

Policy is in the details

Using external validity to help policy makers

John E. Eck

University of Cincinnati

Criminologists lament that their voices are seldom heard in the rooms of policy. Although all interest groups suffer similar distress, we criminologists do have a point. Much criminal justice policy is what some of my colleagues call “quackery” (Latessa, Cullen, and Gendreau, 2002); it is devoid of not only evidence but also of logic. Some approaches to crime and justice receive little policy support, despite their supporting evidence. And as important, some evidence-based policies are misinterpreted, misapplied, and misused.

Criminologists may not be able to prevent dumb ideas from being applied, and they may have difficulty getting policy makers to adopt evidence-based practices. Still, criminologists do have some ability to reduce the misinterpretation, misapplication, and misuse of their evidence. They can be clear about how policy works. A lack of clarity muddles and obfuscates criminological policy discourse.

Clarity about how a policy works is every bit as important as clarity about statistical methods, research designs, and measures of constructs. Science demands that studies be described in sufficient detail that they can be replicated. A study that cannot be replicated becomes the subject of grave suspicion, as is exemplified by cold fusion (Taubes, 1993). Similarly, criminologists who desire to influence policy must describe the policies they study with sufficient clarity that a policy maker could implement it, should that the policy be desirable, or avoid it if the policy is undesirable. Clarity is less a matter of style than of providing the right information. What is this information? This is the subject of this essay.

I argue that policy relevance is related to the external validity. Not external validity as viewed by Berk and MacDonald (2010, this issue), but as viewed by the people who may consider applying a study’s findings. Criteria for external validity (Shadish, Cook, and Campbell, 2002) provide a useful list of information authors should provide policy makers. I apply these criteria to the study by Berk and MacDonald to illustrate how information about these criteria can aid

Direct correspondence to John E. Eck University of Cincinnati, School of Criminal Justice, 600 Dyer Hall Cincinnati, OH 45221-0389 (e-mail: John.eck@uc.edu).

policy makers and why the absence of such information hinders interpretation, application, and use.

External Validity and Practical Utility

What makes a piece of research useful to a criminal justice policy maker? External validity provides a useful way to answer this question. External validity concerns how generalizable a study's findings are to other times, places, and circumstances (Shadish et al., 2002). Researchers pay homage to external validity in principle but less so in practice. The reason is that a single study can do little to address it. Whereas internal validity (i.e., the degree to which the causal conclusions of a study are justified by the research design) is a matter of falsifying alternative explanations of something that occurred in the past, external validity is about providing evidence for some hypothetical future outcome. The best a researcher can do with a single study is to show that there are no obvious insurmountable barriers to success in the future. Nevertheless, from the policy maker's perspective, external validity criteria provide a handy way of assessing the use of a particular study.

Expanding on Cronbach (1982), Shadish et al. (2002) developed five dimensions of external validity: units, treatments, outcomes, settings, and mechanism. Units refer to the cases examined in the research: Will the study's results change if the units are altered? Treatments refer to the operationalization of the intervention: Will the study's results change if the intervention is implemented differently? Outcomes describe how the dependent variable is operationalized: Will the study's results change if these measures are taken differently? Settings describe the context within which the treatment is applied: Will the study's results change when applied in a different jurisdiction by a different agency? Mechanism refers to the process by which the treatment produces the results within the setting: Will the study's results change if the treatment triggers a different process in a new setting? To make this clearer, let us apply these dimensions to Berk and MacDonald's (2010) study.

Learning from Policing Homeless

Berk and MacDonald (2010) provide an excellent case study for examining clarity. Their article is clearly relevant. Homeless encampments are not a trivial problem in U.S. cities. Police, public health, and other public agencies need to grapple with what to do about them (Chamard, 2010). As important, the policy examined seems to be practical—regardless of one's feelings about the desirability of removing homeless encampments, there is no dispute that this can be done.

Berk and MacDonald's (2010) study has a reasonable degree of internal validity. Their evaluation design has eliminated most alternative explanations for the crime reduction after the pilot and full treatments. However, although there is still the possibility of some unknown confounding variable responsible for the crime drops, the uncertainty over program efficacy is probably far less than the maximum threshold tolerable by most policy makers.

Units of Homelessness

To what units should the policy maker apply this policy? The introduction suggests places, but the study was carried out in a police district. “Place” is a theoretical term embedded in routine activity theory. It is a geographical entity far smaller than a police district or neighborhood.

Sherman, Gartin, and Buerger (1989) defined a place as “a fixed physical environment that can be seen completely and simultaneously, at least on its surface, by one’s naked eye” (p. 31). Eck and Weisburd (1995) state, “[p]laces . . . are specific locations with the larger social environment. They can be as small as the area immediately next to an automatic teller machine or as large as a block face, a strip shopping center, or an apartment building. Often places are thought of as addresses, specific types of businesses, or block faces” (p. 3). A homeless encampment is a place. Central Division, in Los Angeles, is a police command area, and it is definitely not a place.

Why is this important? Places have three basic roles in crime. First, from a routine activity theory perspective, they are where offenders physically come into contact with targets (Felson and Boba, 2010). Much can be done to prevent such meetings or to prevent such meetings from turning into crime (Eck, 2002). Second, places can be convergent settings (Felson, 2003), which are meeting places for offenders that facilitate co-offending there or elsewhere. And finally, places can be corrupting spots that generate crime at other locations. For example, a metal scrap recycling yard, by serving as a receiver of stolen metal, helps produce copper theft from other places. In all three roles, places are contact sites—not metaphorical contact, but real physical contact. A homeless encampment can take on any of these three roles. A neighborhood or police district can take on none.

Were the homeless encampments the treatment units or was the division? Central Division was unlikely to have been one large homeless encampment. And it is not likely that encampments were dispersed evenly about the division. Therefore, the police probably did not treat the entire division. They probably treated the camps. In contrast, one could argue that a few moderate-to-large homeless camps may have crime implications for the wider area within which they are embedded. The diffusion of crime from the camps was unlikely to have been uniform throughout the division, however. And it is not likely that encampment connected crimes ended abruptly at the boundaries of the division.

What we have is the possibility that the police treated places, but the evaluation was conducted at an area level. A real place-based design would have been preferable. Treated and untreated encampments could have been observed, and the crime in their buffer areas could have been counted.

The authors are rightfully concerned about the use of law enforcement to deal with problems created by homelessness. But their evaluation design gives the possibly misleading impression that entire neighborhoods need to be treated. Policy makers who take a neighborhood approach may not only increase the monetary costs of their interventions but also are likely to increase the social costs of their actions. Precise place-focused interventions might be able to reduce the unintended negative consequences.

Treatment of Homeless Encampments

We do not know where the police acted. Do we know more about what they did? What does removing a homeless encampment entail? How many encampments were addressed? Complicating matters, Berk and McDonald (2010) also state that other forms of enforcement were used outside the encampments, such as crackdowns on drunkenness, prostitution, and drug use; foot patrols against incivilities; undercover actions against drug dealers; an anti-robbery initiative; and a general increase in police visibility. Were these the active ingredients in the treatment? If so, this intervention is not simply one of removing homeless encampments, and the treatment would have been at the area rather than the place level. In fact, homeless encampment removal may not be the actions that drove down crime; it may be the other enforcement activities that accompanied the removals. If this is true, then using these enforcement tactics and leaving the encampments in place would have been as effective. And removing homeless encampments without using these other enforcement activities would have little or no impact on crime.

The lack of treatment specificity makes it difficult for scrupulous policy makers to determine what they should do. If policy makers simply removes homeless encampments, then they do not know how many to remove or whether this strategy is sufficient. If they implement law enforcement crackdowns, then they do not know whether this will work in the absence of camp removal. And if they do both, then they may be wasting resources and making their operations excessively harsh.

Outcomes of Homeless Encampment Treatment

Combining a heterogeneous set of crimes into broad categories such as violent, property, and disorder helps assure that there are sufficient events for statistical tests. However, it is likely that each category is dominated by a few types—for example, assaults may dominate the violence category, thefts from vehicles and assorted larcenies probably are most of the property offenses, and drug possession or prostitution might be the dominant disorder event.

Not knowing what these dominant categories are puts the policy maker in the position of not knowing precisely what types of crime reduction she is buying. A policy maker considering removing homeless encampments to reduce burglaries, or rapes, or public urination might think again if he knew that the Los Angeles experience does not extend to these offenses.

Settings of Homeless Encampment Treatment

If policy makers adopt the policy described in the research, they will implement it in a setting different from that in which it was tested. Although the researcher cannot know who will implement the policy where, the researcher can describe the setting with sufficient detail that the policy maker can make such a comparison. Because units are embedded in settings, lack of clarity about units will reduce setting specificity. If the unit is the police district, then the setting includes the surrounding districts. If the unit is the homeless encampment place, then the setting is the police district.

The setting matters because we cannot rule out the possibility that the treatment interacts with different settings differently, producing different outcomes. Policy makers seem to understand this intuitively when they state, “Your treatment may have worked in your city, but it is unlikely to work here.” The technical term for this is “context sensitivity.” One of the major reasons for replication is to learn how sensitive an intervention is to context. But to replicate, we have to be clear about both units and setting.

Here are some unknown facts about the district setting (assuming units are encampments) that might influence outcomes: the spatial clustering of encampments, the size distribution of encampments within the district, and the proximity of encampments to crime and disorder hot spots. The treatment may work well against spatially concentrated encampments but poorly when they are spread over a large area (or vice versa). If a few big encampments predominate, then taking these out may be useful, but if the encampments are many and small, then the intervention might be more difficult to implement. If the encampments are close to the sites of crime, then the causal link between encampments and crime is more obvious, so we would expect the treatment to work better than if encampments are a considerable distance from crime hot spots.

The setting also influences how we can interpret claims about displacement and diffusion of benefits. The standard practice for determining whether displacement or diffusion has occurred is to create a buffer around the treatment unit and handle it like a secondary treatment site. A separate control site is selected so that it is unlikely to be contaminated by the treatment. Because the context of a treatment interacts with that treatment, a setting is inappropriate for a control (Guerette and Bowers, 2009). The use of adjacent districts to Central as controls violates this rule. If offenders displaced to peripheral districts from Central, then the surrounding district crime trends would have been influenced by the treatment (by the same logic, diffusion of benefits would also contaminate these pseudo controls). In short, a policy maker cannot assume the displacement and diffusion findings are correct.

Mechanisms from Treating Homeless Encampments in Their Setting

The final clarity issue involves how the treatment brings about the outcome. It is standard in the hard sciences to show that there is a clear path from a manipulation to the outcome and not just a correlation. Thagard (1999), for example, shows that medical scientists ignored random controlled experiments that indicated that bacteria created stomach ulcers, until researchers demonstrated that a species of bacteria of the type that could cause the ulcers could survive and flourish in a human stomach.

What was the mechanism by which homeless encampment closures influenced crime? There are at least four, each with its own relevance to policy makers: victim removal, offender removal, place control, and side effect.

Victim Removal. The authors imply that removing the camps removed the possible victims. If the homeless were the primary victims of the crimes and disorders measured, then we would have to assume that they reported their victimizations to the police routinely. Although possible,

there are several arguments against this mechanism: The lack of access to phone and lack of trust in police are the two most obvious arguments. In fact, if the homeless were the victims, then closing encampments might suppressed reporting—reducing crime reporting through homeless advocacy groups and reducing (former) residents' trust in police. If a policy maker is interested in reducing harms to the homeless, then this is a serious concern.

Offender removal. If a significant portion of the offenses in Central Division were perpetrated against nonhomeless victims by homeless people living in encampments, then a different mechanism was triggered: the disruption of homeless offending patterns. How, exactly, is unclear. Perhaps these homeless offenders scattered to offend elsewhere. Perhaps, the encampments were convergent settings (Felson, 2003) that allowed offending homeless to meet and coordinate criminal activities. Although we do not know if this were true, this explanation does not have to contend with the crime reporting problem of the first explanation. This mechanism might appeal to policy makers with a disposition toward encampment removal and might be abhorrent to policy makers who oppose such interventions.

Place control. Eliminating the camps could be a form of place management (Madensen and Eck, 2008). Eliminating these places would reduce the chances of vulnerable victims meeting willing offenders in an uncontrolled setting. This explanation runs afoul of the same criticism as the victim mechanism, but we cannot rule it out. If this is the mechanism, then a variety of encampment control mechanisms might be possible for policy makers: ranging from closing them entirely and making camping sites untenable to introducing regulatory processes within camps.

Side effect. Finally, it is possible that the police activity in Central associated with removing the encampments (but not the removal themselves) disrupted criminal activity by the nonhomeless. This is plausible if the encampments were near crime hotspots (even if the homeless were not involved in the hot spots). If true, then the same or greater crime reduction might have been achieved by focusing on crime hot spots and on prolific offenders and ignoring encampments. And, if true, the elimination encampments in other settings may not get the crime reduction results observed in this study.

Improving Policy Relevance

This case study illustrates that external validity can provide a useful way to describe policy relevance. We should put this lesson to use. Just as we expect authors to describe their data, designs, and measures in some detail so we can assess internal, construct, and statistical validity (Shadish et al., 2002), we should also expect authors writing on policy to describe the dimensions of external validity. Policy makers may not understand the importance of the statistical methods, research design, or even measurement, but they probably can understand the relevance of units, treatment, outcomes, setting, and mechanism. We should ask the following questions “Has the study described these dimensions in sufficient detail and clarity that a policy maker could faithfully implement this policy?” And if these elements cannot be described with clarity,

then we should ask the following question: “Has the study clearly stated that there is a lack of relevant information about these dimensions?”

I have made the claim that policy relevance goes beyond competent research. Policy relevance is not simply about “what works.” Policy relevance is also about “*what* works *where*, *when*, and with *whom*.” And policy relevance also concerns “*how* does it work under different circumstances.” If criminologists are to be useful, then they must be clear about these basics. And if the circumstances surrounding the study prevent the researchers from providing these basics, then the researchers should be clear about what they do not know. Only then can we claim that we have done our best to prevent our results from being misinterpreted, misapplied, or misused.

References

- Berk, Richard and John MacDonald. 2010. Policing the homeless: An evaluation of efforts to reduce homeless-related crime. *Criminology & Public Policy*. This issue.
- Chamard, Sharon. 2010. *Homeless Encampments. Problem-Oriented Guides for Police. Problem-Specific Guides Series, no. 56*. Washington, DC: U.S. Department of Justice, Office of Community Oriented Policing Services.
- Cronbach, Lee Joseph. 1982. *Designing Evaluations of Educational and Social Programs*. San Francisco, CA: Jossey-Bass.
- Eck, John E. 2002. Preventing crime at places. In (Lawrence W. Sherman, David Farrington, Brandon Welsh, and Doris Layton MacKenzie, eds.), *Evidence-Based Crime Prevention*. New York: Routledge.
- Eck, John E. and David Weisburd. 1995. Crime places in crime theory. In (John E. Eck and David Weisburd, eds.), *Crime and Place*. Crime Prevention Studies, Vol. 4. Monsey, NY: Criminal Justice Press.
- Felson, Marcus. 2003. The process of cooffending. In (Martha J. Smith and Derek B. Cornish, eds.), *Theory for Practice in Situational Crime Prevention*. Crime Prevention Studies, Vol. 16. Monsey, NY: Criminal Justice Press.
- Felson, Marcus and Rachel Boba. 2010. *Crime and Everyday Life*, 4th Edition. Thousand Oaks, CA: Sage.
- Guerette, Rob T. and Kate J. Bowers. 2009. Assessing the extent of crime displacement and diffusion of benefits: A systematic review of situational crime prevention evaluations. *Criminology*, 47: 1331–1369.
- Latessa, Edward J., Francis T. Cullen, and Paul Gendreau. 2002. Beyond correctional quackery: Professionalism and the possibility of effective treatment. *Federal Probation*, 66: 43-49.
- Madensen, Tamara and John E. Eck. 2008. Violence in bars: Exploring the impact of place manager decision-making. *Crime Prevention & Community Safety*, 10: 111–125.
- Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for General Causal Inference*. New York: Houghton Mifflin.

Sherman, Lawrence W., Patrick R. Gartin, and Michael E. Buerger. 1989. Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, 27: 27–55.

Taubes, Gary. 1993. *Bad Science: The Short Life and Weird Times of Cold Fusion*. New York: Random House.

Thagard, Paul. 1999. *How Scientists Study Diseases*. Princeton, NJ: Princeton University Press.

John E. Eck is Professor of Criminal Justice at the University of Cincinnati's School of Criminal Justice. He has written extensively on police reform efforts, particularly Problem-Oriented Policing, and on the prevention of crime at places. He received his bachelors and masters degrees from the University of Michigan, and his doctorate in criminology from the University of Maryland.