012-CrimJournal-Guts.indd 163 12/21/09 7:14:25 PM

fidelity as a key issue in understanding weak program impacts (Farrington, Ohlin, and Wilson, 1986; Weisburd, 1993).

Moreover, our examination of the largest effects in the studies often led to much more robust findings. In turn, it is not always disingenuous to focus on specific outcomes because they are sometimes the main concern of the intervention (see Weisburd and Green, 1995). In addition, when we examine pre/post studies, we do in fact find much stronger impacts for POP approaches. Although prior reviews suggest that these methods are likely to overstate program effectiveness strongly (Lipsey and Wilson, 1993; Weisburd et al., 2001), we think that the combination of findings in our study, and their overall consistency, adds strength to our general conclusions.

What is most surprising in our review is that so small a group of studies met our main inclusion threshold. As noted, POP is one of the most important and widely implemented police innovations of the last two decades. The small group of studies in our review allows us to come to a solid conclusion regarding the promise of POP, but it does not allow statistical conclusions regarding the types of approaches that work best for specific types of problems. We think it a major public policy failure that the government and the police have not invested greater effort and resources in identifying the POP approaches and tactics that work best to combat specific types of crime. The portfolio of studies is at best serendipitous and does not represent any concerted public effort to either assess the effectiveness of POP as an approach or understand the mechanisms that would make it more successful.

Moreover, a much larger number of studies is needed to draw strong generalizations regarding the possible effectiveness of POP across different types of jurisdictions and different types of police agencies. Is POP more difficult to implement in smaller police agencies where resources are limited? Is it more useful as a strategy in jurisdictions with more serious crime problems? The portfolio of available studies does not allow us to draw conclusions about such contextual factors and suggests that U.S. policing has adopted POP widely without an evidence base for deciding where and when it should be used.

The central conclusion of our review is that POP as an approach has significant promise to ameliorate crime and disorder problems broadly defined. The most successful studies in this review covered problems that ranged from parolee recidivism, to violence in hot spots, to drug markets. But this diversity of programs and approaches also should bring caution to any conclusions drawn from our study. These studies often involve overlapping interventions such as hot spots policing or community policing. Indeed, many policing interventions are so multifaceted that it can be difficult to isolate the impact of any one aspect of the treatment. At the same time, with POP, it is important to remember that we are not evaluating a particular police strategy per se. Instead, we are evaluating a process police use to develop strategies. Despite few eligible studies, we find a statistically significant but modest positive impact for the effects of that process on crime and disorder.

09012-CrimJournal-Guts.indd 164 12/21/09 7:14:25 PM

References[20]

**Anselmo, Diego. 2002. *Hispanic Robbery Initiative: Reducing Robbery Victimization and Increasing Trust of Police and Financial Institutions in a Hispanic Community.* Charlotte-Mecklenburg, NC: Charlotte-Mecklenburg Police Department. Herman Goldstein Award Finalist.

**Arlington Police Department. 2006. *Stopping Open-Air Drug Sales on West Cedar Street, in Arlington, Texas.* Herman Goldstein Award Submission.

**Aspin, Mark. 2006. *Trafford Park Security Initiative—Reducing Crime in Trafford Park.* Safer Trafford Partnership. Manchester, UK. Tilley Award Submission.

*Baker, Thomas E. and Loreen Wolfer. 2003. The crime triangle: Alcohol, drug use, and vandalism. *Police Practice and Research,* 4: 47–61.

Bazemore, Gordon and Allen W. Cole. 1994. Police in the "laboratory" of the neighborhood: Evaluating problem-oriented policing strategies in a medium sized city. *American Journal of Police,* 8: 119–147.

Borenstein, Michael, Larry V. Hedges, Julian P. T. Higgins, and Hannah R. Rothstein. 2009. *Introduction to Meta-Analysis.* West Sussex, UK: Wiley.

Boruch, Robert, Brooke Snyder, and Dorothy DeMoya. 2000. The importance of randomized field trials. *Crime & Delinquency,* 46: 156–180.

Braga, Anthony A. 1997. *Solving Violent Crime Problems: An Evaluation of the Jersey City Police Department's Pilot Program to Control Violent Places.* Unpublished doctoral dissertation, Rutgers, The State University of New Jersey.

Braga, Anthony A. 2001. The effects of hot spots policing on crime. *Annals of the American Academy of Political and Social Sciences,* 578: 104–125.

Braga, Anthony A. 2002. *Problem-Oriented Policing and Crime Prevention.* Monsey, NY: Criminal Justice Press.

Braga, Anthony A. 2007. *Effects of Hot Spots Policing on Crime. A Campbell Collaboration Systematic Review.* Retrieved March 26, 2009 from db.c2admin.org/doc-pdf/Braga_HotSpotsPolicing_review.pdf.

**Braga, Anthony A., David M. Kennedy, Elin J. Waring, and Anne M. Piehl. 2001. Problem-oriented policing, deterrence, and youth violence: An evaluation of Boston's Operation Ceasefire. *Journal of Research in Crime and Delinquency,* 38: 195–225.

*Braga, Anthony A., David Weisburd, Elin J. Waring, Lorraine Green Mazerolle, William Spelman, and Francis Gajewski. 1999. Problem-oriented policing in violent crime places: A randomized controlled experiment. *Criminology,* 37: 541–580.

Buerger, Michael. 1993. *Convincing the Recalcitrant: Reexamining the Minneapolis RECAP Experiment.* Unpublished doctoral dissertation, Rutgers, The State University of New Jersey.

**Buffalo Police Department. 2001. W*orkable Solutions to the Problem of Street Prostitution in Buffalo, NY.* Herman Goldstein Award Finalist.

20. Eligible studies included in the main analysis are denoted with a single asterisk (*). Eligible pre/post studies included in our analysis are denoted with two asterisks (**).

09012-CrimJournal-Guts.indd 165 12/21/09 7:14:25 PM

Bullock, Karen and Nick Tilley. 2003. *Crime Reduction and Problem-Oriented Policing.* Portland, OR: Willan.

Bullock, Karen, Rosie Erol, and Nick Tilley. 2006. *Problem-Oriented Policing and Partnerships: Implementing an Evidence-Based Approach to Crime Reduction.* Cullompton, UK: Willan.

Bureau of Justice Statistics. 2006. *Law Enforcement Management and Administrative Statistics (LEMAS): 2003 Sample Survey of Law Enforcement Agencies.* Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

**Burton, Ken. 1998. *Cowboys: A Problem Solving Initiative. Arlington (TX) Police Department.* Herman Goldstein Award Submission.

**Burton, Steve. 2006. *Safer Travel at Night: Transport for London.* London Transport Policing and Enforcement Directorate, UK. Herman Goldstein Award Winner.

Campbell, Donald and Jean Russo. 1999. *Social Experimentation.* Thousand Oaks, CA: Sage.

Capowich, George E. and Janet A. Roehl. 1994. Problem-oriented policing: Actions and effectiveness in San Diego. In (Dennis P. Rosenbaum, ed.), *The Challenge of Community Policing: Testing the Promises.* Thousand Oaks, CA: Sage.

**Capowich, George, E., Janice A. Roehl, and Christine M. Andrews. 1995. *Evaluating Problem-Oriented Policing: Process and Outcomes in Tulsa and San Diego.* Washington, DC: National Institute of Justice, Office of Justice Programs.

**Cator, Marcus. 2006. *Operation Mullion: Reducing Anti-Social Behaviour and Crime in and Around Mayfield School.* Hampshire Constabulary, UK. Tilley Award Finalist.

Chalmers, Iain. 2003. Trying to do more good than harm in policy and practice: The role of rigorous, transparent, up-to-date evaluations. *Annals of the American Academy of Political and Social Science,* 589: 22–40.

Clarke, Ronald V. 1992a. Situational crime prevention: Theory and practice. *British Journal of Criminology,* 20: 136–147.

Clarke, Ronald V. 1992b. *Situational Crime Prevention: Successful Case Studies.* Albany, NY: Harrow and Heston.

Clarke, Ronald V. 1997. *Situational Crime Prevention: Successful Case Studies,* 2nd Edition. Albany, NY: Harrow and Heston.

Clarke, Ronald V. and Derek B. Cornish. 1972. *The Controlled Trial in Institutional Research: Paradigm or Pitfall for Penal Evaluators?* London, UK: Her Majesty's Stationery Office.

**Clarke, Ronald V. and Gisela Bichler-Robertson. 1998. Place managers, slumlords and crime in low rent apartment buildings. *Security Journal,* 11: 11–19.

**Clarke, Ronald V. and Herman Goldstein. 2002. *Reducing Theft at Construction Sites: Lessons from a Problem-Oriented Project.* Washington, DC: U.S. Department of Justice.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences,* 2nd Edition. Hillsdale, NJ: Lawrence Erlbaum.

Cook, Thomas and Donald Campbell. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Boston, MA: Houghton Mifflin Harcourt.

09012-CrimJournal-Guts.indd 166 12/21/09 7:14:25 PM

**Coombs, Adrian. 2006. *Policing the Glastonbury Festival.* Avon and Somerset Constabulary, UK. Tilley Award Submission.

Cordner, Gary W. 1986. Fear of crime and the police: An evaluation of a fear-reduction strategy. *Journal of Police Science and Administration,* 14: 223–233.

Cordner, Gary and Elizabeth P. Biebel. 2005. Problem-oriented policing in practice. *Criminology & Public Policy,* 4: 155–180.

**Davies, Amanda. 2006. *Operation Clean Up.* Staffordshire Police, UK. Tilley Award Submission.

**Donaghy, Jim. 1999. *Northfields Project: Project Brings Peace Back to City Estate.* Leicestershire Constabulary, UK. Tilley Award Winner.

Duval, Sue J. and Richard L. Tweedie. 2000. A non-parametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association,* 95: 89–98.

**Earle, Julie and Alan Edmunds. 2004. *Operation Cobra: Tackling Vehicle Crime in the City of Portsmouth.* Hampshire Constabulary, UK. Tilley Award Runner-Up.

Eck, John E. 2002a. Preventing crime at places. In (Lawrence Sherman, David P. Farrington, Brandon C. Welsh, and Doris L. MacKenzie, eds.), *Evidence-Based Crime Prevention.* New York: Routledge.

Eck, John E. 2002b. Learning from experience in problem-oriented policing and crime prevention: The positive functions of weak evaluations and the negative functions of strong ones. In (Nick Tilley, ed.), *Evaluation for Crime Prevention.* Monsey, NY: Criminal Justice Press.

Eck, John E. and William Spelman. 1987. *Problem Solving: Problem-Oriented Policing in Newport News.* Washington, DC: Police Executive Research Forum.

**Evans, Marvin. 1998. *PRIDE: Police Response to Incidents of Domestic Emergencies. Reducing Domestic Violence Related Homicides.* Newport News Police Department, UK. Herman Goldstein Award Finalist.

Farrington, David P. 1983. Randomized experiments on crime and justice. In (Michael H. Tonry, ed.), *Crime and Justice: A Review of Research.* Chicago, IL: University of Chicago Press.

Farrington, David P. 2003. Methodological quality standards for evaluation research. *Annals of the American Academy of Political and Social Science,* 587: 49–68.

Farrington, David P. 2006. Methodological quality and the evaluation of anti-crime programs. *Journal of Experimental Criminology,* 2: 329–327.

Farrington, David P. and Anthony Petrosino. 2001. The Campbell Collaboration Crime and Justice Group. *Annals of the American Academy of Political and Social Science,* 578: 35–49.

Farrington, David P., Denise C. Gottfredson, Lawrence Sherman, and Brandon C. Welsh. 2002. The Maryland Scientific Methods Score. In (Lawrence Sherman, David P. Farrington, Brandon C. Welsh, and Doris L. MacKenzie, eds.), *Evidence-Based Crime Prevention.* New York: Routledge.

09012-CrimJournal-Guts.indd 167 12/21/09 7:14:25 PM

Farrington, David P., Martin Gill, Sam J. Waples, and Javier Argomaniz. 2007. The effects of closed-circuit television on crime: Meta-analysis of an English national quasi-experimental multi-site evaluation. *Journal of Experimental Criminology,* 3: 21–38.

Farrington, David P., Lloyd E. Ohlin, and James Q. Wilson. 1986. *Understanding and Controlling Crime: Toward a New Research Strategy.* New York: Springer-Verlag.

**Forrester, David, Mike Chatterston, and Ken Pease. 1988. *The Kirkholt Burglary Reduction Project, Rochdale.* London, UK: Home Office Crime Prevention Unit.

Goldstein, Herman. 1979. Improving policing: A problem-oriented approach. *Crime & Delinquency,* 24: 236–258.

Goldstein, Herman. 1990. *Problem-Oriented Policing.* New York: McGraw-Hill.

**Green, Lorraine. 1996. *Policing Places with Drug Problems.* Thousand Oaks, CA: Sage.

**Hall, Daniel M. 1995. *Drug Activity at Birch Street.* Norfolk Police Department, UK. Herman Goldstein Award Submission.

Heckman, James and Jeffrey A. Smith. 1995. Assessing the case for social experimentation. *Journal of Economic Perspectives,* 9: 85–110.

**Herzog, Sergio. 2002. Does proactive policing make a difference in crime? An implementation of problem-solving policing in Israel. *International Journal of Comparative and Applied Criminal Justice,* 26: 29–52.

Hill, Bradford. 1962. *Principles of Medical Statistics.* New York: Oxford University Press.

**Holderness, Terry. 1998. *"Ten-4": The Transient Enrichment Network: A Community Collaboration to Reduce Homelessness.* Fontana Police Department, CA. Herman Goldstein Award Submission.

Hope, Timothy. 1994. Problem-oriented policing and drug market locations: Three case studies. In (Ronald V. Clarke, ed.), *Crime Prevention Studies,* vol. 2. Monsey, NY: Criminal Justice Press.

**Hopkins, Matt. 2004. Targeting hot spots of alcohol-related town centre violence: A Nottinghamshire case study. *Security Journal,* 17: 53–66.

**Jordan, James T. 2001. *School impact project 2000: Dorchester High School.* Boston Police Department, MA. Herman Goldstein Award Submission.

Kansas City Police Department. 1977. *Response Time Analysis: Executive Summary.* Kansas City, MO: Board of Commissioners.

Kelling, George L., Tony Pate, Duane Dieckman, and Charles E. Brown. 1974. The *Kansas City Preventive Patrol Experiment: Technical Report.* Washington, DC: Police Foundation.

Kennedy, David M. 2006. Old wine in new bottles. Policing and the lessons of pulling levers. In (David Weisburd and Anthony A. Braga, eds.), *Police Innovation: Contrasting Perspectives.* New York: Cambridge University Press.

Kennedy, David M., Anthony A. Braga, Anne M. Piehl, and Elin J. Waring. 2001. *Reducing Gun Violence: The Boston Gun Project's Operation Ceasefire.* Washington, DC: National Institute of Justice, U.S. Department of Justice.

*Knoxville Police Department. 2002. *The Knoxville Public Safety Collaborative.* Knoxville, TN. Herman Goldstein Award Submission.

09012-CrimJournal-Guts.indd 168 12/21/09 7:14:25 PM

**Landry, Steve. 1999. *Operation First Strike.* Phoenix Police Department, AZ. Herman Goldstein Award Submission.

Lipsey, Mark. 1990. *Design Sensitivity: Statistical Power for Experimental Research.* Thousand Oaks, CA: Sage.

Lipsey, Mark. 2000. Statistical conclusion validity for intervention research: A significant (p < .05) problem. In (Leonard Bickman, ed.), *Validity and Social Experimentation: Donald Campbell's Legacy.* Thousand Oaks, CA: Sage.

Lipsey, Mark and David B. Wilson. 1993. The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist,* 48: 1181–1209.

Lipsey, Mark. and David B. Wilson. 2001. *Practical Meta-Analysis.* Thousand Oaks, CA: Sage.

**Lopez, Yolanda. 2001. *Reducing Juvenile Graffiti Vandalism: A Study of Community Policing Problem-Solving in San Benito, Texas.* Unpublished Master's thesis, University of Texas.

Lösel, Friedrich and Peter Köferl. 1989. Evaluation research on correctional treatment in West Germany: A meta-analysis. In (Herman Wegener, Freidrich Lösel, and Jochen Haisch, eds.), *Criminal Behavior and the Justice System.* New York: Springer.

**Maguire, Mike and Hilary Nettleton. 2003. *Reducing Alcohol-Related Violence and Disorder: An Evaluation of the "TASC" Project.* London, UK: Home Office Research.

*Mazerolle, Lorraine Green, James F. Price, and Jan Roehl. 2000. Civil remedies and drug control: A randomized field trial in Oakland, California. *Evaluation Review,* 24: 212–241.

Mazerolle, Lorraine Green and Janet Ransley. 2005. *Third Party Policing.* New York: Cambridge University Press.

**Mazerolle, Lorraine Green, Justin Ready, William Terrill, and Elin Waring. 2000. Problem oriented policing in public housing: The Jersey City evaluation. *Justice Quarterly,* 17: 129–158.

Mazerolle Green, Lorraine, David W. Soole, and Sacha Rombouts. 2005. *Drug Law Enforcement: The Evidence.* Fitzroy, Australia: Turning Point Alcohol and Drug Centre.

Mazerolle Green, Lorraine, David W. Soole, and Sacha Rombouts. 2008. *Street Level Drug Law Enforcement: A Meta-Analytic Review.* Retrieved March 26, 2009 from db.c2admin.org/doc-pdf/Mazerolle_DrugLaw_Enforcement_review.pdf.

**McDonald, Barbara B. 2000. *Taking Back our Blocks: The Avers Story.* Chicago Police Department, IL. Herman Goldstein Award Submission.

**McNerlin, Julie and Samantha Allen. 2003. *Tackling City Centre Assaults in Foyle District Command Unit.* Police Service of Northern Ireland, Belfast. Tilley Award Winner.

**Metro Dade Police Department. 1996. *Tourist-Oriented Police Program.* Miami, FL. Herman Goldstein Award Submission.

**Middleham, Neil and Caroline Marston. 2004. *Mole Hills from Mountains.* Lancashire Constabulary, UK. Herman Goldstein Award Finalist.

**Murdie, Robert. 2003. *Campaign "Get Home Safe."* Police Service of Northern Ireland, Belfast. Tilley Award Winner.

09012-CrimJournal-Guts.indd 169 12/21/09 7:14:26 PM

National Research Council. 2004. Effectiveness of police activity in reducing crime, disorder and fear. In (Wesley Skogan and Kathleen Frydl, eds.), *Fairness and Effectiveness in Policing: The Evidence.* Washington, DC: The National Academies Press.

Pawson, Ray and Nick Tilley. 1997. *Realistic Evaluation.* Beverly Hills, CA: Sage.

**Pearson, Gareth and Steve Armes. 2004. *The Hopwood Triangle: Revitalizing a Depressed Neighbourhood in Lancashire.* Lancashire Constabulary, UK. Herman Goldstein Award Finalist.

**Pease, Ken. 1991. The Kirkholt Project: Preventing burglary on a British public housing estate. *Security Journal,* 2: 73–77.

**Peel Regional Police. 1996. *The Turner-Fenton Project: Reducing School Disorder with CPTED.* Brampton, Canada. Herman Goldstein Award Winner.

Perry, Amanda and Matthew Johnson. 2008. Applying the Consolidated Standards of Reporting Trials (CONSORT) to studies of mental health provision for juvenile offenders: A research note. *Journal of Experimental Criminology,* 4: 165–185.

Perry, Amanda, David Weisburd, and Catherine Hewitt. In press. Are criminologists reporting experiments in ways that allow us to assess them? *Journal of Experimental Criminology.*

Poyner, Barry. 1981. Crime prevention and the environment—Street attacks in city centres. *Police Research Bulletin,* 37: 10–18.

**Prince, Jean and Valerie Spicer. 1999. *Intersecting Solutions: How Consistent Police Enforcement, Partnerships with the Community and Environmental Change Restored Order and Civility to an Urban Intersection.* Vancouver Police Department, Canada. Herman Goldstein Award Finalist.

Rothstein, Hannah R. 2008. Publication bias as a threat to the validity of meta-analytic results. *Journal of Experimental Criminology,* 4: 61–81.

Rubenser, Lorie. 2005. Unofficial use of problem-oriented policing: An analysis at the department and individual officer level. *Southwest Journal of Criminal Justice,* 2: 23–40.

Scott, Michael S. 2000. *Problem-Oriented Policing: Reflections on the First 20 Years.* Washington, DC: U.S. Department of Justice.

Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston, MA: Houghton-Mifflin.

**Sheard, Mike. 1997. *The Elite Arcade: Taming a Crime Generator. Reducing Crime and Disorder in an Arcade.* Delta Police Department, CO. Herman Goldstein Award Finalist.

Sherman, Lawrence and John E. Eck. 2002. Policing for crime prevention. In (Lawrence Sherman, David P. Farrington, Brandon C. Welsh, and Doris L. MacKenzie, eds.), *Evidence-Based Crime Prevention.* New York: Routledge.

*Sherman, Lawrence, Michael Buerger, and Patrick Gartin. 1989. *Repeat Call Address Policing: The Minneapolis RECAP Experiment.* Washington, DC: Crime Control Institute.

09012-CrimJournal-Guts.indd 170 12/21/09 7:14:26 PM

Sherman, Lawrence, Denise C. Gottfredson, Doris L. MacKenzie, John E. Eck, Peter Reuter, and Shawn D. Bushway. 1997. *Preventing Crime: What Works, What Doesn't, What's Promising.* Washington, DC: U.S. Department of Justice.

**Siggs, Richard. 2005. *Operation Dodger: Policing the street community in Brighton and Hove.* Sussex Police, UK. Tilley Award Winner.

**Smith, Andy. 2004. *Safe and Secure—Twenty Four Seven.* Staffordshire Police, UK. Tilley Award Finalist.

**Smith, Andy. 2005. *Nowhere to Run to Nowhere to Hide: Neighborhood Burglary Reduction.* Staffordshire Police, UK. Herman Goldstein Award Finalist.

Solé Brito, Corina and Tracy Allan. 1999. *Problem-Oriented Policing: Crime-Specific Problems, Critical Issues and Making POP Work,* vol. II. Washington, DC: Police Executive Research Forum.

Spelman, William and John E. Eck. 1987. *Problem-Oriented Policing.* Washington, DC: U.S. Department of Justice.

*Stokes, Robert, Neil Donahue, Dawn Caron, and Jack R. Greene. 1996. *Safe Travel to and from School: A Problem-Oriented Policing Approach.* Washington, DC: U.S. Department of Justice.

*Stone, Sandra S. 1993. *Problem-Oriented Policing Approach to Drug Enforcement: Atlanta as a Case Study.* Unpublished doctoral dissertation, Emory University.

**St. Petersburg Police Department. 1996. *Protect "Respect."* St. Petersburg, FL. Herman Goldstein Award Submission.

**Tai, Rudy and Robert C. Smith. 1998. *San Ysidro Boulevard: Reducing Crime in a Bordertown Block.* San Diego Police Department, CA. Herman Goldstein Award Finalist.

**Thistlethwaite, Edward and Paolo Pertica. 2002. *The Tower Project.* Lancashire Constabulary, UK. Tilley Award Submission.

**Thomas, Clive. 2001. *Bristol Anti-Robbery Strategy.* Avon and Somerset Constabulary, UK. Tilley Award Submission.

*Thomas III, George R. 1998. *Coordinated Agency Network (C.A.N.).* San Diego Police Department, CA. Herman Goldstein Award Submission.

Tita, George, K. Jack Riley, Greg Ridgeway, Clifford Grammich, Allan F. Abrahamse, and Peter W. Greenwood. 2003. *Reducing Gun Violence: Results from an Intervention in East Los Angeles.* Santa Monica, CA: RAND.

*Tuffin, Rachel, Julia Morris and Alexis Poole. 2006. *An Evaluation of the Impact of the National Reassurance Policing Programme.* London, UK: Home Office Research.

Weisburd, David. 1993. Design sensitivity in criminal justice experiments. In (Michael H. Tonry, ed.), *Crime and Justice: A Review of Research.* Chicago, IL: University of Chicago Press.

Weisburd, David. 1997. *Reorienting Crime Prevention Research and Policy: From the Causes of Criminality to the Context of Crime.* Washington, DC: U.S. Government Printing Office.

09012-CrimJournal-Guts.indd 171 12/21/09 7:14:26 PM

Weisburd, David. 2003. Ethical practice and evaluation of interventions in crime and justice: The moral imperative for randomized trials. *Evaluation Review,* 27: 336–354.

Weisburd, David and John E. Eck. 2004. What can the police do to reduce crime, disorder and fear? *Annals of the American Academy of Social and Political Sciences,* 593: 42–65.

*Weisburd, David and Lorraine Green. 1995. Policing drug hot spots: The Jersey City drug market analysis experiment. *Justice Quarterly,* 12: 711–735.

Weisburd, David, Cynthia Lum, and Anthony Petrosino. 2001. Does study design affect research outcomes in criminal justice? *The Annals of the American Academy of Political and Social Sciences,* 578: 50–70.

Weisburd, David, Cody W. Telep, Joshua C. Hinkle, and John E. Eck. 2008. *The Effects of Problem-Oriented Policing on Crime and Disorder.* Retrieved March 26, 2009 from db.c2admin.org/doc-pdf/Weisburd_POP_review.pdf.

**White, Michael D., James J. Fyfe, Suzanne P. Campbell, and John S. Goldkamp. 2003. The police role in preventing homicide: Considering the impact of problem-oriented policing on the prevalence of murder. *Journal of Research in Crime and Delinquency,* 40: 194–225.

**Williams, Paul, Paul White, Michael Teece, and Robert Kitto. 2001. *Problem-Oriented Policing: Operation Mantle—A Case Study.* Canberra, Australia: Australian Institute of Criminology.

Wilson, David B. 2009. The importance of a comprehensive document search as part of systematic reviews. *Journal of Experimental Criminology.*

David Weisburd is the Walter E. Meyer Professor of Law and Criminal Justice at the Hebrew University Law School, and the Distinguished Professor of Administration of Justice at George Mason University. He is the 2010 recipient of the Stockholm Prize in Criminology.

Cody W. Telep is a doctoral student in the Justice, Law, and Crime Policy Program in the Administration of Justice Department at George Mason University. His research interests include innovations in policing, police education, and advancing evidence-based policy.

Joshua C. Hinkle is an assistant professor in the Department of Criminal Justice at Georgia State University. His research interests include policing, criminological theory, and evidence-based crime policy.

John E. Eck is a professor in the School of Criminal Justice at the University of Cincinnati. He studies crime prevention and problem-oriented policing.

09012-CrimJournal-Guts.indd 172 12/21/09 7:14:26 PM

POLICY ESSAY

PROBLEM-ORIENTED POLICING

Setting a higher standard for the evaluation of problem-oriented policing initiatives

Anthony A. Braga
Harvard University
University of California, Berkeley

Approximately 30 years have passed since Herman Goldstein (1979) first articulated the basic principles of problem-oriented policing. During those years, problem-oriented policing emerged as a widely practiced approach to crime prevention in police departments throughout the United States as well as throughout the world. Problem-oriented policing enjoys broad support from federal agencies, professional policing groups, and a small cadre of scholars interested in effective crime prevention practices. Given this support, it was surprising to learn that Weisburd, Telep, Hinkle and Eck (2010, this issue) identified only ten evaluations that met the minimum methodological standards of their systematic review—a comparison group that did not receive the problem-oriented policing treatment. Their systematic review also identified 45 before–after intervention studies without a strong control or comparison group.

To some observers, a ratio of one strong evaluation to every four or five weak evaluations might not sound overly concerning. Weak evaluations, unfortunately, provide less valid answers to policy questions when compared with well-designed quasi-experiments and randomized controlled trials (Campbell and Stanley, 1966; Cook and Campbell, 1979). Several crime and justice scholars suggest that a "moral imperative" exists in pursuing the most rigorous evaluation designs to discover whether a program is effective (e.g., Boruch, 1975; Weisburd, 2003). Isolating the effects of treatments or programs from other confounding aspects of selection or design is viewed as one of the evaluator's most important obligations to society. When the evaluation evidence base is informed largely by weak designs, practitioners risk implementing certain treatments or programs as effective crime prevention practices when they are not, which can lead to significant economic and social costs.

Much of the academic work on problem-oriented policing seeks to improve practice by refining key steps in the process such as encouraging the in-depth analysis of problems and searching

Direct correspondence to Anthony A. Braga, John F. Kennedy School of Government, Harvard University, 79 JFK Street, Cambridge, MA 02138 (e-mail: anthony_braga@harvard.edu).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 173 12/21/09 7:14:26 PM

for innovative responses that go beyond traditional enforcement. Unfortunately, although these academics press practitioners to conduct more rigorous assessments of implemented prevention strategies, they also tend to be dismissive of the use of more rigorous research designs, such as randomized controlled trials, in evaluating the crime prevention effects of problem-oriented policing initiatives. This apparent bias against more rigorous designs is certainly a contributing factor to the weak scientific evidence base for the problem-oriented policing approach. Beyond increasing the openness of problem-oriented policing scholars to include stronger research designs, investments need to be made to develop strong working relationships with police practitioners so that opportunities can be created to conduct more rigorous evaluations.

Can Problem-Oriented Policing Be Properly Evaluated Using More Rigorous Designs?

The problem-oriented policing movement has benefited greatly from the hard work of a few academics, such as Herman Goldstein, Ronald V. Clarke, Nick Tilley, John E. Eck, Johannes Knutsson, and others who are concerned about the quality and effectiveness of policing. These academics have made a concerted effort to improve the quality of problem analysis, identify alternative approaches to deal with the underlying conditions and situations that cause recurring problems to exist, as well as disseminate examples of successful problem-oriented policing projects through conference participation, the production of practitioner-friendly guides, and strategic partnerships with governmental agencies such as the U.S. Office of Community-Oriented Policing Services. Indeed, these scholars deserve much credit for stimulating interest in the problem-oriented policing approach in police departments around the world.

These academics are vocal in lamenting the poor quality of practitioner-led assessments of problem-oriented policing projects and conclude that police departments need considerable support from academic partners to carry out evaluations (e.g., Clarke, 1998). Although they advocate for a range of evaluation methodologies, these academics are often skeptical of using more rigorous evaluation designs to evaluate problem-oriented policing projects (Eck, 2002; Knutsson, 2009; Pawson and Tilley, 1997; Tilley, 2009). Randomized controlled trials, in particular, are described in a discouraging manner. They suggest that problem-oriented policing interventions are highly customized for specific local problems and, as such, often are not amenable to rigorous designs that rule out threats to internal validity. Experimental evaluations of problem-oriented policing also are described as "academic research projects" with low external validity that does not reflect the reality of problem-oriented policing as it is practiced in the field (Knutsson, 2009: 20).

These problem-oriented policing scholars suggest that single group before–after designs are the modal assessment methodology to evaluate problem-oriented policing interventions, and they argue that it is much more feasible and fruitful to decrease threats to internal validity through different means such as adding qualitative evidence, using time-series data, investigating possible alternative casual hypotheses, or being guided by precise theory (Eck, 2002; Knutsson, 2009; Pawson and Tilley, 1997; Tilley, 2009). This pragmatic view has generated compelling

09012-CrimJournal-Guts.indd 174 12/21/09 7:14:26 PM

case studies and a growing body of practical evidence that supports problem-oriented policing as a promising approach to crime prevention (Skogan and Frydl, 2004; Weisburd and Eck, 2004). Unfortunately, this practical approach to generate knowledge on improved police crime prevention practices is problematic in regard to generating more rigorous scientific evidence in the long term. If academics interested in exploring the effects of problem-oriented policing on crime are not being encouraged by their peers to use more rigorous research designs, then problem-oriented policing will never have a strong evaluation research base. As "evidence-based" approaches, with their premium on rigorous evaluation designs, gain currency in policy circles, funding for continued expansion of the problem-oriented policing agenda might be threatened.

In reality, fewer police departments will be willing to host more intrusive research designs such as randomized controlled trials. As Eck (2002) described, there are many reasons why police departments might not be willing to play host to an experiment—the department might have other pressing business, the department's leadership might not be tested, and other reasons. When a police department is willing to engage a more rigorous design, the external validity of the findings are called into question because the host departments and their contexts are described as not representative of all agencies or contexts that use the intervention. Drawing on Campbell and Stanley (1966), Eck (2002: 104) identified this issue as a problem of the "interaction of selection and treatment."

Although this point is legitimate, all problem-oriented policing evaluations suffer from external validity concerns regardless of the degree of internal rigor in the evaluation research design. Problem-oriented policing is primarily an analytic approach to crime prevention that requires customizing interventions to highly localized crime and disorder problems. What works in preventing a street robbery problem in the public areas of Harvard Square in Cambridge, Massachusetts, might not work when applied to repeated robberies occurring in the London Underground subway system. Appropriate interventions need to be applied in both contexts that are based on careful analysis of the conditions that create the compelling criminal opportunities. Neither the evaluation findings of a carefully constructed single group before–after design nor the findings of a randomized controlled experiment will travel perfectly across these settings. Problem-oriented policing evaluations of many forms provide valuable guidance to police officers struggling with real-world problems. However, given the highly customized nature of effective problem-oriented policing interventions, it is important to recognize that the generalizability of specific crime prevention practices identified in an effective application of the approach might be limited, regardless of the evaluation approach used.

Randomized controlled trials and more rigorous quasi-experimental designs can be developed to maximize their external validity. A well-designed test can collect the necessary qualitative and quantitative data to understand the processes that underlie problem-oriented policing and "why" it might impact crime and disorder problems. For instance, the units of analysis in the evaluation design—whether they are problem places, problem people, hot products, or some other unit—must be analyzed thoroughly to ensure equivalent treatment and control groups.

09012-CrimJournal-Guts.indd 175 12/21/09 7:14:27 PM

Although this task is time consuming, it is by no means insurmountable through established matching techniques (that range from descriptive exercises to more formal propensity score analyses) and subsequent allocation from like pairs or statistically comparable blocks to treatment and control conditions. The thorough description of problem units provides police practitioners with important information regarding the problems addressed in the study so that they can determine whether any effective interventions are applicable to problems faced in their local jurisdictions.

Rigorous evaluation designs can accommodate the reality of problem-oriented policing as practiced in the field. An important consideration in gaining real-world insights on the crime prevention value of problem-oriented policing in dealing with crime and disorder problems is to test the general approach rather than any specific set of tactics. In two randomized controlled experiments (Braga and Bond, 2008; Braga, Weisburd, Waring, Mazerolle, Spelman, and Gajewski, 1999), the treatment was the meta-method known as "problem-oriented policing," which comprised several specific operational tactics implemented by the officers to control the identified problems at treatment places.[1] Careful quantitative and qualitative documentation of the content of these problem-oriented interventions provided police practitioners with a clear sense of the types of responses to address problem places so that they can develop knowledge on a range of strategies that could be plausibly effective in their jurisdictions. Indeed, both experiments documented interventions that closely resembled problem-oriented policing as it is usually practiced in the field—"shallow" problem solving characterized by weak problem analysis and limited response development.

When interventions are described in sufficient detail, further analyses, such as mediation analysis, can be used to determine the key elements of the intervention that seem to be driving any observed crime control gains. Using this approach, a randomized controlled trial of problem-oriented policing in Lowell, Massachusetts, found that the strongest crime prevention gains at crime and disorder hot spots were driven by situational prevention strategies rather than by misdemeanor arrest strategies or short-term, police-led social service strategies (Braga and Bond, 2008). Because many jurisdictions deal with crime and disorder hot spot locations, the basic findings of this study provide important guidance to police executives interested in maximizing their ability to manage similar problem places.

Creating the Opportunity to Conduct More Rigorous Evaluations

Academics must work to create the opportunities to test problem-oriented policing interventions using more rigorous designs with higher degrees of internal validity. Unfortunately, past partnerships between academics and police practitioners often are characterized by role conflicts such as researchers reporting the "bad news" that an evaluated program was not effective in preventing crime (Weisburd, 1994). For academic researchers, success or failure matters less

1. The Braga and Bond (2008) randomized controlled trial was completed after Weisburd et al. (2010) finished their systematic search for eligible problem-oriented policing studies and, as such, is not included in their review.

09012-CrimJournal-Guts.indd 176 12/21/09 7:14:27 PM

than their commitment to the development of knowledge of what does or does not work in preventing crime. For police practitioners, this news could be interpreted as their personal failure, and the skepticism of academics could be viewed as irritating. The police also can view traditional research and evaluation roles played by academics (often involving data collection and analysis *after* programs have been developed) as not particularly helpful in their efforts to prevent crime.

In recent years, however, the demand for the participation of academic researchers in problem-oriented policing projects has increased because the police have come to recognize the importance of strategic information products in developing effective crime prevention interventions. Academics can help the police by conducting research on urban crime problems to focus limited enforcement, intervention, and prevention resources on high-risk offenders, victims, and places. Problem-oriented interventions based on research insights have been associated with a 60% reduction in youth homicide in Boston (Braga, Kennedy, Waring, and Piehl, 2001) and a 40% total homicide reduction in Indianapolis (McGarrell, Chermak, Wilson, and Corsaro, 2006). These success stories have made academic researchers an important part of new crime prevention initiatives. For example, the U.S. Department of Justice-sponsored Project Safe Neighborhoods (PSN) initiative provided each of the 94 U.S. Attorney's districts with funds to hire academic research partners to help understand and address serious gun violence problems in local jurisdictions. A recent national evaluation of PSN found that treatment cities with high levels of implementation, which include in-depth problem analysis research to tailor prevention strategies, were associated with declines in violent crime (McGarrell et al., 2009).

The increasing police openness to the assistance of academics in understanding the nature of crime problems provides an important opportunity to develop higher quality evaluations of problem-oriented policing interventions. As a partner in the problem-oriented policing process, academics can make a case for using a more rigorous research design to assess the implemented interventions. Many current police executives understand the importance of determining "what works" in police crime prevention efforts.[2] Savvy police executives also understand that simple pre–post analyses of time-series data are not scientifically rigorous enough to provide strong evidence of a program effect. Indeed, the strong skepticism and dismissive assertions made by certain academics in their assessment of the role of innovative policing strategies in the 1990s crime drop (e.g., Levitt, 2004) made an impression on the police profession. Although some police executives (most notably former New York Police Commissioner and former Los Angeles Police Chief William Bratton) attacked the credibility of "ivory tower" academics in assessing police crime prevention strategies, others such as Redlands, California, Police Chief James Bueerman and former Jersey City, New Jersey, Police Chief Francis Gajewski opened

2. For instance, the Police Executive Research Forum (PERF), which has police chiefs completely comprise its board of directors, states "substantial and purposeful academic study is a prerequisite for acquiring, understanding and adding to the body of knowledge of professional police management" as one of its organizational principles. PERF also published an edited volume of essays called *Police Program Evaluation* that makes the case that meaningful program evaluation often requires going beyond piecemeal observations or simple "before and after" comparisons (Hoover, 1997).

09012-CrimJournal-Guts.indd 177 12/21/09 7:14:27 PM

their departments to academics looking to conduct rigorous tests of problem-oriented policing strategies (e.g., Braga et al., 1999; Weisburd, Morris, and Ready, 2008).[3]

Although other important factors undoubtedly exist, ongoing collaborative relationships between police executives and academics create opportunities to conduct rigorous evaluations of problem-oriented policing initiatives. In an inspiring description of her rewarding experiences as an "embedded" criminologist in California's prison reform efforts, Joan Petersilia (2008) rightly observed that "timing is *everything*" when working with policy makers and criminal justice practitioners. Her observation is particularly salient when considering the legal, political, institutional, and resource constraints inherent in the successful execution of randomized controlled trials in criminal justice agencies.

Throughout the course of a long-standing collaborative relationship with a police executive, "timing" directly affected my ability to conduct more rigorous evaluations of problem-oriented policing initiatives. In 1997, I formed a collaborative relationship with Edward F. Davis III when he was Superintendent of the Lowell Massachusetts Police Department. During the next 6 years, my colleagues and I worked with the Lowell Police on a series of problem analyses of gang violence problems and conducted a quasi-experimental evaluation of a problem-oriented intervention to guide their gang violence reduction efforts (Braga, McDevitt, and Pierce, 2006; Braga, Pierce, McDevitt, Bond, and Cronin, 2008). In 2004, Davis expressed a desire to make a substantive contribution to the policing field by conducting a more rigorous test of the effects of problem-oriented policing strategies on crime and disorder hot spots (see Braga and Bond, 2008). We collaborated on the design of a randomized controlled experiment and jointly made a successful argument to the Massachusetts Executive Office of Public Safety for research funds. It is important to note that, at that point in time, Davis had served as the chief of the Lowell Police for nearly 10 years, crime was low in the city, research funding was available, and Davis had gained considerable trust and legitimacy among community members and politicians through his commitment to community policing. In other words, the timing was ripe for experimentation.

In December 2006, Davis was appointed Commissioner of the Boston Police Department. Boston was facing a troubling increase in gun violence when he took charge of the police force. Davis immediately led the department in developing a problem-oriented policing initiative that targeted gun violence hot spots in Boston that borrowed heavily from the lessons of the Lowell experiment and other evidence-based crime prevention practices. However, at that point in time, I could not pursue a randomized controlled trial of the new program. The resulting evaluation design was the modal single group before–after design supplemented by qualitative evidence and longer time-series data. Indeed, the conditions were different in Boston compared with Lowell a few years earlier. His mandate from Mayor Thomas M. Menino was clear. Davis needed to

3. In his memoir on crime fighting in Boston and New York City, Bratton (1998: 289) stated that "Criminologists apparently still had a hard time accepting the reality of our success. I made a conscious decision to take on the academics, to challenge conventional wisdom about crime in America and prove that effective policing can make a substantial impact on social change."

09012-CrimJournal-Guts.indd 178 12/21/09 7:14:27 PM

make Boston's neighborhoods safe by addressing the gun violence wherever it presented itself in the city. He was also a new chief who needed to develop the necessary political support to permit potentially controversial decisions such as creating treatment and control conditions in gun violence hot spots.

These experiences suggest that academics need to be "commonsense opportunists" who are sensitive to the varying needs of their practitioner partners throughout the course of a long-term working relationship. Sometimes academics can pursue more rigorous evaluation designs. Other times academics will need to respect the political environment and simply provide program development and implementation advice that is rooted in solid research evidence.

Concluding Thoughts

The purpose of this policy essay was not to make an argument that all evaluations of problem-oriented policing need to be randomized controlled trials. Indeed, weaker evaluations can, and do, produce important findings. Case studies on the problem-oriented policing process can provide rich insights and much-needed guidance to practitioners struggling with substantive public safety problems. Nevertheless, the Weisburd et al. (2010) review suggests that the need is for more rigorous research in this important area. Weaker evaluations outnumber the stronger evaluations by almost five times. Problem-oriented policing scholars, who are mostly likely to be in the position to conduct such studies, should recognize the importance of developing a stronger evidence base for the approach and foster an environment that encourages more rigorous research designs.

To some observers, long-term working relationships as described in this essay might violate the purported scientific necessity to separate program developer and evaluator roles (for discussion, see Petrosino and Soyden, 2005). However, unless some convincing evidence of widespread evaluator bias is associated with such arrangements, these collaborative arrangements seem necessary to put academics in the position of being able to conduct higher quality evaluations of problem-oriented policing initiatives. As David Olds (2009) suggested in his recent essay in support of "disciplined passion," balancing scientific integrity with the practical challenges that are associated with program evaluation in real-world settings needs to be addressed through higher standards for reporting trials, better peer review, improved investigator training, and rigorous collegial support of those who choose this line of work.

Although this essay focused on evaluation and assessment issues, it is important to recognize that a large gap continues to exist between Goldstein's (1979) vision of problem-oriented policing and its implementation by most police departments. Research has documented that it is difficult for police officers to implement problem-oriented policing strategies (Clarke, 1998; Cordner and Biebel, 2005; Eck and Spelman, 1987). Deficiencies exist in all phases of the problem-oriented policing process. Growing evidence suggests that even the imperfect implementation of problem-oriented policing (so-called "shallow" problem solving) generates crime prevention gains (Braga and Weisburd, 2006). Improving the problem-oriented policing process, however, could produce even stronger crime control effects. Police executives interested in improving the

09012-CrimJournal-Guts.indd 179 12/21/09 7:14:27 PM

capacity of their agency to deal with crime through more robust problem solving should seek the involvement of academics the same way that they seek assistance from community partners and other city departments. Beyond evaluation, academics can be helpful in improving the quality of problem analysis and broadening the search for alternative responses.

Federal funding initiatives (similar to PSN) that provide support for academic–police practitioner partnerships could raise the quality of problem-oriented policing projects and improve our knowledge base on effective practice. Unfortunately, the number of academics with the experience and expertise in working with police departments on problem analysis, response development, and evaluation is relatively small. Although additional funding provides an incentive, the real challenge is to increase these collaborations by educating police departments on the benefits of working with academics and to encourage uninvolved academics to learn more about and participate in problem-oriented policing projects. Establishing new and more productive relationships will require flexibility and openness on both parts. Academics cannot be condescending in their relationships with the police; they must demonstrate a genuine interest in helping the police to be more effective, and they should select research methodologies that produce credible results but are not disproportionately demanding given the limits of available data as well as the nature of the strategy being evaluated (Goldstein, 2003). The police must recognize that quality research takes time and patience, researchers must be objective and require facts to support conclusions, and evaluations might produce results that could be interpreted as critical of past and current operations (Goldstein, 2003).

In the brief space given to this policy essay, all factors necessary to produce effective working partnerships between academics and police practitioners cannot be covered adequately. It does seem obvious that both camps as well as the general public would benefit greatly from a sustained collaborative relationship focused on understanding the nature of crime problems and on determining what works in dealing with the underlying conditions and situations that cause them to persist.

References

Boruch, Robert F. 1975. On common contentions about randomized field experiments. In (Robert F. Boruch and Henry L. Reicken, eds.), *Experimental Testing of Public Policy: The Proceedings of the 1974 Social Sciences Research Council Conference on Social Experimentation*. Boulder, CO: Westview Press.

Braga, Anthony A. and Brenda J. Bond. 2008. Policing crime and disorder hot spots: A randomized controlled trial. *Criminology*, 46: 577–608.

Braga, Anthony A., David M. Kennedy, Elin J. Waring, and Anne M. Piehl. 2001. Problem-oriented policing, deterrence, and youth violence: An evaluation of Boston's Operation Ceasefire. *Journal of Research in Crime and Delinquency*, 38: 195–225.

Braga, Anthony A., Jack McDevitt, and Glenn L. Pierce. 2006. Understanding and preventing gang violence: Problem analysis and response development in Lowell, Massachusetts. *Police Quarterly*, 9: 20–46.

09012-CrimJournal-Guts.indd 180 12/21/09 7:14:27 PM

Braga, Anthony A., Glenn L. Pierce, Jack McDevitt, Brenda J. Bond, and Shea Cronin. 2008. The strategic prevention of gun violence among gang-involved offenders. *Justice Quarterly*, 25: 132–162.

Braga, Anthony A. and David L. Weisburd. 2006. Problem-oriented policing: The disconnect between principles and practice. In (David Weisburd and Anthony A. Braga, eds.), *Police Innovation: Contrasting Perspectives*. Cambridge, MA: Cambridge University Press.

Braga, Anthony A., David L. Weisburd, Elin J. Waring, Lorraine Green Mazerolle, William Spelman, and Francis Gajewski. 1999. Problem-oriented policing in violent crime places: A randomized controlled experiment. *Criminology*, 37: 541–580.

Bratton, William. 1998. *Turnaround: How America's Top Cop Reversed the Crime Epidemic.* New York: Random House.

Campbell, Donald T. and Julian C. Stanley. 1966. *Experimental and Quasi-Experimental Designs for Research.* Chicago, IL: Rand McNally.

Clarke, Ronald V. 1998. Defining police strategies: Problem solving, problem-oriented policing, and community-oriented policing. In (Tara O'Connor Shelley and Anne Grant, eds.), *Problem-Oriented Policing: Crime-Specific Problems, Critical Issues, and Making POP Work.* Washington, DC: Police Executive Research Forum.

Cook, Thomas D. and Donald T. Campbell. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Boston, MA: Houghton Mifflin Company.

Cordner, Gary and Elizabeth Perkins Biebel. 2005. Problem-oriented policing in practice. *Criminology & Public Policy*, 4: 155–180.

Eck, John E. 2002. Learning from experience in problem-oriented policing and situational prevention: The positive functions of weak evaluations and the negative functions of strong ones. In (Nick Tilley, ed.), *Evaluation for Crime Prevention*. Monsey, NY: Criminal Justice Press.

Eck, John E. and William Spelman. 1987. *Problem-Solving: Problem-Oriented Policing in Newport News.* Washington, DC: Police Executive Research Forum.

Goldstein, Herman. 1979. Improving policing: A problem-oriented approach. *Crime & Delinquency*, 25: 236–258.

Goldstein, Herman. 2003. On further developing problem-oriented policing: The most critical need, the major impediments, and a proposal. In (Johannes Knutsson, ed.), *Problem-Oriented Policing: From Innovation to Mainstream.* Monsey, NY: Criminal Justice Press.

Hoover, Larry. 1997. *Police Program Evaluation*. Washington, DC: Police Executive Research Forum.

Knutsson, Johannes. 2009. Standards of evaluations in problem-oriented policing projects: Good enough? In (Johannes Knutsson and Nick Tilley, eds.), *Evaluating Crime Reduction Initiatives.* Monsey, NY: Criminal Justice Press.

Levitt, Steven D. 2004. Understanding why crime fell in the 1990s: Four factors that explain the decline and six that do not. *Journal of Economic Perspectives*, 18: 163–190.

09012-CrimJournal-Guts.indd 181 12/21/09 7:14:28 PM

McGarrell, Edmund, Steven Chermak, Jeremy Wilson, and Nicholas Corsaro. 2006. Reducing homicide through a "lever-pulling" strategy. *Justice Quarterly*, 23: 214–229.

McGarrell, Edmund, Natalie Kroovand Hipple, Nicholas Corsaro, Timothy Bynum, Heather Perez, Carol A. Zimmerman, and Melissa Garmo. 2009. *Project Safe Neighborhoods—A National Program to Reduce Gun Crime.* East Lansing: Michigan State University.

Olds, David. 2009. In support of disciplined passion. *Journal of Experimental Criminology*, 5: 201–214.

Pawson, Ray and Nick Tilley. 1997. *Realistic Evaluation.* Thousand Oaks, CA: Sage.

Petersilia, Joan. 2008. Influencing public policy: An embedded criminologist reflects on California prison reform. *Journal of Experimental Criminology*, 4: 335–356.

Petrosino, Anthony and Haluk Soyden. 2005. The impact of program developers as evaluators on criminal recidivism: Results from meta-analyses of experimental and quasi-experimental research. *Journal of Experimental Criminology*, 1: 435–450.

Skogan, Wesley G. and Kathleen Frydl (eds.). 2004. *Fairness and Effectiveness in Policing: The Evidence.* Washington, DC: National Academies Press.

Tilley, Nick. 2009. What's the "what" in "what works?" Health, policing, and crime prevention. In (Johannes Knutsson and Nick Tilley, eds.), *Evaluating Crime Reduction Initiatives.* Monsey, NY: Criminal Justice Press.

Weisburd, David. 1994. Evaluating community policing: Role tensions between practitioners and evaluators. In (Dennis P. Rosenbaum, ed.), *The Challenge of Community Policing: Testing the Promises.* Thousand Oaks, CA: Sage.

Weisburd, David. 2003. Ethical practice and evaluation of interventions in crime and justice: The moral imperative for randomized trials. *Evaluation Review,* 27: 336–354.

Weisburd, David and John E. Eck. 2004. What can police do to reduce crime, disorder, and fear? *Annals of the American Academy of Political and Social Science*, 593: 42–65.

Weisburd, David, Nancy A. Morris, and Justin Ready. 2008. Risk-focused policing at places: An experimental evaluation. *Justice Quarterly*, 25: 163–200.

Weisburd, David, Cody W. Telep, Joshua C. Hinkle and John Eck. 2010. Is problem-oriented policing effective in reducing crime and disorder? Findings from a Campbell systematic review. *Criminology & Public Policy*. This issue.

Anthony A. Braga is a Senior Research Associate in the Program in Criminal Justice Policy and Management and Lecturer in Public Policy at Harvard University's John F. Kennedy School of Government, as well as a Senior Research Fellow in the Berkeley Center for Criminal Justice at the University of California, Berkeley. He is also a fellow of the Academy of Experimental Criminology and former visiting fellow at the U.S. National Institute of Justice. His research focuses on working with criminal justice agencies to develop crime prevention strategies to deal with urban problems such as gang violence, illegal gun markets, and violent crime hot spots. He received his M.P.A. from Harvard University and his Ph.D. in criminal justice from Rutgers University.

09012-CrimJournal-Guts.indd 182 12/21/09 7:14:28 PM

Whither problem-oriented policing

Nick Tilley
University College London

Herman Goldstein's (1979, 1990) vision of problem-oriented policing (POP) has not been difficult to sell, at least in principle, and in the 30 years since he first mooted it many scholars and police agencies have bought into it. POP promises a move away from managerial priorities, standardized (and often ineffective) ways of working, and law enforcement as an end in itself. It focuses instead on relevant community concerns, the open-minded pursuit of ethical and effective solutions to recurrent problems, and attention to results to learn better how to deal with similar problems in the future. It also suggests an efficient and effective way of handling heavy demands on police time, by resolving problems rather than simply by responding to them incident by incident.

The idea of focusing on real recurrent problems, working out why they have persisted despite conventional responses, analyzing their underlying sources, figuring out what might be done to ameliorate or remove them on the basis of that analysis, and then checking whether the strategy has produced its expected benefits, seems like plain common sense. Who could disagree? POP is also appealing because of its associations with evidence-based practice and with the application of scientific methods to deal with real community problems (Laycock, 2004). POP's core concern is with what works practically to deal with specific problems, whether the response is delivered by the police or some other agency or both, provided that what works is also ethical. It eschews the authority of tradition, convention or superstition. It comprises a thoroughly "modern" approach to policing.

POP has been widely attempted on both sides of the Atlantic. In the United States many police agencies have adopted it (Scott, 2000). In the United Kingdom it has also been popular (Read and Tilley, 2000), albeit that following the Crime and Disorder Act (1998), which has broadened the responsibility for crime and disorder to a range of local agencies beyond the police, the second "P" now generally refers to partnership rather than "policing". The thinking

Karen Bullock, Gloria Laycock, and Aiden Sidebottom kindly read and commented on earlier versions of this essay. Their help is gratefully acknowledged. Direct correspondence to Nick Tilley, Department of Security and Crime Science, University College London, Second Floor Brook House, 16 Torrington Place, London WC1E 7HN, United Kingdom (e-mail: n.tilley@ucl.ac.uk).

© 2010 American Society of Criminology

09012-CrimJournal-Guts.indd 183 12/21/09 7:14:28 PM

behind "Problem Oriented Partnership", however, is fundamentally the same as that set out by Goldstein.

In view of its evident enduring appeal both in the United States and in other jurisdictions, notably the United Kingdom, it seems proper and timely to ask whether the promise of POP has been delivered. Is POP a "best buy"? Is it even a "good" one? Weisburd, Telep, Hinkle, and Eck (2010, this issue) look for an answer by sifting meticulously through the now voluminous literature to pick out relevant studies that meet their methodological standards (the inclusion of a non-treatment comparison group, at least one crime/disorder outcome, and enough data to calculate an effect size) to see what, on balance, is to be found from them. Have those of us who are associated with POP bought (or sold) a pig in a poke: something that we are told (or tell) about and believe in without actually seeing (or showing) the product itself?

Relatively little of the quite large literature on POP was eligible for inclusion once it was filtered through the relevance and methodological quality control vetting mechanisms: 4 studies in all from 5,282 initial database hits and 628 that related specifically to POP, plus another 6 studies that were identified by a most careful search of the grey literature. To my relief, and slight surprise, the overall verdict from the 10 studies that met the eligibility criteria was positive, although the overall effects are said to be "modest but statistically significant". Weisburd et al. (2010) were troubled at the shortage of studies that met their criteria and, thus, the material on which they felt they could base their conclusions. Moreover, even amongst the 10 studies that did meet the technical specifications for inclusion, Weisburd et al. report widespread implementation weaknesses: Only 1 of the 10 "reported full implementation without any significant problems".

Why should I have been surprised that even a small effect size was found across the 10 studies? There are three main reasons. First, POP is used, at least in theory, when there are persistent problems that have resisted conventional, known responses. The specifics of the problem were interrogated to find some point of intervention that would produce a reduction in the problem, or ideally its elimination. Think of the parallels in medicine. There is a pattern of symptoms. Current understanding and associated treatment methods have failed to deal with the issue. The problem is interrogated carefully and a new treatment is applied. What success rate would we expect if we took a representative sample of such innovative treatments? At what rate would we expect success to exceed that which would have been achieved by the natural course of the disease, with no treatment? Pretty low, I would hazard. Solving unsolved knotty problems is difficult. That any detectable impact could be found across a small number of reasonably typical POP efforts seems to me to be quite remarkable. Second, Weisburd et al. (2010) report patchy implementation. That a small net impact across studies was found with patchy implementation on top of intrinsically tricky issues that have not responded to conventional treatment seems doubly remarkable. Third, in medicine we have established bodies of theory which have survived empirical tests, highly trained researchers, massive funding, and substantial levels of institutional support for devising solutions to knotty problems. In policing we often have clever and enthusiastic practitioners, but their starting point is much less

09012-CrimJournal-Guts.indd 184 12/21/09 7:14:28 PM

promising than it is for those addressing the equivalent types of knotty problem in medicine. Yet a positive effect was still detected.

Am I fully convinced by the welcome findings at which I am slightly surprised? Regrettably there are at least four reasons not to be fully confident in them. The first relates to causal attribution. Much can be done by the police service as part of a problem-oriented policing package, especially when a formally (and expensively) evaluated project is undertaken. For example, the most enthusiastic and experienced police officers can be seconded, more supervisory attention provided, performance measures tracked, and greater publicity and extra resources can be distributed. These in themselves could produce the effects rather than the process of problem-orientation per se.[1] The second relates to comparison areas. Local crime rates are susceptible to wide short-term fluctuations and unless rates have tracked one another closely for a suitably long period of time then the counterfactual provided by the comparison area is dubious. The third reason relates to the compensatory efforts that can be made by those in comparison areas to match or better the results of the experimental area or to under-perform in the interests of securing future resources. The point is that there are patterns of adaptation in local areas that can be tricky to control for. The fourth relates to the specificity of local problem solving that focuses on the distinctive idiosyncrasies of the particular presenting problem and the reduced scope this provides for finding appropriate equivalent comparison areas. In some studies multiple control areas and multiple experimental areas are used to try to overcome at least some of these difficulties, but not in all cases. There is much devil in the details of any initiative and these details are steamrolled over in meta-evaluations.

Weisburd et al. (2010) recognize some of the problems of evaluating POP using standard experimental and quasi-experimental methods. They acknowledge that the POP experts they consulted as part of their review were skeptical of reliance on studies with this design as a basis for coming to a verdict about POP. In response, Weisburd et al. carefully searched out pre–post studies that fell somewhat short of the stricter methodological criteria met by the initial 10 studies. They found 45 further studies. Many were entries for awards that are made on either side of the Atlantic for excellence in problem-oriented efforts. Unsurprisingly, much larger effect sizes were found. Prizes are not generally awarded for projects that have failed to have had any effect and there are likely publication biases in relation to those that are published. It is difficult to determine on this basis what such studies say about the effectiveness of POP as it is ordinarily delivered, although they comprise fairly convincing evidence that, at least sometimes, POP processes can lead to strategies that reduce crime.

Overall the Campbell review, notwithstanding the slightly skeptical comments made so far in this essay, makes a pretty good case that problem-oriented work can, at least sometimes, reduce crime and disorder. It accepts that evaluations using standard (quasi-experimental or experimental) methods may sometimes not be possible. It also notes that implementation problems are more or less chronic features of POP as it is practiced. These are all helpful find-

1. They may also, of course, make the POP work that was assessed in the evaluations atypical.

09012-CrimJournal-Guts.indd 185 12/21/09 7:14:28 PM

ings, particularly given the strong pedigree of Weisburd et al. (2010) and the assiduousness with which they went about their task.

It is important to remember, however, the strict limitations of even the best of randomized controlled trials (RCTs; the standards for which, for example in terms of blinding/masking those who designed, delivered and received the program and also who collected and analyzed the evaluation data, were probably met by none of the studies included). Even the purest of RCTs necessarily focus on spatio-temporally specific populations. Generalizations beyond these are strictly acts of faith, as was recognized long ago by David Hume (1739). Moreover the assumptions about the "uniformity of nature" that lie (perfectly reasonably) behind work in the physical sciences are much more problematic in the social sciences, where the domain is constituted by reflective, adaptable, intentional agents. All that can be said about effectiveness (which is a lot) is that the evidence assembled strongly suggests that POP can reduce crime.

What are the policy implications of the Weisburd et al. (2010) findings and of what else we know of POP? The terms of the invitation to write this essay from the editorial staff of *Criminology & Public Policy* were clear:

> The ideal policy essay considers salient issues in the available relevant literature—with a special focus upon the article to which the essay responds—and does not conclude that more and better research is needed to inform public policy; instead, the best policy essays to date have identified and discussed specific and actionable policy steps based upon the current state of knowledge (limited as it may be).

I am sympathetic to this brief. I do not think more studies that meet the testing Campbell standards are needed to establish the case for implementing POP. The possibility of human flight was established with one successful example. Thereafter the agenda for research and development was to improve the reliability and efficiency of flight. Thanks in part to Weisburd et al. (2010), there now is convincing evidence that POP can reduce crime and disorder. The research and development agenda for POP is now that of improving its efficiency, reliability in producing intended outcomes. This is undoubtedly, however, where many problems lie.

Weisburd et al. (2010) refer to chronic implementation weaknesses in what is undertaken in the name of POP. These weaknesses are of three kinds. The first relates to the conduct of POP. The Scanning, Analysis, Response, and Assessment (SARA) steps that were sketched by Eck and Spelman in 1987 (and required as one of the eligibility criteria for the studies included in Weisburd et al.'s review), can be done better or worse. In many cases the delivery of SARA is rather desultory and formulaic: The motions may be gone through nominally but what is done falls far short of the rather more demanding processes of specific problem identification, detailed analysis for causal pinch-points and careful trialling and adaptation of thoughtfully chosen interventions (Knutsson, 2003; Knutsson and Clarke, 2006; Leigh, Read, and Tilley, 1998).

The second type of implementation problem relates to the delivery of chosen measures on the ground. Chronic difficulties have been identified in this, as with all crime reduction projects

09012-CrimJournal-Guts.indd 186 12/21/09 7:14:29 PM

and, indeed, with public policy more generally (Bullock, 2007; Bullock and Tilley, 2003; Hope and Murphy, 1983; Laycock and Tilley, 1995). Even though a strategy might have been well chosen, persuading relevant stakeholders to act on it consistently proves very difficult.

The third implementation problem relates to the requirements for evaluation. Recalcitrant organizations often fail to adapt and maintain their practices in accordance with the strict requirements of action and inaction called for in experimental and quasi-experimental evaluation designs (Bullock and Tilley, 2003; Hanmer, 2003; Homel et al., 2004). Enacting POP properly, which involves adaptation and adjustment in the light of experience, may conflict with doing POP in a way that is most readily evaluated using the methods preferred by the Campbell Collaboration. Maybe different methods are needed.

What can be done to help remedy these implementation weaknesses in the interests of improving the yield from POP in future? Let us take each issue in turn. In the end we will find that developments that might help one will have implications for the others.

The Conduct of POP

A great deal has already been done to support improvements in the conduct of POP. Many of these emanate from the Center for Problem-Oriented Policing (popcenter.org), which gathers together a wide range of resources that aim to enhance the delivery of POP. They include:

- Clarke and Eck's *Become a Problem-Solving Crime Analyst* (2003). This guide (now available in many languages) is aimed at supporting the analysis function which underpins POP but contains a wealth of information about the process as a whole
- More than 50 guides (whose preparation has been supported by the U.S. Department of Justice), which aim to help practitioners think through common community problems, together with parallel guides on some commonly used responses and on tools that can be used in problem solving
- The full text of the first 18 volumes of the *Crime Prevention Studies* series
- Teaching materials, for example, a model curriculum for teaching POP

The effects of this growing resource on the fidelity of what is delivered in the name of POP to its underlying theory are not yet clear, but the materials provided aim to address the widespread implementation weaknesses that have been identified. For any one agency (or partnership) that wishes to deliver POP in a way that reflects the thinking behind the approach, the popcenter.org Web site provides extensive advice on what needs to be done and how to do it.

The British approach has been to make attention to crime and disorder problems using a POP approach a statutory responsibility. The Crime and Disorder Act (1998) followed many previous years during which a partnership approach had been promoted by successive governments, going back to the Interdepartmental Circular on Crime Prevention issued in 1984 (Home Office, 1984). The growth of the partnership approach has been accompanied by much guidance and exhortation supporting problem solving, a great deal of which is available through the Home Office Web site (crimereduction.homeoffice.gov.uk/).

09012-CrimJournal-Guts.indd 187 12/21/09 7:14:29 PM

More recently "neighborhood policing" has been championed in the United Kingdom, with the establishment of small, dedicated Neighborhood Policing Teams (NPTs), which cover the whole country. Although their aims are wider than those of POP, NPTs have the same process at their heart. Some central support had been provided for problem solving by NPTs and, indeed, one of the studies included in Weisburd et al.'s (2010) top ten that meet their methodological inclusion criteria (Tuffin, Morris, and Poole, 2006) was of a pilot of this initiative (although the later rolling out of the approach produced rather less promising findings, albeit that the reason—again—seems largely to do with implementation rather than substance; see Quinton and Morris, 2008).

In both the United Kingdom and the United States annual problem-solving conferences are held, where practitioners can learn from one another and from national and international authorities on POP. These conferences are also occasions for making awards that are designed to act as incentives for high-quality work, some of which is reflected amongst the studies included in the Campbell Collaboration review when it extended its reach beyond those reports satisfying its preferred methodological standards.

Many police services (and partnerships) are employing analysts using increasingly sophisticated software for the collection, mapping and analysis of crime and disorder data as well as other local information that could be important for understanding local problems. In London's Metropolitan Police, for example, more than 400 analysts are employed, and most are highly trained. This situation is very different from that in 1979, when Goldstein first thought of POP, and provides an information infrastructure that would not then have been dreamed of.

Is there anything else that can be done? Can these developments, which certainly appear to be conducive to the development of POP, be expected over a period of time to stimulate strongly implemented versions? One would hope so. If we take POP to be the crime and disorder counterpart to scientific medicine and health improvement, however, rather more substantial changes may be needed in addressing many problems.

Much of the most notable POP work has involved criminologists, for example the Boston Gun Project (Kennedy, Braga, Piehl, and Waring, 2001; Kennedy, Piehl, and Braga, 1996), the Rochdale burglary reduction work (Forrester, Chatterton, and Pease, 1988; Forrester, Frenz, O'Connell, and Pease, 1990) and the construction site theft reduction work in Charlotte–Mecklenburg (Clarke and Goldstein, 2003). In each case there was substantial involvement from informed and able professional social scientists. In this sense the conduct of the projects was similar to the ways in which researchers in health work alongside clinicians to try to find solutions to health problems. In health, the orientation to evidence-based practice is embraced both by the researchers and practitioners. In policing there is no such tradition. The exceptional POP work is undertaken in exceptional conditions. It seems likely that POP would improve by closer work and greater mutual understanding between the police (and their partners) and the research community, where there is a shared commitment to finding practical solutions to crime and disorder problems.

09012-CrimJournal-Guts.indd 188 12/21/09 7:14:29 PM

What the research community brings is not so much the common sense that is used in some problem solving as it deals with fairly simple issues, as an immersion in much tested theory that might be applied to more complex presenting problems. Members of the research community also bring strong analytic techniques. These can be brought to bear when determining how a given intervention, that aims to apply a certain theory in a particular problem context, will generate specific expected outcomes. The research community's habits of thought and of work tend to be missing amongst those who value their "can do" practical skills. Situational crime prevention, with its own emphasis on action research and well-developed theories, has been found to be a comfortable academic bedfellow to POP, although there may be other traditions in criminology (and other disciplines) that could also be drawn on.

The SARA (Scanning, Analysis, Response, Assessment) mnemonic has provided an easy way of explaining what is involved in POP, and it is used for headings for the aforementioned award entries. It is sometimes used as a "model" to be followed by police services that embrace POP, with attendant forms to be completed (and boxes to be ticked) to show that the steps have been followed. Yet critics argue that this is too simple and too formulaic. It is not difficult to re-describe normal practice in SARA terms. SARA is unexceptionable, perhaps, because it is so undemanding—producing pseudo-POP, rather than real POP. Instead of SARA, Ekblom (2005) proposed the 5Is (Intelligence, Intervention, Implementation, Involvement, Impact), which are intended better to capture what POP calls for and ideally does. Previously, Read and Tilley (2000) had likewise proposed PROCTOR (PROblem, Cause, Tactic/Treatment, Output, Result) as an alternative to SARA, but that was little taken up. SARA has, however, taken hold and has undoubtedly helped promote POP. But it has been a mixed blessing.

It might be neither possible nor useful to dislodge SARA. Being clearer about what is involved in the iterative processes of problem solving, as problem to effective solution evolves, might, however, help improve the quality of problem solving. In the fields of engineering and health, recursive processes of problem redefinition and solution refinement that are informed by analyses and trialling, seem characteristically to take place rather than the simple stage-by-stage path that might seem to be implied by SARA. The rather messy and often slow and frustrating worlds of problem solving in engineering, medicine and health appear alien in relation to crime and disorder problem solving, where quick fixes are frequently called for. Habits of collaborative work with universities and an increasingly educated police and partnership workforce might improve the working conception and working practices of POP.

In the field of health improvement, which has many parallels with that of POP, and where the substance of what is done is often problem-solving, William Deming's recursive PDSA (Plan-Do-Study-Act) cycles are often used. The underlying philosophy is that of continuous improvement. In practice this often means being attentive to specific outcome problems (for example those related to "adverse drug events", "surgical site infections", "central line bloodstream infection", "intensive care unit mortality rate", and "falls"). The problem is diagnosed. Measures and mechanisms are identified to address them and they are trialled and refined. Researchers and clinical staff work together with a common agenda. There is much to learn in this for improving

09012-CrimJournal-Guts.indd 189 12/21/09 7:14:29 PM

POP (see ihi.org/IHI/, which attempts to do for health largely what popcenter.org attempts for crime and disorder). A shared commitment to improvement, often involving the research community, seems to comprise an effective way of delivering routine and systematic attention to problems and performance weaknesses and to have driven collective efforts to provide better patient care and better clinical outcomes.

The Delivery of Specific Measures

The implementation of specific initiatives and their measures has proven difficult in many contexts. The issues here are not confined to interventions that are made as part of POP. Within POP, however, one factor relevant to the choice of measures has to do with their practicability and one aspect of delivering POP may relate to encouragement of their implementation. It is rarely the case that sustainable solutions to the recurrent problems addressed in POP are in the hands of the police alone, and they are also often not in the hands of members of the formal partnerships that operate in the United Kingdom either. Given this circumstance, the application of leverage on third parties has been deemed part of POP. Indeed, one of the Community-Oriented Policing Services (COPS) guides relates specifically to this (Scott and Goldstein, 2005). Even when measures have been agreed, in practice they may not be applied properly—and yet another COPS guide relates to this issue (Brown and Scott, 2007).

There is much devil in the detail in the design of interventions if they are to be agreed on and put in place as intended. As the COPS guides indicate, close attention to implementation issues has to be part of the delivery of POP if interventions are to be effective. Implementation is, however, largely missing from the SARA process. It is, at most, buried within "response", which relates mainly to measures and their expected efficacy rather than processes for their installation and operation. Issues of implementation may seem rather dull, but systematic attention to them is evidently needed if POP is to maximize its potential outcome benefits.

It is worth noting again the parallels with health. It is widely, if not universally, recognized amongst health professionals that hand hygiene is crucial to the prevention of much unintended hospital and clinician-induced infection. Yet it is also recognized that there is a great deal of hospital-induced infection, much of which is attributable to routine lapses in hand hygiene, especially by more senior physicians (Pittet, Simon, Hugennot, Pessoa-Silva, Sauvan, and Perneger, 2004). Part of the improvement/problem-solving in hospitals has related precisely to hand hygiene failings. The improvement orientation has led to initiatives that have attempted to understand why hands remain unwashed and to devising measures that will improve rates with the expected knock-on effect that infections will then fall. Sometimes quite simple measures, such as the relocation of washing facilities had a substantial effect on rates of hand-washing compliance.[2]

Where known sources of crime problems are likewise understood well (or the solution is obvious), the problem for POP might relate more to the implementation of measures than to their selection. Issues of intervention implementation are of importance, thus, not only because

2. I owe this example to Dr. Kath Kirkland, from Dartmouth College. I am most grateful to her.

09012-CrimJournal-Guts.indd 190 12/21/09 7:14:29 PM

of their implications for the conduct of evaluations, but also because they are crucial to the very design of responses to problems. They should properly be considered part of POP.

Evaluation

Evaluation has been notoriously weak within POP (Bullock, Erol, and Tilley, 2006; Knutsson, 2009) as it has been within crime prevention more generally (Ekblom and Pease, 1995; Tilley, 2000). Internal, agency-based evaluations have tended to be poor and to be more concerned with searching for any available signs of success than with a dispassionate test of the efficacy of what has been put in place. Weisburd et al. (2010) were disappointed at the small number of evaluations that meet their methodological vetting criteria. What kind of practicable evaluation is needed?

The evaluation methods favored by the Campbell Collaboration are poor at discriminating what it is about programs that work (where they do), what explains failure (when this is found), and what conditions are needed for successes and failures. Because of this, alternative methods have been sought and it is these that are likely to prove more useful in informing future practices. These alternatives relate both to the evaluation of individual initiatives and to methods of synthesizing evidence to inform policy and practice.

The COPS guides provide examples of alternative synthesis methods in practice. Pawson (2006) described a realist method, which eschews the traditional "what works?" question in favor of, "What works for whom in what circumstances, and how?" Pawson's approach draws on a much wider range of evaluative and other evidence for synthesis in the interests of policy and practice than that made use of in Campbell reviews, where most fall by the wayside. With regard to individual evaluations, Pawson and Tilley (1997) highlighted the limitations of the Campbell Collaboration preferred experimental and quasi-experimental evaluation methods and suggested realist alternatives. Using arguments that are congruent with those of Pawson and Tilley, Eck (2002, 2005) stressed the need for external validity in the design of evaluations of POP initiatives. This is not the chief concern of experimental and quasi-experimental methods, which emphasize internal validity. Eck and Madenson (2009) stressed, in particular, the need for detailed and theoretically derived outcome-pattern signatures to test the ideas that go into the kinds of interventions selected in problem solving.

The real threats to internal validity that experimental and quasi-experimental methods attempt to address need to be recognized. But it also needs to be acknowledged that the experimental and quasi-experimental methods are inadequate on their own for deriving lessons for problem solving and intervention selection in specific situations. Different methods serve different needs (see Tilley, 2006). The "realist" requirement that the outcomes accord with what is theoretically required makes heavy demands for a program to be deemed successful. Tilley and Hopkins (1998), for example, found that an initiative centered on reducing crime against business by focusing on repeat victimization seemed to succeed when assessed in the kinds of terms preferred by Campbell, but it failed in realist terms. It was not reduction in repeats that drove the fall in incidence. It was not clear what had done so. Later researchers suggested that

09012-CrimJournal-Guts.indd 191 12/21/09 7:14:29 PM

what might have been produced was an "anticipatory benefit" brought about by heavy publicity when Princess Anne, the Princess Royal, launched the project prior to measures being put in place (Smith, Clarke, and Pease, 2002). Without taking account of this as the (possible) active ingredient, it might otherwise have wrongly been presumed that it was the measures actually applied within the project that generated the crime reduction.

In addition to helping with evaluation, one of the ways collaborating criminologists working with POP practitioners might help is with articulating in advance the working theory behind the intervention. The dual benefits of surfacing the theory are (a) that it would open it to critical scrutiny in advance of implementing the chosen measure and (b) that it would clarify what could realistically be expected from it (its expected "signature") and hence what needs to be tracked to see if it is being effective.

If POP is to contribute to cumulative knowledge about how to deal with problems, then researcher involvement and evaluation cannot be confined to the conduct of technically strong experimental and quasi-experimental evaluations of the sort preferred by the Campbell Collaboration. They require more theory and more subtlety. The Boston Gun Project and subsequent theory and practice developments are, perhaps, the best examples we have so far (Kennedy, 2008; Kennedy, Piehl, and Braga, 1996; Kennedy et al., 2001). At its best this is what is done in medical research also, where explication of theory and mechanism precedes clinical trial and evaluation, and where methods of evaluation are increasingly going beyond the RCT, although without dispensing it altogether (see Berwick, 2008).

Conclusion

I was pleasantly surprised at the positive findings from the top ten studies included in the Campbell review of POP. I think it clinches the potential, and when that is done more of the same types of study may no longer be needed. I would not, however, have been persuaded that POP was a lost cause had the findings been negative. The reason is that POP describes the application of scientific methods and science rarely comes up with quick fixes. Most initial efforts fail!

POP goes back 30 years. There has been some progress in its application, but it could be better. Much is already being done in efforts to help it improve. The suggestions made in this short essay comprise modest additions.

Most importantly, although the police and their partners may be able to deliver solutions to relatively straightforward problems by simply attending to them, the research community has a substantial and interesting role to play in the problem-solving process where large-scale, recurrent problems that have resisted traditional treatments are being addressed. This role includes the development and articulation of theory as well as the application of well-rehearsed conventional research techniques to assess the effectiveness of the measures applied. In the latter case, conventional "gold standard" methods might not always be fit for purpose. Just as traditional, thoughtlessly applied (albeit technically challenging) police techniques might not be relevant or helpful for many problems, so too thoughtlessly applied (albeit technically challenging) evaluation techniques might not be relevant or helpful for the assessment of many solutions.

09012-CrimJournal-Guts.indd 192 12/21/09 7:14:30 PM

References

Berwick, Donald M. 2008. The science of improvement. *Journal of the American Medical Association*, 299: 1182–1184.

Brown, Rick and Michael S. Scott. 2007. *Implementing Responses to Problems*. Tool Guide 7. Washington, DC: U.S. Department of Justice, Office of Community-Oriented Policing Services.

Bullock, Karen. 2007. *Lost in Translation*. Unpublished Ph.D. Thesis. University College London.

Bullock, Karen and Nick Tilley. 2003. From strategy to action: The development and implementation of problem-oriented projects. In (Karen Bullock and Nick Tilley, eds.), *Crime Reduction and Problem-Oriented Policing*. Cullompton, Devon: Willan.

Bullock, Karen and Nick Tilley. 2003. *Crime Reduction and Problem-Oriented Policing*. Cullompton, Devon: Willan.

Bullock, Karen, Rosie Erol, and Nick Tilley. 2006. *Problem-Oriented Policing and Partnerships: Implementing an Evidence-Based Approach to Crime Reduction*. Cullompton, Devon: Willan.

Clarke, Ronald V. and John E. Eck. 2003. *Become a Problem-Solving Crime Analyst: In 55 Small Steps*. London: Jill Dando Institute of Crime Science.

Clarke, Ronald V. and Herman Goldstein. 2003. *Reducing Thefts at Construction Sites: Lessons from a Problem-Oriented Project*. Washington, DC: U.S. Department of Justice, Office of Community-Oriented Policing Services. Retrieved November 29, 2009 from popcenter.org/Library/RecommendedReadings/ConstructionTheft.pdf.

Eck, John E. 2002. Learning from experience in problem-oriented policing and situational crime prevention: The positive functions of weak evaluations and the negative functions of strong ones. In (Nick Tilley, ed.), *Crime Prevention Studies, Volume 14: Evaluation for Crime Prevention*. Monsey, NY: Criminal Justice Press.

Eck, John E. 2005. Evaluation for lesson-learning. In (Nick Tilley, ed.) *Handbook of crime Prevention and Community Safety*. Cullompton, Devon: Willan.

Eck, John E. and Tamara Madensen. 2009. Using signatures of opportunity structures to examine mechanisms in crime prevention evaluations. In (Johannes Knutsson and Nick Tilley, eds.), *Crime Prevention Studies*, *Volume 24: Evaluating Crime Reduction Initiatives*. Monsey, NY: Criminal Justice Press.

Eck, John E. and William Spelman. 1987. *Problem-Solving: Problem-Oriented Policing in Newport News*. Washington, DC: Police Executive Research Forum.

Ekblom, Paul. 2005. The 5Is framework: Sharing good practice in crime prevention. In (Erich Marks, Anja Meyer, and Ruth Linssen, eds.), *Quality in Crime Prevention*. Hanover, MD: LPR.

Ekblom, Paul and Ken Pease. 1995. Evaluating crime prevention. In (Michael Tonry and David Farrington, eds.), *Crime and Justice: A Review of Research, Volume 19—Building a Safer Society*. Chicago: University of Chicago Press.

Forrester, D., Mike Chatterton, and Ken Pease, with the assistance of Robin Brown. 1988. *The Kirkholt Burglary Prevention Project, Rochdale*. Crime Prevention Unit Paper 13. London: Home Office.

09012-CrimJournal-Guts.indd 193 12/21/09 7:14:30 PM

Forrester, David, Samantha Frenz, Martin O'Connell, and Ken Pease. 1990. *The Kirkholt Burglary Prevention Project, Phase II*. Crime Prevention Unit Paper 23. London: Home Office.

Goldstein, Herman. 1979. Improving policing: A problem-oriented approach. *Crime and Delinquency*, 25: 236–258.

Goldstein, Herman. 1990. *Problem-Oriented Policing*. New York: McGraw-Hill.

Hanmer, Jalna. 2003. Mainstreaming solutions to major problems: Reducing repeat domestic violence. In (Karen Bullock and Nick Tilley, eds.), *Crime Reduction and Problem-Oriented Policing*. Cullompton, Devon: Willan.

Home Office. 1984. *Crime Prevention*. Home Office Circular 8/84. London: Author.

Homel, Peter, Sandra Nutley, Barry Webb, and Nick Tilley. 2004. *Investing to Deliver: Reviewing the Implementation of the U.K. Crime Reduction Programme*. Home Office Research Study 281. London: Home Office.

Hope, Tim and Dan Murphy. 1983. Problems of implementing crime prevention. *The Howard Journal*, 23: 38–50.

Hume, David. 1739, 1999. *An Enquiry Concerning Human Understanding*. Oxford: Oxford University Press.

Kennedy, David M. 2008. *Deterrence and Crime Prevention*. London: Routledge.

Kennedy, David M., Anthony A. Braga, Anne M. Piehl, and Elin J.Waring. 2001. *Reducing Gun Violence: The Boston Gun Project's Operation Ceasefire*. Washington, DC: National Institute of Justice, U.S. Department of Justice, Office of Justice Programs.

Kennedy, David M., Anne M. Piehl, and Anthony A. Braga. 1996. Youth violence in Boston: Gun markets, serious youth violence, and a use-reduction strategy. *Law and Contemporary Problems*, 59: 147–196.

Knutsson, Johannes. 2003. *Crime Prevention Studies, Volume 15: Problem-Oriented Policing—From Innovation to Mainstream*. Monsey, NY: Criminal Justice Press.

Knutsson, Johannes. 2009. Standards of evaluations in problem-oriented policing projects: Good enough? In (Johannes Knutsson and Nick Tilley, eds.), *Crime Prevention Studies, Volume 24: Evaluating Crime Reduction Initiatives*.. Monsey, NY: Criminal Justice Press.

Knuttson Johannes and Ronald V. Clarke. 2006. *Crime Prevention Studies, Volume 20: Putting Theory to Work—Implementing Situational Prevention and Problem-Oriented Policing*. Monsey, NY: Criminal Justice Press.

Laycock, Gloria. 2004. New challenges for law enforcement. *European Journal on Criminal Policy and Research,* 10: 39–53.

Laycock, Gloria and Nick Tilley. 1995. Implementing crime prevention. In (Michael Tonry and David Farrington, eds.), *Crime and Justice: A Review of Research, Volume 19: Building a Safer Society*. Chicago: University of Chicago Press.

Leigh, Adrian, Tim Read, and Nick Tilley. 1998. *Britpop II: Problem-Oriented Policing in Practice*. Crime Prevention and Detection Series Paper 93. London: Home Office.

Pawson, Ray. 2006. *Evidence-Based Policy*. London: Sage.

Pawson, Ray and Nick Tilley. 1997. *Realistic Evaluation*. London: Sage.

09012-CrimJournal-Guts.indd 194 12/21/09 7:14:30 PM

Pittet, Didier, Anne Simon, Stéphane Hugennot, Carmen Lúcia Pessoa-Silva, Valérie Sauvan, and Thomas V. Perneger. 2004. Hand hygiene among physicians: Performance, beliefs, and perceptions. *Annals of Internal Medicine*, 141: 1–9.

Quinton, Paul and Julia Morris. 2008. *Neighbourhood Policing: The Impact of Piloting and Early National Implementation*. Home Office On-Line Report 01/08. London: Home Office.

Read, Tim and Nick Tilley. 2000. *Not Rocket Science: Problem-Solving and Crime Reduction*. Crime Reduction Research Series 6. London: Home Office.

Scott, Michael S. 2000. *Problem-Oriented Policing: Reflections of the First Twenty Years.* Washington, DC: U.S. Department of Justice, Office of Community-Oriented Policing Services.

Scott, Michael S. and Herman Goldstein. 2005. *Shifting and Sharing Responsibility for Public Safety Problems*. Problem-Oriented Guides for Police Response Guide Series No. 3. Washington, DC: U.S. Department of Justice, Office of Community-Oriented Policing Services.

Smith, Martha J., Ronald V. Clarke, and Ken Pease. 2002. Anticipatory benefits in crime prevention. In (Nick Tilley, ed.), *Crime Prevention Studies*, *Volume 13: Analysis for Crime Prevention*. Monsey, NY: Criminal Justice Press.

Tilley, Nick. 2000. The evaluation jungle. In (Scott Ballintyne, Ken Pease, and Vic McLaren, eds.), *Secure Foundations: Key Issues in Crime Prevention, Crime Reduction, and Community Safety*. London: IPPR.

Tilley, Nick. 2006. Knowing and doing: Guidance and good practice in crime prevention. In (Johannes Knuttson and Ronald V. Clarke, eds.), *Crime Prevention Studies, Volume 20: Putting Theory to Work—Implementing Situational Prevention and Problem-Oriented Policing*. Monsey, NY: Criminal Justice Press.

Tilley, Nick and Matt Hopkins. 1998. *Business as Usual: An Evaluation of the SBCI Programme*. Crime Prevention and Detection Series Paper 95, London: Home Office.

Tuffin, Rachel, Julia Morris, and Alexis Poole. 2006. *An Evaluation of the National Reassurance Policing Programme*. Home Office Research Study 296. London: Home Office.

Weisburd, David, Cody W. Telep, Joshua C. Hinkle, and John Eck. 2010. Is problem-oriented policing effective in reducing crime and disorder? Findings from a Campbell systematic review. *Criminology & Public Policy*. This issue.

Nick Tilley, OBE, AcSS, belongs to the Department of Security and Crime Science at University College London. He is interested in theoretically informed applied social science, especially as this relates to the prevention of crime and disorder. He has worked closely with the Home Office and many British police services in attempting to develop effective and ethical ways of reducing crime. He has published a dozen books and more than 150 reports, chapters and journal articles, mostly relating to policing, crime reduction and program evaluation methodology. His most recent books include *Crime Prevention* (Willan, 2009) and *Evaluating Crime Reduction Initiatives* (edited with Johannes Knutsson, Criminal Justice Press, 2009).

09012-CrimJournal-Guts.indd 195 12/21/09 7:14:30 PM