EXPERIMENTAL CRIMINOLOGY AND THE FREE-RIDER DILEMMA*

Accepted for publication, *British Journal of Criminology* (2020).
Please do not cite or circulate this pre-proof without prior permission.

Authors

Johann Koehler[1] & Tobias Smith[2]

[1] Department of Social Policy, London School of Economics and Political Science
[2] Jurisprudence and Social Policy, University of California, Berkeley

Abstract

Experimental criminology promises a *public good*: when experiments generate findings about criminal justice interventions, everyone benefits from that knowledge. However, experimental criminology also produces a *free-rider problem*: when experiments test interventions on the units where problems concentrate, only the sample assumes the risk of backfire. This mismatch between who pays for criminological knowledge and who rides on it persists even after traditional critiques of experimental social science are addressed. We draw from medicine and economics to define experimental criminology's free-rider problem and expose a dilemma. Either we distribute the costs of producing policy-actionable knowledge to the entire beneficiary population. Or we justify isolating the risk of experimental harm on that class of the population where ethical concerns are most acute.

Keywords

Ethics; experimental criminology; free-rider; iatrogenesis; medical model; risk of harm

Word count

9,332

* For invaluable feedback at assorted stages in this article's production, we thank David Humphreys, Keith Humphreys, Mona Lynch, Rob MacCoun, Amanda Sheely, David Thacher, and annual meeting attendees of the Law & Society Association and the American Society of Criminology, and seminar series participants at UC Berkeley in Jurisprudence and Social Policy and at the LSE in International Social and Public Policy.

> The true dilemma … is the question not of *whether* to be unfair, but of *to whom* to be unfair.
> —Lawrence Sherman (2007: 318)

PROLOGUE

Experimental criminologists wish to evaluate whether interventions alleviate social ills. In designing evaluations, we face a dilemma: if we evaluate the intervention where the ill concentrates most acutely, we can optimise for statistical power, because experimental criminology teaches that treatment effects are easiest to perceive among the most acutely afflicted. Alternatively, if we design an evaluation that tests the intervention on the afflicted and the unafflicted alike, we can optimise for distributive justice, because ethics teaches that all beneficiaries should share the costs of an experiment equitably. The dilemma triggers an uncomfortable choice: we can prioritise method or equity, but *we cannot do both*.

Consider the city of Minneapolis. In the 1980s, within a single decade spanning Chief Bouza's tenure steering the Minneapolis Police Department, the city was transformed into a laboratory for rigorous policing experiments. In quick succession, Minneapolis witnessed the first policing experiment that randomly assigned mandatory arrest for domestic violence (Sherman & Berk 1984), another that randomly assigned hot spots to receive saturated police presence (Sherman & Weisburd 1995), another that randomly assigned a special unit to work intensively with half of the addresses responsible for generating the most calls for service (Buerger 1993), and another that assigned a neighbourhood watch officer to some city blocks but not others (Pate et al. 1987). Crime dropped in the first two experiments' treatment units; the last two reported no significant differences between either condition. But in *all four* studies — not just the ones that reported a crime drop — we learned something new about how to design an effective police service.

Within a few short years, Minneapolis had been refashioned into criminology's bridgehead for an Experimenting Society (Campbell 1969). Subsequent experiments conducted in police

departments across the United States revealed equivocal results for mandatory arrest of domestic violence and clear reductions in crime for hot spots policing (Sherman et al. 1992; Braga et al. 2019). Insights from the Minneapolitan experiments guided police practice for years to come. The practice of submitting justice interventions to rigorous testing during Bouza's tenure also stands for a proposition: concentrating rigorous experimental interventions on an acutely impacted sample is an effective and ethical way to provide a benefit of criminological knowledge to the world.

Imagine things had gone differently after Chief Bouza left Minneapolis. What if, from the late-'80s onward, criminologists insisted upon testing the effectiveness of criminal justice interventions *only* in Minneapolis? A methodologist would complain that testing an intervention's effects in Minneapolis alone presents an external validity problem. If the goal is to extrapolate from an intervention's effectiveness 'over there' in Minneapolis to its effectiveness 'over here,' then reliance on only those tests that were conducted 'over there' will not cut the mustard (Cartwright & Hardie 2012). Instead, sound justice policy depends on fostering a robust and ubiquitous culture of rigorous evaluation the world over.

But ethicists would also surely raise a cry about experimenting only in Minneapolis. After all, the equitable distribution of cost and benefit is a bedrock principle of distributive justice. People everywhere benefit from the laboratory of Minneapolis, because experimental knowledge is a public good. Yet not everyone bore the costs of the laboratory of Minneapolis. The people of Minneapolis bore those costs alone. They assumed the burden of testing the interventions that did not work. And they assumed the risk that every intervention — no matter how well conceived — could have done worse than nothing. Any of them could have backfired and done harm. And while the possible backfire varies in nature across each experiment (*e.g.*, hot spots policing could have

increased crime; mandatory arrest could have worsened domestic violence; *etc.*), the risk of backfire does not. No matter the intervention, the experimental sample bears the uniform cost of increased *risk*. Because of this uneven distribution of risk, experimenting only in Minneapolis would be unjust, even if it were methodologically sound.

A thought experiment in which we conduct experiments on the effectiveness of criminal justice interventions only in Minneapolis may sound fanciful. But, as we show in this article, it is in fact in a meaningful ethical sense what experimental criminology proposes we do. Experimental criminologists call for experimentation on units where effects will be easiest to observe. Those units may not be concentrated geographically (as in the example of Minneapolis's hot spots) but their concentration is nonetheless a persistent feature of crime's distribution that experimental criminologists propose we tap. Although the methodologist may have few qualms about repeated experimentation on those units, the ethicist should have many. Moreover, the ethicist's worry will grow to the extent we sample on those units because their vulnerability might be methodologically attractive.

The ethical concern we highlight in this article arises because of a free-rider problem. Free-riders appear wherever one class of people enjoys a good without bearing the burden of its production. It is a market inefficiency that arises from exploiting a benefit at no internalised cost. Minneapolis took on risks, and we learned something from them. Criminological experiments create precisely such a market inefficiency by isolating the risks of harm to those units of a population from whom the public good — knowledge about what works — is extracted. Free-riders present more than a problem of economic or political theory; in experimental criminology, they also pose a problem of research ethics. Free-riding is a matter of distributive injustice.

Experimental criminology has drawn its fair share of methodological and ethical critique (*e.g.*, *see* Baunach 1980; Berk 2005; Erez 1987; Hollin 2008; Nagin & Sampson 2019; Sampson et al. 2013). The settled wisdom, however, concludes that "folklores about experiments being 'unethical' or 'impractical' are (in most cases) just that" (Sampson 2010:490). Yet the concern we raise here is a new one that takes seriously the arguments that proponents of experimental criminology have themselves raised. Indeed, we find that the free-rider problem arises only once we recognise the public good that experimental criminology offers.

THE FREE-RIDER PROBLEM

Distributive justice refers to the branch of ethics that takes up the question of how to distribute goods and harms fairly in a society. We owe its perhaps most famous articulation to John Rawls, whose Difference Principle states that "social and economic inequalities … are to be of the greatest benefit to the least-advantaged members of society" (Rawls 2001: 42-3). The Difference Principle is featured in the *Belmont Report* (1979: B.3), which summarises ethical principles for human subjects research. The report notes three pillars of ethical practice: respect for persons, beneficence, and (implicitly distributive) justice. None of those principles is absolute. Indeed, the *Belmont Report* appreciates that flexibility and a balancing of factors best serves the interests at stake in human subjects research (Levine & Skedsvold 2008). Flexibility is important for research ethics in criminology, where ethical carve-outs, compromises, and side-constraints may be necessary for sound work to proceed (Armstrong et al. 2014; Carlen 2016). Surprisingly, the American Society of Criminology's recently-promulgated Code of Ethics (2017) — unlike that of the Academy of Criminal Justice Sciences (2000: III.B.16) and the British Society of Criminology (2015: 4.1-4.3) — foregoes altogether both of *Belmont*'s beneficence and justice pillars. The omission of distributive justice in the ASC's Ethics Code coincides with the formalisation of experimental criminology's research programme. We take it as uncontroversial that a research

programme that allows some to free-ride on the risk of harm that others assume creates an ethical problem, for criminology no less than for other disciplines (Cowburn et al. 2017; Graebsch 2000; Israel & Hay 2011; Reiter 2009). We engage a debate over these concerns — so far absent — in both the ASC Ethics Code and experimental criminology's research agenda.

Experimental criminology's free-rider problem is cause for urgent concern — but not alarm — for three reasons: First, the free-rider problem is not unique to experimental criminology. Rather, it is a problem endemic to research approaches that provide a public good through evaluations that internalise the costs of knowledge production and externalise the benefits. Experimental medicine is an example of another such approach. Medical researchers test new drugs and treatments in experiments before approving them for use in the general population. Bioethicists who are alert to the free-rider problem in medical research offer instructive paths forward for experimental criminology. We apply those lessons to experimental criminology for the first time here. Second, the free-rider problem is limited. Free-riders appear in only those research designs that manipulate the social world in ways that produce a risk of backfire (that is, experiments that can make things worse). Therefore, the free-rider problem attaches only to research designs that satisfy the manipulation criterion. Finally, as an economic matter, free-riding is just an inefficient distribution of benefit and cost. What elevates free-riding to a matter of urgent ethical concern is that experimental criminologists call for distributing those benefits and costs in a way that is not only inefficient; it is also *unjust*.

Although we discuss some ways to address the free-rider problem, we forewarn that no approach will please all readers. Dilemmas, by their nature, present an uncomfortable choice between two suboptimal alternatives. So, too, the dilemma we identify forces an unappealing choice between compromising our commitments to either method or equity.

We begin by explaining how the research process can produce a free-rider problem. We then discuss how medicine — to which criminology often looks as a model — grapples with the free-rider problem. Next, we show how experimental criminology has institutionalised principles that logically call for the production of knowledge for the general good through the selective application of risk of harm. This conclusion derives from premises contained in three canonical texts in experimental criminology: Joan McCord's *Cures that Harm* (2003), Lawrence Sherman's *The Power Few* (2007), and David Weisburd's *Moral Imperative for Randomized Trials* (2003). When taken together, these three articles articulate a research programme that isolates the risk of possible backfire effects on a sample of the population where normative concerns are at their most acute, without distributing the associated risks of experimentation throughout the population that benefits from this research. This research programme produces the free-rider problem: experimental criminology purchases knowledge about interventions that work, for the benefit of the many, with the coin of risk tithed from the at-risk few. In searching for a solution to the free-rider problem, a dilemma forces the choice between optimising either statistical power or distributive justice. We conclude by deducing why *we cannot do both*.

FREE-RIDERS AND RESEARCH ETHICS

Social scientists and policy researchers observe free-rider problems throughout the social world. Wikipedia users outnumber those who donate for the website's upkeep; neighbourhood-wide property values appreciate despite only selective investment; non-voters benefit from voters' behavior; out-of-state motorists ride locally-funded streets; ranchers let cattle graze on public land; underpaying NATO member-states enjoy the protections of multilateral defence treaties; un-unionised laborers benefit from the gains of collective bargaining; and punitive municipalities send a disproportionate share of prisoners to state-funded institutions. In each instance, free-riders

appear because the consumption of a public good misaligns with the costs of its provision (Cooter & Ulen 2012).

Social scientists and policy researchers can also *produce* free-rider problems through the research process. We test the effectiveness of pharmaceuticals in the Global South that we consume in the Global North; we test pedagogical techniques in the laboratories of inner-city charter schools and use them in advantaged schools in the suburbs; and we test for the point at which the un-vaccinated benefit from the herd immunity of the vaccinated. Useful policy knowledge is a public good, but not everyone shares in its production. The free-rider, so the argument follows, reaps the benefits of that useful policy knowledge without bearing the burdens of undergoing the research necessary to generate it.

Economists recognise two solutions to the free-rider problem, both of which are unsatisfactory in this context. The first solution is demand-side: free-riders could 'opt out' of enjoying the public good. For example, what if British patients were to refuse prescriptions of pharmaceuticals whose effectiveness had been trialled only in, say, sub-Saharan Africa? Or, what if British doctors decided not to prescribe those pharmaceuticals? Declining to use the policy knowledge that research has produced would certainly go some way to fixing the problem. But in most cases there is no practical way for a free-rider to 'opt out' of the benefit, nor can free-riders be prevented from enjoying it. This is the nature of public goods — they are nonexcludable. Thus, although some people may decline to bear the cost of participation in research, they cannot similarly forgo the benefit by withdrawing from a world that generates knowledge that may apply to them or similarly-situated others.

The second unsatisfactory solution is supply-side: we could do away with consent and mandate universal participation in research. Indeed, bioethicists have cautioned that the postwar

ethical emphasis on informed consent shortchanges distributive justice concerns. For example, in response to the deplorably low rate of medical studies that satisfy target enrolment — which, according to the Institute of Medicine (2010), has fallen as low as 10%, and is now a more serious concern than research funding shortfalls (Schaefer et al. 2009) — some bioethicists caution that research participation may in fact be an ethical necessity. In particular, Rhodes (2005; 2017) proposes a universal prima facie requirement to participate in biomedical research *even for those classes of the population not otherwise considered ill*. She identifies the preoccupation with informed consent as blinding bioethics to the distributive justice concern that everyone is vulnerable to the afflictions that medical research seeks to address. Shifting attention in this way rebalances how we adjudicate ethical concerns that may be in tension with one another. A distributive justice approach soft-pedals the need to procure 'buy-in' from research subjects, and instead draws our attention to how we all benefit from research on others. Thus, although medicine is ostensibly tasked with preventing and curing sickness, nonetheless medical research should involve participation by both the healthy and the sick alike. Because the well may become the ill, at which point they will benefit from medical knowledge, we must all engage in research. Medical experiments ought to be like taxes — a shared burden for a public good.

A presumption of participation in research may seem straightforward for those familiar with population-level vaccination and immunisation studies, which routinely require the participation of a healthy cohort; or longitudinal public health studies, which necessitate testing before the onset of ailments; or epidemiological studies, which follow health trends such as the composition of community-members' microbiomes (Rhodes 2017: 321-322). However, when we move from observational research designs to experimental ones, the introduction of a manipulation criterion renders the ethical concern much more acute. In particular, it is controversial to encourage *more*

participation for populations currently identified as vulnerable. It is, after all, only a prima facie requirement — and therefore one that may not prevail against a competing right not to participate. But that competing right would have to clear its own bar: for example, in a widely-discussed bioethics article, Rhodes (2005) advocates greater efforts to include groups such as children who may not themselves be legally eligible to give consent. A research participation requirement proposal might seem far-fetched and fanciful, but it cannot be easily dismissed.

Criminologists should care about a free-rider problem in medical research because of how criminology has modeled its epistemological and normative commitments. Standard codes of research ethics frown upon imposing risk on the few in exchange for a benefit that accrues to the entire distribution (Emanuel et al. 2000). The *Belmont Report*, for example, cautions that the benefits of research ought not to accrue disproportionately to a class that does not itself participate in the research, and it moreover counsels against selecting subjects "simply because … of their compromised position." (1979: B3) This directive retains a persistent clarity and force even amid institutional review boards' expanded remits, ambiguous articulation of ethical principles, and intensified enforcement of protections against putatively risky research (Feeley 2007; Haggerty 2004; Rothman 1991). Yet the American Society of Criminology's Code of Ethics is silent on this particular point — that silence does not, however, lessen the concern's moral thrust. When we reap the results of an experiment, we all prosper regardless of the direction of its sign, the magnitude of its effect, and the amplitude of its associated error bar.

In that light, we can extract a dilemma for experimental criminology from Rhodes's proposal. One option is for criminology to embrace the analogy to medicine and consider how the discipline might formulate a research programme that addresses the distributive justice problem from which free-riders emerge. The alternative is to distinguish between medicine and criminology so as to

unburden criminology from the ethical requirement to mandate research participation. As we will show below, the leading lights of experimental criminology have drawn explicit analogies to medicine, and in so doing have embedded a free-rider problem within the research process.

PREMISES OF THE CONCERN

In this section, we examine three canonical texts that both produce experimental criminology's free-rider problem and illustrate its importation from medical research. We select these three texts because they are widely recognised as central to experimental criminology's development, and each one has been authored by a past president of the Academy of Experimental Criminology. They therefore stand as authoritative expressions of experimental criminology's mainstream. Moreover, these pieces exemplify the high standard of intellectual rigour that experimental criminology promises. All three texts make claims about experimental criminology's ethical obligations, and taken together they stand for a set of claims about the field's normative and epistemological commitments. Each also establishes a premise in our argument: the premise to which we refer in *Cures that Harm* is McCord's empirical finding of iatrogenic risk; in *The Power Few*, it is Sherman's hypothesis regarding sampling strategy; in *The Moral Imperative*, it is Weisburd's ethical precept about research design. When we put these three pieces together, we can see how the free-rider problem is one of experimental criminology's own creation.

McCord's *Cures That Harm*

Joan McCord's *Cures that Harm: Unanticipated Outcomes of Crime Prevention Programs* (2003) was a watershed in experimental criminology. McCord summarised the research to date from the Cambridge-Somerville experiment, a half-century longitudinal study following the outcomes for at-risk youth assigned to receive an intensive therapeutic battery of social intervention and guidance. What scholars observed was surprising: the programme had "iatrogenic" effects — it backfired. Like five other interventions directed at at-risk groups that

McCord identified, participants who received the intervention ended up worse off than those in the control. In addition to the iatrogenesis observed among all the interventions McCord canvassed, the studies also shared other features: they were all designed with the best of intentions; they were all rigorously evaluated; and they were all applied to at-risk subjects.

Synthesising these studies led McCord to two conclusions: first, "[s]ocial programs deserve to be treated as serious attempts at intervention, with possibly toxic effects, so that a science of intervention can prosper." (McCord 2003: 17) The second insight followed from the first: a principal task for the evaluation researcher working on crime and justice is one of preventing Cures That Harm. Yet while *Cures That Harm* forcefully positioned the need to focus on whether an intervention backfires as a primary concern for criminological analysis, it left unstated a feature common to the studies McCord marshalled: the research participants' social position. In particular, McCord under-emphasised that on each occasion in which scholars tested interventions that later reported iatrogenic outcomes, the units selected for study were uniformly drawn from a deliberately 'at-risk' sample in a fashion that targeted those high-risk characteristics. As scholars have extended the argument put forward in *Cures That Harm*, they have taken to heart the need to design interventions that do not backfire (Gottfredson 2010; Linning & Eck 2018; Weisburd et al. 2011; Zane et al. 2016), but they have not taken up the concern that in criminological research the harm of that backfire has consistently fallen upon groups already designated as vulnerable — indeed, that those groups were selected for study precisely because they were vulnerable.

Far from taking caution on this front, subsequent criminological scholarship has hardened its resolve that at-risk units are the units of a population on which interventions *ought* to be tested. This resolve is encapsulated in the Power Few Hypothesis, which provides that we may be best positioned to observe effects if we test on units most likely to maximise statistical power, which

are precisely the populations about whom McCord was most concerned. The result is that the silent clause within McCord's admonition has become one of experimental criminology's organising principles.

Sherman's *Power Few*

Lawrence Sherman took *Cures That Harm* as a forceful point of departure in his 2006 Joan McCord Prize Lecture, *The Power Few: Experimental Criminology and the Reduction of Harm* (Sherman 2007: 299-300; *see also* 2019). Where McCord was primarily concerned with the ethical imperative to identify interventions that backfire, Sherman argued that the guiding ethical imperative for the criminologist was to identify interventions that work, because "no matter how much negative evidence we may produce as gatekeepers barring harmful programs, our best chance to reduce human misery is with positive evidence of programs that work well" (p.300). The recourse for the criminologist to know what works, Sherman hypothesised, might be to conduct experiments that leveraged the statistical effects of the 'power few.'

The thrust of Sherman's argument was that experiments in criminology are beset with a challenge that is endemic to all research in public policy: problem units are axiomatically concentrated in one tail of the distribution (Farrell 2015): well-rehearsed examples refer to crime's concentration at places (*e.g.*, Sherman et al. 1989; an observation that leads one observer to ratify a "law" of crime's concentration at place, Weisburd 2015), and among offenders (*e.g.*, Wolfgang et al. 1972), offending years (*e.g.*, Steffensmeier et al. 1989), and victims (*e.g.*, Pease 1998). Sherman showed that if we want to study whether an intervention works to alleviate a problem, we may either study its effect on units as though the problems they present are evenly distributed throughout the population, or we may instead recognise that problems concentrate in one tail of a skewed distribution. But, so the hypothesised argument follows, at least studying the effectiveness

of interventions on the power few isolates the research subjects where a strong treatment 'dose' can be applied to units presenting acute levels of the problem.

Sherman was attuned to the ethical arguments that experiments targeting the power few may raise. In fact, he adapted the term 'power few' from the popular science writer Malcolm Gladwell, whose reportage indicted the use of a power few policy to tackle an intractable homeless population in Denver (2006). Gladwell wrote that Denver spared considerable expense and trouble by concentrating intensive social services on a small sample of problem individuals who cost the city vast sums each per year in police, ambulance, and related municipal fees. Targeted preventative services for these individuals — including housing and a full-time caretaker — were cheaper than blanket municipal interventions applied irrespective of intensity of need. Gladwell acknowledged that the programme was cost-effective, but he wondered whether it was fair for a select minority to soak up a majority of the city's resources just because they made such a mess, while the quiet suffering of less publicly troublesome cases was implicitly ignored. Gladwell's point was to note the tension between distributive justice and utilitarian solutions. Or as he put it, "[power few] problems leave us with an unpleasant choice. We can be true to our principles or we can fix the problem. We cannot do both" (Gladwell, qtd. in Sherman 2007: 303).

Gladwell's moralism did not persuade Sherman, who concluded that "The true dilemma Gladwell points us to is the question not of *whether* to be unfair, but of *to whom* to be unfair" (p. 318, emphasis in original). Sherman insisted that criminology stands on consequentialism's side. Nonetheless, he believed that criminology may gain some moral purchase when its interventions applied to the power few provide services rather than punishments. He justified the intervention in Denver, for example, by pointing out that the Denver programme was a net good for its recipients, *and* that its knowledge and second-order effects trickled up to society at large. In this case, the

free-rider problem posed little concern because the Denver programme accrued benefits both for the target recipient and the population at large and the treatment left no-one worse off.

Yet that conclusion sidesteps a central feature of McCord's thesis from *Cures that Harm*. As in the Denver programme, all of McCord's six test cases were expected to provide a benefit to recipients, yet in all cases the recipients still endured a harm. Indeed, at the outset each experiment appeared to provide a positive service. Yet in each case, researchers observed worse outcomes among those in the treatment group than among those in the control group. In other words, whenever we test an intervention, no matter how beneficial on its face, we must be mindful that participants assume a risk in the intervention. It may backfire. And regardless of whether it does, researchers gain an insight, and subjects suffer a risk of possible harm.

Gladwell was right to be worried about distributive justice. However, if we take McCord's thesis seriously, Gladwell's ethical worry pointed in the wrong direction. Gladwell was concerned with the unequal distribution of benefits to the power few. But the participants in randomised controlled trials always assume a risk, no matter how benign the intervention. In Gladwell's case the risk yielded a benefit to the experimental sample, but McCord shows us that this need not always be the case. Taking on board McCord's observations and reading through Sherman's proposition that effects may be most readily observable through experiments testing subjects selected from the power few, then what we should be most concerned about is not unequal benefit, but unequal distribution of risk of harm among the power few.

Weisburd's *Moral Imperative*

In a 2003 article, *Ethical Practice and Evaluation of Interventions in Crime and Justice: The Moral Imperative for Randomized Trials*, David Weisburd claimed that experimental criminology rested upon a moral duty that "develops from our professional obligation to provide valid answers to questions about the effectiveness of treatments, practices, and programs." (Weisburd 2003: 350)

The argument is neat: sound criminal justice policy ought to rely on valid inferences about causal effects; randomised trials provide the most valid attributions of cause and effect; therefore, we ought — wherever possible and whenever the policy question is so susceptible — to prize randomised controlled trials.

The claim is compelling. Weisburd insists that the researcher has a moral duty to produce the most internally valid research possible so as to maximise benefit and minimise harm in the policy that follows. Policy can do good if based in sound evidence, or it can do harm if settled on methodologically faulty findings.

> From an ethical standpoint, isolating the effects of treatment is one of the evaluator's most important obligations to society. Stating that a certain treatment or protocol is effective when it is not will lead to significant societal costs, economic and social. Moreover, failing to recognize the harms of treatments or interventions can lead to much suffering on the part of the individuals receiving treatment, or communities that expect benefit not harms from them.
>
> (Weisburd 2003: 339)

Setting Weisburd's claim in dialogue with McCord's puts his claim in starker relief: given that interventions always carry the risk of harm, we had better contain that risk within the experiment, rather than visit it on everyone through the final policy. If the treatment group were randomly selected from the population, Weisburd's argument would raise no distributive justice concerns and thus no free-rider problem. But, Sherman's power few hypothesis closes the door on random selection from the population. Instead, we resort to a sample drawn from society's problem units.

The power few need not be society's most disadvantaged, but criminology rarely looks elsewhere. Instead, we find that criminology's preferred sample is typically the same narrow subset of society. Whether it is "rebellious youth from ghastly families" (as McCord describes the Somerville Study participants, pp. 18), "high-risk, previously convicted, recently released

offenders," or "hot spot" high-crime neighborhoods (two of the four examples of the power few that Sherman provided, pp. 310), or "high-priority victims" or "felonious few" in Sherman's (2019) reconsideration of the nomenclature, these samples consistently draw from among the dispossessed. It is not enough that these groups may, in fact, benefit from the treatment because, as McCord makes clear, these groups nonetheless assume a risk of a possible null result or even iatrogenesis. And it is precisely in assigning the risk of a null result or backfire to only these groups rather than society at large that Weisburd argues experimental criminology acts on an ethical imperative.

EXTERNALIZING THE INTERNALITY

Together, McCord's, Sherman's, and Weisburd's arguments coalesce as a set of interlocking propositions. Those propositions ground the ethics of experimental criminology in the public good that its findings produce. Rigorous experiments doubtless provide socially beneficial criminological knowledge. But the combination of the risk of backfire, sampling on the power few, and the containment of risk to the experimental sample produces conditions under which the general population rides for free.

Consider that we purchase the knowledge an experiment produces with the coin of a risk of possible harm (*i.e.*, backfire). Risk, here, refers narrowly to the uncertainty about whether a good or bad outcome, howsoever construed, will transpire from an experiment. Moreover, McCord shows that backfire harms are unpredictable with regard to frequency, form, and intensity. The experimentalist must acknowledge that observed effects are an outcome of a process that probabilistically produces harms as well as benefits. It is precisely because there is a risk of backfire that Weisburd thinks experimental research is so important. He grounds the moral imperative to experiment in its capacity to provide valid observations of an effect; that validity is useful insofar as it ensures that harmful interventions proceed no farther than the experiment alone.

This is a programme of iatrogenic containment: Weisburd is concerned with whether we can internalise the costs of experimentation while we also externalise the benefits.

If we follow Weisburd's reasoning, then the answer is always yes. The knowledge that experiments produce helps society as a whole. The thinking goes that the knowledge gained from having tested an innovation helps design a better justice system. For example, when an experiment reveals a more effective technique or a new safeguard against a harmful practice, the ambition is that knowledge is disseminated for use elsewhere. The previous section explained that knowledge about an intervention's effects is purchased with the coin of risk that an intervention may ultimately backfire. But this risk is not shared throughout the population. To the contrary, although information about interventions that 'work' is disseminated throughout the population, only the power few pay the price to acquire it.

Containing risk within an experiment for the benefit of the many could be distributively just if experiments reliably improved outcomes. But McCord, who stresses that backfire is both unpredictable and unavoidable, takes that option off the table. Experimenters must instead recognise that some cures harm. We might also wish that containing risk within an experiment for the benefit of the many could be distributively just if experimental subjects were drawn randomly from the population. But Sherman's hypothesis takes that off the table, too. Experimenters are instead encouraged to sample on the power few. Finally, we might also hope that containing risk within an experiment could be distributively just if the knowledge that an experiment produces were to apply exclusively to the experimental sample. For example, perhaps power few units are qualitatively distinct — as a matter of kind, not degree — from other, 'weak many' units in the distribution from which they are drawn. But Weisburd takes that option off the table as well.

Experimenters instead advance the moral imperative when we export insights from rigorous criminological tests to the design of interventions that we deliver to other groups.

We deduce from McCord's, Sherman's, and Weisburd's arguments that although the entire population will benefit from experimental knowledge, only a select few will bear the costs of producing it. Indeed, that is the point. If experimental criminology's ethical imperative is to produce evidence that is valid, then the power few hypothesis may be a step forward because it takes seriously the observation that crime concentrates in some units within the population. However, if the ethical imperative requires that we not only produce valid knowledge but also distribute the risk intrinsic to the production of this knowledge equitably, then the power few hypothesis may be instead a step backward, because it internalises the costs of an experiment on the same group every time. In sum, the problem for experimental criminology is not just how to observe an effect. It is also about how to distribute the costs of doing so.

THE DILEMMA

What are we to do? Criminologists have been attentive to the ethical concerns that experiments raise for a long time, and have often borrowed useful insights from medicine. Indeed, medicine presents itself at the onset in McCord's title, *Cures that Harm*. McCord borrows further in her vocabulary — iatrogenesis — a concept whose labcoat origins loom large. Sherman, too, frequently looks to medicine for guidance throughout his oeuvre (2009). Weisburd also laments the "famine" in experimental criminology and the "feast" in medical research (*see also* Shepherd 2003).

Medicine offers helpful cues for many ethical concerns that confront experimental criminology. Criminologists note that some of those cues have calmed the worry that randomised experiments can be conducted ethically (Lum & Yang 2005; Weisburd 2010). But confidence in the ethical viability of criminological experiments generally focuses on issues arising *within* the

experimental design, such as those pertaining to withholding treatment, group assignment based on treatment need, when experiments ought to be discontinued, and the vulnerability of research subjects (*e.g.*, Baunach 1980; Boruch et al. 2000; Erez 1987; Farrington & Welsh 2006; *etc.*). Those discussions leave untouched the ethical concerns we raise here, which extend beyond the experimental samples alone and implicate the experimental participants' relationship to the general population.

The structure of the free-rider problem makes proposed fixes that focus solely on ethical concerns within the sample a non-starter. Take, for example, the question of whether voluntary participation in research solves the free-rider problem. Ethicists might note that informed consent is an important provision in codes of research ethics because it ensures that experimental samples include only those subjects who volunteered to participate. Because research subjects who willingly submit to participate in an experiment freely bear the costs of doing so, their willing submission might defuse the ethical worry associated with backfire risk within the sample. The problem with this approach is that while it focuses on cost and benefit within the experimental sample, it overlooks the distributive justice concern that reaches outside the sample to the population beyond, and from which the free-rider problem emerges. The free-rider problem shifts the ethical concern that voluntary participation proposes to solve, namely one that pits the interests of eligible volunteers against eligible non-volunteers, to one that instead pits the interests of people on whom we propose to experiment against people on whom we do not propose to experiment. We should worry about a research programme in which we consistently ask some to choose whether they are willing to participate in research that benefits others who do not participate. No matter how much buy-in we get from volunteering participants, free-riders remain.

Along similar lines, we could condition approval to conduct criminological experiments upon first demonstrating low risk through a preliminary battery of observational research designs. This staggered sequence of escalating methodological rigour before introducing studies with a manipulation criterion analogises to the graduation from exploratory Phase I and II research to confirmatory Phase III and IV research in therapeutic development (Friedman et al. 2015). This solution also fails to solve the free-rider problem because, as McCord teaches, even if we could minimise the risk that an intervention might backfire, we cannot eliminate it. And this is the basis of our moral imperative to experiment: experiments contain the risk of backfire within the experimental sample. Insofar as each experiment poses a discrete instance of distributive injustice, reducing the number of experiments reduces the scale of distributively unjust experimentation. But reducing the scale of the problem is not the same as solving it. And besides, reducing the scale of experimentation runs directly counter to the stated platform of experimental criminology, because more experiments generate more robust findings that can provide more policy solutions at scale to the population outside the sample.

For the same reason, we could even consider abandoning experiments entirely. But given experimental criminology's many benefits, although this proposal seems parsimonious, it is also both extreme and untenable in that it throws the baby out with the bathwater. The matter at issue here is the narrower one, which we believe has gone unnoticed, of adjudicating the costs and benefits that experimental criminology's free-rider problem implicates.

This leaves a dilemma. Either we can accept the bioethical proposal to require full participation in the research process, and thus to experiment on both the sick and the healthy alike. In criminology, this would leave the analogy to medicine intact, but it would require experimenting as zealously on the 'weak many' as the 'power few.' Or, we can contest the applicability to

criminology of medicine's proposal to experiment on the healthy. Doing so triggers two further questions: first, why, and under what conditions does the medical analogy break down? Second, if medicine's solutions do not fit, can criminology come up with another solution to the free-rider problem?

Consider each option in turn. We should be encouraged that a medical template is already in place for what to do about the free-rider problem. Moreover, medicine is a particularly appealing place to look because bioethicists deal with the free-rider problem specifically in the context of the research process. Briefly, to apply this template to criminology would entail reconsidering the sites and subjects of experimental samples. Bioethicists insist that medicine has much to learn from the healthy as well as the sick, and that the field will overcome the free-rider problem when — and only when — it also puts healthy participants to experimental use. Thus, just as bioethicists argue that an ethical obligation exists to experiment on the healthy and the sick alike, so too criminologists would be obliged to focus not just on the power few but also the weak many.

Experimental criminologists might well be concerned about the loss to statistical power that would accompany leaving the power few hypothesis behind. They also might doubt that there is *anything* to be learned by experimenting on whatever criminology's equivalent of medicine's 'healthy' subjects might be. If so, then criminology will have to reckon with the medical analogy's limits. Plenty of criminologists have pointed in this direction before (Garland 1985; Katz & Abel 1984; MacNamara 1977), and thoughtful critique of the medical analogy's applicability to criminology is not new (*e.g.*, Thacher 2001; 2015). For example, where medical practitioners share the goal of providing care to patients (McNeill 1993; Veatch 1987), criminal justice pits competing interests and harms against one another (Thacher 2015). To be sure, criminal justice practitioners address competing ethical imperatives between roles in an adversarial system (*e.g.*, public defender

and prosecutor; parole officer and community social worker), and they may contain competing ethical imperatives within roles beholden to assorted constituents (*e.g.*, a correctional officer's duties to personal safety, to the safety of their charge, or to the facility's institutional mandate). Although the criminal justice researcher may come from one of these communities and so be bound to one of these professional ethical imperatives (*e.g.*, the refinement of police practice or service provision to survivors of domestic violence), she need not be so situated. Given criminal justice's value-plurality, it could be challenging to imagine what use could be derived from analogising to medicine's call to experiment on the 'healthy.'

That leaves the second solution. We could ethically continue to experiment on the power few if we can address two queries. First, we would need to know why medicine's requirement to experiment on criminology's equivalent of the 'healthy' is inapplicable; second, we would need to devise a different solution to the free-rider problem. This is not insuperable, but overcoming it does necessitate a change of perspective. Criminologists may need to look further afield, to other proposed remedies to free-rider problems in further-removed contexts than the research process.

Economists, for example, offer two workarounds to a free-rider problem: either public policy may subsidise the private provision of the public good, or the government may provide the public good itself. Both solutions internalise externalities by taxing free-riders, either directly or indirectly, such as through a system of political bargaining or regulation (Groves & Ledyard 1977). Solutions of this sort make intuitive sense for public goods like national defense and clean air. For example, when one of two smog-producing factories installs a smog remediation device to meet a municipal clean air ordinance, the other factory will free-ride unless it is forced to subsidise the remediating factory for the device's installation and upkeep. In this model, payment subsidies

neutralise the factory's riding for free. Analogously, we might entertain the thought experiment of a cap-and-trade model of experimental criminology.

However, squaring the standard economic solutions with experimental criminology is a stretch. Proposals to regulate the 'market' of research participants by 'taxing' the weak many to 'subsidise' the power few readily inspire ridicule and outrage. We suspect that few readers — let alone members of the general public — will endorse paying a fee to residents of a city's hot spots to subsidise the costs of producing knowledge from a policing experiment. Likewise, we would not stomach paying money to repeat domestic violence offenders to continue testing experimental interventions on them, even if the results would yield a public good. Perhaps, then, we would need a more creative solution to experimental criminology's free-rider problem.

The outlook is similarly bleak if we look even further afield than economics for solutions to the free-rider problem. Political scientists argue that states solve free-rider problems of governance by finding ways to incorporate smaller units of government into a larger polity (Cooter & Siegel 2010). This resembles the second of the economists' proposals, which recruits government to provide the public good itself. In experimental criminology, this quickly begins to resemble the bioethical proposal to force full participation in research experiments. This, in turn, soon reverts to the thorny matter of preserving the medical analogy and abandoning experimentation on the power few, only this time with the authority of state force adding further complexity.

Moreover, stringent experimental protections for vulnerable populations do not fix the problem, either. Ethicists in development studies, for example, rightly caution that the ethical stakes of conducting field experiments on vulnerable subjects are much higher than otherwise, and research designs should adjust accordingly (*e.g.*, Dionne et al. 2016; Teele 2014). But, as we have argued above, even the strongest protections for vulnerable research subjects would not mitigate

the free-rider problem: such protections attach only to subjects *within* the trial, whereas free-riders emerge from the experimental subjects' relationship to the world beyond.

CONCLUSION

We began this article with a thought experiment: what if, following the success of a rigorous experimental research programme in Minneapolis in the 1980s, investigators insisted that all subsequent criminological experiments ought to take place *only* in Minneapolis? Imagine — just for a moment — that this proposition turned out to be not only methodologically tenable, but moreover that it was methodologically desirable; would it also be distributively just? Would it be ethical to experiment on a handful of Minneapolitans over and over again if we knew doing so would generate useful knowledge for the benefit of all of society? Moreover, would an assessment that Minneapolis is an especially vulnerable city, and that further testing must take place there *because* the properties of vulnerability yield better insights, sharpen or blunt our worry?

Our hypothetical Minneapolis example may seem outlandish, but experimental criminology asks us to take it seriously. As a maturing field, experimental criminology has articulated a research programme that can test the effectiveness of criminal justice interventions with a methodological rigour that evaluation studies from earlier decades lacked. Yet the refinements that have improved the methodological quality of criminological programme evaluation coalesce as an interlocking set of propositions that, when taken together, present a free-rider problem. Namely, problems concentrate among a group that shoulders the risk that an experiment might backfire, and others ride on the benefit for free.

We deduce the free-rider problem from three classic texts in experimental criminology. Joan McCord's *Cures That Harm* shows that even interventions that are intended to benefit the treatment group may in fact cause harm (*i.e.*, backfire). Lawrence Sherman's *The Power Few* hypothesises that experiments that concentrate on the units in a sample where a phenomenon

concentrates (*i.e.*, the power few) may offer the best chance of observing an effect. And David Weisburd articulates a moral imperative to contain the risk of neutral or backfire treatments within the experimental population, so as to avoid exposing the public at large to the risk of harmful policies pursued without an evidence base.

Each of these texts alone provides a laudable insight for experimental criminology; but, in concert, these foundational texts present an ethical bind. Testing interventions on those units that have been selected because of where problems concentrate will also, perforce, concentrate the risk of treatment backfire there. Test interventions on units at one end of the distribution, experimental criminology counsels, and in so doing we produce knowledge that everyone else can use. This research programme amounts to a call for an unjust distribution of risk and reward.

We find that a way forward for the free-rider dilemma in experimental criminology will take one of two forms. The first form continues the analogy to the medical model. In that case, we could take lessons from bioethics and internalise experimental criminology's externality by rejecting the power few hypothesis on normative grounds. Doing so, in turn, requires experimenting on criminology's equivalent of the 'healthy' and revisiting assumptions about how experimental criminology distributes benefits and burdens. A second form reaffirms the power few hypothesis. In that case, we could stipulate that the power few contain the units in the population where useful criminological knowledge can most likely be gleaned. Maybe crime is not like medicine; although medicine may benefit from studying the healthy, criminology sees no benefit from doing the same. If this is the case, criminology is still stuck with the free-rider problem, and experimentalists must find another way to meet the ethical burden.

WORKS CITED

Academy of Criminal Justice Sciences (2000) *Code of Ethics*. Available at: https://perma.cc/5FBQ-2A47 (accessed 22 July 2020).

American Society of Criminology (2017) *Policies and Procedures Manual, Appendix Q: Code of Ethics*. Available at: https://perma.cc/X5UV-E4D4 (accessed 22 July 2020).

Armstrong R, Gelsthorpe R and Crewe B (2014) From paper ethics to real-world research: Supervising ethical reflexivity when taking risks in research with 'the risky.' In: Lumsden K and Winter A (eds.) *Reflexivity in Criminological Research: Experiences with the Powerful and the Powerless*. Basingstoke: Palgrave MacMillan. pp. 207-219.

Baunach PJ (1980) Random assignment in criminal justice research—Some ethical and legal issues. *Criminology* 17(4): 435-44.

Belmont Report (1979) National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, Department of Health, Education and Welfare. Washington, DC: United States Government Printing Office.

Berk RA (2005) Randomized experiments as the bronze standard. *Journal of Experimental Criminology* 1(4): 417-433.

Boruch RF, Victor T and Cecil JS (2000) Resolving ethical and legal problems in randomized experiments. *Crime and Delinquency* 46(3): 330-353.

Braga AA, Turchan B, Papachristos AV and Hureau DM (2019) Hot spots policing of small geographic areas effects on crime. *Campbell Systematic Reviews* 15(3): e1046.

British Society of Criminology (2015) *Statement of Ethics for Researchers*. Available at: https://perma.cc/C5V2-LQK9 (accessed 22 July 2020).

Buerger M (1993) *Convincing the recalcitrant: An examination of the Minneapolis RECAP experiment*. Unpublished Ph.D. dissertation. Newark, NJ: Rutgers University.

Campbell DT (1969) Reforms as experiments. *American psychologist*, *24*(4), 409-429.

Carlen P (2016) Ethics, politics and the limits of knowledge. In: Adorjan M and Ricciardelli R (eds.) *Engaging with Ethics in International Criminological Research*. Routledge. pp. 52-65.

Cartwright N and Hardie J 2012. *Evidence-based policy: A practical guide to doing it better*. Oxford University Press.

Cooter R and Siegel N (2010) Collective action federalism: A general theory of Article I, Section 8. *Stanford Law Review*, 63: 115-186.

Cooter R and Ulen T (2012) *Law and economics*, 6th ed. Boston, MA: Pearson.

Cowburn M, Gelsthorpe L and Wahidin A (eds.) (2017) *Research Ethics in Criminology: Dilemmas, Issues and Solutions*. London: Routledge.

Dionne KY, Harawa A and Honde H (2016) The ethics of inclusion when experimenting in impoverished settings. In: Desposato D (ed.) *Ethics and experiments: Problems and solutions for social scientists and policy professionals.* New York: Routledge. pp. 17-41.

Emanuel E, Wendler D and Grady C (2000) What makes clinical research ethical? *JAMA* 283(20): 2701-2711.

Erez, E (1987) Randomized experiments in correctional context: Legal, ethical, and practical concerns. *Journal of Criminal Justice* 14(5): 389-400.

Farrell G (2015) Crime concentration theory. *Crime Prevention and Community Safety* 17(4): 233-248.

Farrington DP and Welsh BC (2006) A half century of randomized experiments on crime and justice. *Crime and Justice* 34(1): 55-132.

Feeley MM (2007) Legality, social research, and the challenge of Institutional Review Boards. *Law & Society Review* 41(4): 757-776.

Friedman LM, Furberg CD, DeMets DL, Reboussin DM and Granger CB (2015) *Fundamentals of clinical trials*, 5th ed. Springer.

Garland D (1985) The criminal and his science: A critical account of the formation of criminology at the end of the nineteenth century. *British Journal of Criminology* 25(2): 109-137.

Gladwell M (2006) Million-dollar Murray: Why problems like homelessness may be easier to solve than to manage. *The New Yorker*, 13 February 2006. Available at: https://perma.cc/PY74-E8SK (accessed 22 July 2020).

Gottfredson D (2010) Deviancy training: Understanding how preventive interventions harm. *Journal of Experimental Criminology* 6(3): 229-243.

Graebsch C (2000) Legal issues of randomized experiments on sanctioning. *Crime & Delinquency*, 46(2): 271-282.

Groves T and Ledyard J (1977) Optimal allocation of public goods: A solution to the 'free rider' problem. *Econometrica*, 45(4): 783-810.

Haggerty K (2004) Ethics creep: Governing social science in the name of research ethics. *Qualitative Sociology* 27(4): 391-414.

Hollin CR (2008) Evaluating offending behaviour programmes: Does only randomization glister? *Criminology & Criminal Justice* 8(1): 89-106.

Institute of Medicine (2010) *Value in Healthcare: Accounting for Cost, Quality, Safety, Outcomes, and Innovation*. Washington, DC: National Academies Press.

Israel M and Hay I (2011) Research ethics in criminology. In: Gadd D, Karstedt S and Messner SF (eds.) *Sage Handbook of Criminological Research Methods.* London: Sage. pp.500-509.

Katz J and Abel CF (1984) The medicalization of repression: Eugenics and crime. *Contemporary Crises* 8(3): 227-241.

Levine FJ and Skedsvold PR (2008) Where the rubber meets the road: Aligning IRBs and research practice. *PS: Political Science & Politics* 41(3): 501-505.

Linning SJ and Eck JE (2018) Weak intervention backfire and criminal hormesis: Why some otherwise effective crime prevention interventions can fail at low doses. *British Journal of Criminology* 58(2): 309-331.

Lum C and Yang SM (2005) Why do evaluation researchers in crime and justice choose non-experimental methods? *Journal of Experimental Criminology*, *1*(2): 191-213.

MacNamara DEJ (1977) The medical model in corrections: Requiescat in pace. *Criminology* 14(4): 439-448.

McCord J (2003) Cures that harm: Unanticipated outcomes of crime prevention programs. *ANNALS of the American Academy of Political and Social Science* 587(1): 16-30.

McNeill PM (1993) *The Ethics and Politics of Human Experimentation*. Cambridge, UK: Cambridge University Press.

Nagin DS and Sampson RJ (2019) The real gold standard: measuring counterfactual worlds that matter most to social science and policy. *Annual Review of Criminology* 2: 123-145.

Pate A, McPherson M and Silloway G (1987) The Minneapolis community crime prevention experiment: Draft evaluation report. *Washington, DC: Police Foundation*.

Pease K (1998) *Repeat Victimization: Taking Stock*. Crime Detection and Prevention Paper Series, Paper 90. London: Home Office.

Rawls J (2001) *Justice as Fairness: A Restatement*. Cambridge, MA: Harvard University Press.

Reiter K (2009) Experimentation on prisoners: Persistent dilemmas in rights and regulations. *California Law Review* 97(2): 501-566.

Rhodes R (2005) Rethinking research ethics. *The American Journal of Bioethics*, 5(1): 7-28.

Rhodes R (2017) When is participation in research a moral duty? *The Journal of Law, Medicine & Ethics*, 45(3): 318-326.

Rothman D (1991) *Strangers at the Bedside: A History of How Law and Bioethics Transformed Medical Decision Making*. New York: Basic Books.

Sampson RJ (2010) Gold standard myths: Observations on the experimental turn in quantitative criminology. *Journal of Quantitative Criminology* 26(4): 489-500.

Sampson RJ, Winship C and Knight C (2013) Translating Causal Claims: Principles and Strategies for Policy-Relevant Criminology. *Criminology & Public Policy* 12(4): 587-616.

Schaefer GO, Emanuel EJ and Wertheimer A (2009) The obligation to participate in biomedical research. *Jama*, 302(1): 67-72.

Shepherd J (2003) Explaining feast or famine in randomized field trials: Medical science and criminology compared. *Evaluation Review* 27(3): 290-315.

Sherman LW (2007) The power few: Experimental criminology and the reduction of harm. *Journal of Experimental Criminology* 3(4): 299-321.

Sherman LW (2009) Evidence and liberty: The promise of experimental criminology. *Criminology & Criminal Justice* 9(1): 5-28.

Sherman LW (2019) Burying the 'Power Few': Language and Resistance to Evidence-Based Policing. *Cambridge Journal of Evidence-Based Policing*, 3(4): 73–79.

Sherman LW and Berk RA (1984) The specific deterrent effects of arrest for domestic assault. *American Sociological Review*, 49(2): 261-272.

Sherman LW, Gartin PR and Buerger ME (1989) Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology* 27(1): 27-56.

Sherman LW, Schmidt JD and Rogan DP (1992) *Policing domestic violence: Experiments and dilemmas*. Free Press.

Sherman LW and Weisburd D (1995) General deterrent effects of police patrol in crime "hot spots": A randomized, controlled trial. *Justice Quarterly*, *12*(4): 625-648.

Steffensmeier DJ, Andersen Allan E, Harer MD and Streifel C (1989) Age and the distribution of crime. *American Journal of Sociology* 94(4): 803-831.

Teele DL (2014) Reflection on the ethics of field experiments. In: Teele DL (ed.), *Field experiments and their critics: Essays on the uses and abuses of experimentation in the social sciences*. New Haven, CT: Yale University Press. pp. 115-140.

Thacher D (2001) Policing is not a treatment: alternatives to the medical model of police research. *Journal of Research in Crime and Delinquency* 38(4): 387-415.

Thacher D (2015) Perils of value neutrality. *Research in the Sociology of Organizations* 44: 317-352.

Veatch R (1987) *The Patient as Partner: A Theory of Human-Experimentation Ethics*. Bloomington, IN: Indiana University Press.

Weisburd D (2003) Ethical practice and evaluation of interventions in crime and justice: The moral imperative for randomized trials. *Evaluation Review* 27(3): 336-354.

Weisburd D (2010) Justifying the use of non-experimental methods and disqualifying the use of randomized controlled trials: Challenging folklore in evaluation research in crime and justice. *Journal of Experimental Criminology* 6(2): 209-227.

Weisburd D (2015) The law of crime concentration and the criminology of place. *Criminology*, 53(2): 133-157.

Weisburd D, Hinkle JC, Famega C, and Ready J (2011) The possible 'backfire' effects of hot spots policing: An experimental assessment of impacts on legitimacy, fear, and collective efficacy. *Journal of Experimental Criminology* 7(4): 297-320.

Wolfgang ME, Figlio RM and Sellin T (1972) *Delinquency in a Birth Cohort*. London: Chicago University Press.

Zane SN, Welsh BC and Zimmermann G (2016) Examining the iatrogenic effects of the Cambridge-Somerville Youth Study: Existing explanations and new appraisals. *British Journal of Criminology* 56(1): 141-160.