BOOK REVIEWS


Two centuries ago, a self-educated London magistrate named Patrick Colquhoun published a landmark treatise, The Police of the Metropolis. The book was destined to become as influential on crime control as Karl Marx’s Capital became on the economy. Colquhoun became known as “the architect who designed our modern police” and the book became the blueprint for Sir Robert Peel’s construction of the London Metropolitan Police in 1829, as well as American versions of that institution a few years later. Colquhoun’s treatise opens with this Preface:

Police in this country may be considered as a new Science; the properties of which consist not in the Judicial Powers which lead to punishment, and which belong to Magistrates alone; but in the Prevention and Detection of Crimes, and in those other Functions which relate to internal regulations for the well ordering and comfort of civil society.

After Colquhoun, prevention, detection, and regulation became the goals of policing, with visible uniformed patrol its primary method. But as often happens with great ideas, much of the content of Colquhoun’s book was lost when it was put into practice. The extent to which police ever adopted Colquhoun’s ideas is as hotly debated as the extent to which Soviet Communism was faithful to Marx. On one point, however, there is widespread academic consensus. Since at least World War II, the American police have let the process of policing displace Colquhoun’s concern for the results of police work.

This insight was first and most convincingly articulated by University of Wisconsin Law Professor Herman Goldstein, whose Problem-Oriented Policing has already become the blueprint for a new police institution and the most influential treatise on police since

1 Patrick Colquhoun, The Police of the Metropolis (7th ed. 1806).
3 Wilbur R. Miller, Cops and Bobbies (1977).
4 Colquhoun, supra note 1, at preface.
Colquhoun's. In a 1979 article that introduced the conceptual framework of his book, Goldstein compared modern police to bus drivers who cannot stop to pick up passengers because to do so would make them fall behind schedule. He noted the obsession of the 20th Century police reform movement with the internal management of police organizations, and its virtually complete lack of concern with the substantive tasks police were asked to accomplish. Giving police chiefs civil service terms was an important reform issue; policing domestic violence was not. Educational requirements for police recruits was an important issue; reducing armed robbery was not. Controlling police corruption was an important issue; preventing barroom brawls was not. There was too much concern with administering the hospital and not enough concern with treating the patients. In the metaphor of business, no one was paying attention to the bottom line.

Goldstein went on to prescribe a strategy for policing for results. He advocated abandoning the criminal law as the organizing framework for defining problems, replacing it with substantive patterns of behavior. "Aggravated assault," for example, is far too broad a category of events to comprise a clear target for police efforts to control the problem. But youth gang rumbles, domestic stabbings, barfights and driveby shootings among drug dealers each constitute discrete, homogeneous crime patterns that police can diagnose and attack. "Taking a problem apart" was the most important step Goldstein advocated, to be followed by a diagnosis of the problem's causes and some hypotheses about the problem's possible solutions. It is this emphasis on clear definition of highly specific problems as targets for police attack which gives the problem-oriented policing strategy its name.

The 1979 article caused a national sensation in the police reform community. As one commentary observed, "Goldstein's prescription was so penetratingly simple and so profoundly sensible that it not only spawned a number of experiments in police agencies to attempt to put his ideas into practice but also triggered nearly a decade of debate over what his prescription really meant and how it might be realized." Goldstein's book reflects and reviews the first decade of the idea's impact, richly illustrated with case studies of problem selection, definition, diagnosis, solutions and evaluations.

The book begins with a restatement and elaboration of the basic 1979 framework. It then proceeds to assess the early experiences with the strategy in Madison (with Goldstein himself participating), Baltimore County, Newport News and London. He explores the major issue of how his own muscular conceptual framework relates to the flabby content of “community policing,” a politically appealing label which has become far more popular than Goldstein’s own. Separate chapters examine the identification of problems, their analysis, the search for better responses to them, and the overall changes in police management needed to make problem-oriented policing work.

The result is at once a magisterial review and a practical guidebook for action. It is a wise and thoughtful commentary based on extensive experience in the field, reflecting great erudition in the substantive police literature. Its calm and balanced tone understates the book’s radical attack on current police practice, with its revolutionary implications for all aspects of the police institution: who becomes police, how they are trained, how they are rewarded, what powers society will grant them. The proposals amount to nothing less than making the police work a matter of brains far more than brawn, a profession as complicated as neurosurgery or nuclear engineering. The fact that many current police (including my own students) are not able to get past the somewhat dense and formal style of the book is unimportant, since Goldstein’s work has many populizers.

I. The Paradox of Results

The most effective and creative populizers, John Eck and William Spelman, use the acronym “SARA” to help readers recall the four basic steps or problem-oriented policing: Scanning the environment for problem selection, Analyzing the problem selected, Responding to the problem in a new and different way, and Assessing the results of the new response. This structure helps clarify the process of policing for results. Yet it also helps reveal a paradox in Goldstein’s path-breaking work, found in the final letter of the acronym. The paradox is that the book’s emphasis on the importance of results is unsupported by any serious concern with the measurement of results. For despite the book’s success in focusing police attention on the effects of what they do, the book is almost cavalier in its disregard for the complexity of causal inference necessary for the assessment of those effects.

This paradox leaves the blueprint unfinished. Goldstein has designed a brilliant foundation for policing over the next several centuries. He has charted the ways in which creative people can build upon that foundation to attack crime problems which have not even been invented yet. But he has omitted a design element that is at least as important as a roof, without which the building cannot last.

The roof we need is some protection against self-delusion about success. Encouraging police to focus on results creates a general temptation to claim success for every attempted solution. From the patrol officer handling a single dispute to a police chief developing a city-wide policy, the inventors of solutions appropriately want to be acclaimed for their good work. The failure of a police effort to produce its desired results is more painful than not attending to results at all. Thomas Edison saw failure as learning one more way not to make a lightbulb, but police whose solutions fail may be emotionally devastated. Unless the rules of determining success or failure are clearly laid out in advance, problem-oriented policing may be destroyed by fraudulent claims of success.

Goldstein does not provide those rules. He does provide some general discussion of the issue, consuming less than three pages. But the discussion lacks clear examples, and seems half-hearted. While other sections of the book offer stern warnings of other dangers, this section shows no hint of admonition.

Business has a bottom line, but it also has the independent audit. Surgery has the patient’s recovery, but it also has the second opinion. Academics write books, but they must also undergo book reviews. Each of these systems is a set of rules for evaluating the results of professional work. The rules may vary in the fairness or accuracy of the assessments they produce, but they all insure that the results are judged by someone independent of those who did the work. Anything else is a conflict of interest. Yet Goldstein offers no clear warning against such conflicts.

That omission may encourage police to continue to take credit when crime goes down, while blaming “social factors” beyond their control when crime goes up. It may encourage police to declare that all their problem-solving efforts a success, regardless of the actual results. It may discourage police from learning the principles of casual inference needed for more rigorous evaluations of their

10 Goldstein, supra note 6.
At worst, it will leave police as ignorant of the results of their work as they were before the book was written.

II. Two Case Studies

Two examples from the book illustrate its paradox of results. Both are case studies which Goldstein presents as strong examples of successful results of policing. One of his conclusions seems quite compelling. The other one violates almost every cannon of evaluation research.

A. The Case of the Vibrating Jukebox

The first case comes from Philadelphia, where a sergeant was exposed to the concept of problem-oriented policing. Assigned to a new district, the sergeant discovered police officers were being dispatched several times a day on a noise complaint about the same bar. On checking, he found there had been 505 calls about the same problem in just the last six months. 

The officers assigned to the 505 calls had always responded to the bar, but never found the noise to be very loud. After some effort to identify the complainant, the police conducted decibel meter studies to show that the noise in the bar did not violate local law. But further analysis showed it was the vibrations from the jukebox that disturbed the bar's neighbor, shaking the common wall between the two buildings. Police then negotiated a solution: the jukebox was moved to another wall, and the calls for police service stopped.

This case study is an archetype of common sense. Not only was the solution painfully simple, so was the assessment of its success. To question whether the solution "worked" is to belabor the obvious. The cause of the problem (the jukebox location) was easily isolated and eliminated. Once that was done, the problem—defined as calls to the police—clearly stopped because of what the police did.

---

11 See, e.g., Thomas D. Cook & Donald T. Campbell, Quasi-Experimentation: Design and Analysis Issues for Field Settings (1979).
12 Goldstein, supra note 6, at 81.
13 This fact alone speaks volumes about the state of modern police work. It may help the uninitiated reader understand why Goldstein's ideas are so important. Only an institution blindly going through the motions of answering calls for service, without any regard for the results of those responses, could allow such a pattern to develop. The fact that different officers may have been dispatched on each separate occasion may help explain it, but does not justify it. It is precisely the need to look for and address such problems that is the basis of Goldstein's strategy.
14 Memorandum from William Bugg to Inspector Edward McGlaughlin, Philadelphia Police Department, March 15, 1988, cited in Goldstein, supra note 6, at 81.
The sequel to the police action was obviously a result of police work, in the sense that police work caused the result to occur. To suggest that the calls may have stopped because the neighbor died, or moved away, or had her phone disconnected for nonpayment of bills, or was threatened by the bar owner, or was caused by any other reason besides the police intervention would be bad manners at best.

But this case is deceptively simple. The close correlation in time between the police action and the cessation of calls makes the hypothesis of police success quite compelling. As a matter of common sense, the odds of a rival hypothesis explaining the cessation of calls seems very low. While some other explanation is admittedly possible, it seems extremely improbable in the circumstances. If every challenge for problem-oriented policing were as simple to assess, there would be no point to the present essay.

But the principle remains: the elimination of rival hypotheses. As philosopher of science Karl Popper has argued, it is not possible to "prove" a causal relationship between two events, or any hypothesis. All one can do is fail to disprove it. One of the best ways to do that is to eliminate rival hypotheses, or develop alternate theories about why a given event occurred—such as the cessation of noise complaints about the bar. The more rival theories are eliminated, the greater the confidence one places in the central hypothesis. Under this philosophy of science, the truth of any theory becomes a matter of degree rather than an either-or question.

Few of the problems police must address, especially the most serious, provide such clear-cut evaluations of police efforts. Goldstein concedes this point, but then dismisses academic-style debate over the results of police-work as likely to "inhibit well-motivated, innovative people in the police field from making any claims about the value of their efforts." That may be true if such debate is conducted crudely, but not if the principles of evaluation are introduced slowly to police culture and presented as integral to the very idea of problem-oriented policing. By omitting proper evaluation techniques from his blueprint, Goldstein runs the risk that "well-motivated, innovative people" will fool themselves into...
believing they are successful, even when they are not. For while he warns against "making bold claims for the value of police strategy without any persuasive evidence,"19 he fails to warn against making bold claims when there is persuasive evidence both for and against success. While he advocates a "middle ground by thinking through more carefully what is involved in making statements about police effectiveness,"20 the book does not provide that thinking. By failing to put the weight of his enormous prestige and authority behind the importance of principles of evaluation, he provides a de facto endorsement of overblown claims of success and leaves more rigorous independent evaluators vulnerable to charges of sour grapes.

B. THE CASE OF THE CONVENIENCE STORE CLERKS

The second case study in the book is a prime example of this failing. In the mid-1980's, the Gainesville Police Department conducted a master-piece of analysis in the spirit of problem-oriented policing. Goldstein justifiably spotlights their effort as a model of diagnosis.21 The problem selected was convenience store robberies, which had been steadily increasing in recent years. To understand the problem, Gainesville police undertook pathbreaking research, far beyond anything ever published in the criminological literature. They discovered, for example, that the prevalence of robbery was much higher for convenience stores than for other types of retail establishments in town. Over a five-year period, 96% of all forty-seven convenience stores were robbed, compared to 36% of the sixty-seven fast food establishments, 21% of the seventy-one gas stations, and 16% of the forty-four liquor stores.22 While they failed to take account of the differences in the number of hours the stores were open and therefore at risk of being robbed,23 they were at least able to confirm the scope of the target problem. They went on to commission an interview study of incarcerated convenience store robbers, and to analyze the pattern of robbery events as reflected in the narrative reports.

The pattern analysis showed that 92% of the robberies occurred when a single clerk was on duty, usually with no other cus-

19 Id.
20 Id.
21 Id. at 80.
tomers present in the store. Often a robber would enter while other customers were present and wait for them to leave before committing the robbery. The interviews with incarcerated robbers showed that they rated having two clerks in the store as an important deterrent compared to other possible factors, although having a big strong clerk alone was about as important to them. These findings together supported the department's working hypothesis that requiring two clerks in the stores at night would serve as a deterrent to convenience store robbery.

The commendably detailed analysis, which was written up for presentation to the City Council and later presented to national scholarly meetings, does have an interesting omission. It gave no consideration to the rival hypothesis that clerks working alone are in an excellent position to commit fraud in reporting robberies as a cover for their own embezzlement. The longstanding problem of fraudulent robbery reports has been cited by former police chiefs in Kansas City and Washington, D.C. One Kansas City example was discovered when repeated robbery reports at one convenience store led police to assign a stakeout team to the location, without informing the clerk on duty. While the team was observing the location, the clerk phoned in a robbery report. The officers immediately entered the premises, where they had seen no one enter or leave for many minutes prior to the clerk's call. A cash register audit showed the clerk to be the thief, and he was charged with the crime. How widespread this problem is may be anyone's guess, but a three year review of robbery records in one southwestern convenience store chain found that at least 9 of the 149 (6%) recorded robberies were discovered to be employee thefts fraudulently reported as robbery. Undetected frauds could be far more numerous.

The Gainesville analysis was impressively detailed, but more as an advocate's brief than as a judicious opinion. It did consider contrary evidence from another interview study of convicted robbers (not necessarily of convenience stores) in another state, but dismissed it as suspect since it was sponsored by the convenience store industry. The analysis was careful in the way lawyers are careful, rather than scientists: attempting to prove an argument rather than

24 Clifton, supra note 22.
25 Interview with Larry Joiner, Kansas City Police Chief (April 1989).
26 Interview with Jerry V. Wilson, former Washington D.C. Police Chief (November 1989).
27 Joiner, supra note 25.
to investigate a hypothesis. Its goal of legislative action was clearly achieved. The recommendation was adopted by the City Council in early 1987, and in the next six months police reported that the number of convenience store robberies dropped by 65%.29

Goldstein's book reports this entire story, along with the Philadelphia jukebox case, as two "classic examples of what is involved in problem-oriented policing: in the identification of problems, their analysis, and the development of an effective response."30 His account is so impressive that the New York Times editorial page, reviewing Goldstein's book, heralded the Gainesville case as a dramatic example of the successful results of the new approach to policing, recommending the approach highly for New York.31 Not incidentally, the Times coverage helped the Gainesville Police Department's advocacy of a Florida state law requiring all convenience stores to have two clerks on duty after dark. If passed, it would be the first such state law in the country, and a major innovation in public policy apparently endorsed by the authority of Goldstein's writings.

Yet nowhere does Goldstein employ the principle of rival hypotheses.32 The only hint of qualification he offers in judging the Gainesville case a success comes in one vague sentence that would mean little to most police (and other) readers: "Predictably, the long-range impact has been challenged by various interest groups, and efforts to respond definitively to the challenges reflect the difficulty in isolating the effect of any single change."33 Careful reflection on this sentence may reveal the principle of rival hypotheses embedded in it, but the reader has to work very hard to extract it. The book clearly does not use the principle for any serious consideration of the possibility that the decline in robberies might have been the result of factors other than the two clerk law.

Goldstein's sentence supports the Gainesville police view that anyone who questions their claim of success must be representing an "interest group", and that alternative explanations should be fought off as an attack on the police department.34 From all appear-

---

29 Clifton, supra note 22.
30 GOLDSTEIN, supra note 6, at 81.
32 See supra note 15.
33 GOLDSTEIN, supra note 6, at 81.
34 Full disclosure requires, for example, that the present writer report that he has been and is currently employed as an expert witness in tort cases involving convenience store security measures, on behalf of both plaintiffs and defendants. In the past, he has also been retained by the National Association of Convenience Stores (NACS) as the supervisor of the report prepared by Wilson, infra note 47, and as a consultant on the
ances, the Gainesville police sincerely believe that the two clerk law has "worked", and that its widespread adoption will help control crime in the U.S. Anyone who disagrees is the enemy. Just as the willingness to defer to police authority in the street encounters arbitrary arrests and even brutality, unwillingness to accept police assertions about solutions to crime problems can provoke outrage among police problem-solvers. Yet this kind of aggressive defensive posture, supported by some evidence and the rejection of any contrary evidence, is precisely what Goldstein should be warning police against. Evangelizing for a single solution, rather than remaining open-minded about the evidence of its effectiveness, may do little to help control crime in America.

If police accepted the principles of evaluation before undertaking any solutions, however, they could go a long way towards avoiding this problem of defensiveness. Goldstein's failure to instruct police in those principles is the most unfortunate omission in the book.

III. PRINCIPLES OF EVALUATION

The basic principles of evaluation are not mysterious, and certainly not beyond the comprehension of police problem-solvers. Understanding the principles does not require graduate education or extensive experience in research. All it requires is an open, logical mind and a little patience, qualities many police in my training classes have brought to their instruction in these principles. Some of the technical jargon may be a bit off-putting, but the principles are clearly explained in many undergraduate research methods texts.

Ever since evaluation pioneer Donald Campbell (with his colleague H. Laurence Ross) evaluated the Connecticut speeding crackdown over two decades ago, his formulation of evaluation principles has warned policymakers against accepting simple before-after changes as proof of effective programs. Without restating his design of better research to evaluate the effectiveness of two clerks. He has, however, no current or anticipated future financial ties to the convenience store industry.

37 Albert J. Reiss, Jr., Police Brutality—Answers to Key Questions, 10 TRANS-ACTION 5 (July-August 1968).
38 Personal observations of the Minneapolis Police Department's Repeat Call Address Policing Unit (January-December 1987) and a representative of the Gainesville Police Department at the Tallahassee, Florida meeting (October 1990).
39 See, e.g., Herman W. Smith, STRATEGIES OF SOCIAL RESEARCH 61 (1975).
40 Campbell & Ross, supra note 15.
arguments in full, it is appropriate here simply to apply those with
greater relevance to the case of the convenience store clerks. All of
them take the form of rival or alternate hypotheses to the conclusion
that the presence of two clerks at night deters convenience store
robberies.

A. HISTORY

The most basic principle for interpreting before-after change is
called history.41 This principle requires asking whether any other
change in the environment, besides the intervention being evalu-
ated, might account for the observed results. In the case of conven-
ience store robbery, there are some obvious questions to raise about
other changes in the environment. Did the number of convenience
stores in Gainesville decline because of the higher costs imposed by
the law? Did the number operating at night decline? Were there
any active robbery suspects specializing in convenience stores who
were apprehended at about the same time that the two clerk law
went into effect?

The best way to control for history is by examination of trends
in nearby or similar environments which did not adopt the policy
change, using them as a comparison or a control group. Thus we
would ask of Gainesville, what were the trends in convenience store
robberies in nearby jurisdictions? If the trends were up, we might
interpret that as meaning the strategy worked in Gainesville, though
at the price of displacement of Gainesville robbers to one-clerk
store environments in other jurisdictions. If the trends elsewhere
were flat, we would still give more credence to success in Gainesville
than if we did not know that fact. But if the trends nearby were
downward, it would provide strong evidence for a history effect and
weaken our confidence in concluding that the two-clerk policy was a
success.

B. TESTING

A second major principle is called testing, or some aspect of the
intervention and its measurement that may change the behavior of
the subjects.42 In the case of the convenience store robbery, the
measurement of the problem depended completely on reports by
the clerks themselves. When they were working alone at night, it
would be relatively easy for them to report a robbery that never oc-
curred, and then pocket the money. Once a second clerk was pres-

41 SMITH, supra note 39, at 62.
42 Id., at 63.
ent in the store, dishonest clerks were suddenly faced with the necessity of enlisting a co-conspirator, at least if they wished to use robbery as a cover for their embezzlement. Reluctance to trust co-employees could have led them to reduce substantially their rate of reporting robberies. They might conceivably develop plans for staged robberies, in which a friend comes to the store when the available cash is high and takes the money by threat of force to both clerks. But such accomplice jobs are also much harder to orchestrate than simple fraudulent robbery reports.

C. INSTABILITY

The principle of evaluation that even the Gainesville police acknowledged in their analysis, but which Goldstein does not report in his book, is called instability, or chance fluctuations in the frequency of the events being measured. Tests of statistical significance are usually employed to eliminate this rival hypothesis, and the Gainesville report noted the results of the first six months were not tested for statistical significance. Such a test would have been largely pointless, given the small numbers of events per month before and after the law went into effect. Robberies of convenience stores in 1987 were originally reported to have declined to thirty, from seventy-four in 1986. On a monthly basis, which is the appropriate unit of analysis for the relevant statistical tests, this represents a decline from 6.2 to 2.5 convenience store robberies per month, small numbers indeed for a nationally influential conclusion about effective policing. Yet Goldstein’s book does not even mention the actual

43 Id., at 68.
44 Some have made the same comment about the Minneapolis Domestic Violence Experiment, an early attempt to develop problem-oriented policing. Lawrence W. Sherman & Richard A. Berk, The Specific Deterrent Effect of Arrest for Domestic Assault, 49 AM. SOC. REV. 261 (1984); Richard A. Berk & Lawrence W. Sherman, Police Responses to Family Violence Incidents, 83 J. OF THE AM. STAT. ASS’N 70 (1988). Ironically, Goldstein (at 172) cites this viewpoint [Arnold Binder & James W. Meeker, Experiments as Reforms, 16 J. OF CRIM. JUST. 347 (1988)] as evidence for his conclusion that the Minneapolis experiment was “promoted in ways unjustified by its limited nature,” an issue he fails to raise about his own promotion of the Gainesville and other anecdotal cases of problem-oriented policing throughout the book. Yet the Minneapolis experiment was a randomized, controlled experiment finding statistically significant results with a sample of 314 cases, about which even a strong critic of promotion says, “from the standpoint of social science it is an excellent piece of work.” Richard A. Lempert, From the Editor, 18 LAW AND Soc’y REV. 505 (1984). The justification for “promoting” the Minneapolis experiment was fully equal to the justification for the standard “promotion” of medical research findings [Lawrence W. Sherman & Ellen G. Cohn, The Impact of Research on Legal Policy: The Minneapolis Domestic Violence Experiment, 23 LAW AND Soc’y REV. 117, 135 (1989)], especially because the promotion of Minneapolis encouraged a federal agency to fund replications of the original research in other sites which contradicted the Minneapolis results and helped advance police understanding of this problem. Franklyn Dunford,
number of convenience store robberies in the Gainesville experi-
ment; it only reports the percentages.

D. REGRESSION TO THE MEAN

A related principle of evaluation is called regression to the mean, or
the tendency of any extreme developments over time to return to
normal.45 The children of geniuses are almost never quite so smart,
the children of very short men are usually somewhat taller, stock
market booms go bust and busts rebound. In public policy evalua-
tions, the danger of this pattern is that the reason policies are
adopted is the problems they address have recently suffered ex-
remely high levels by historical standards. If the policy is adopted
at the peak of the distribution, a decline in the frequency would be
expected with no intervention. In reporting that the two clerk law
was followed by a 65% reduction in robbery, Goldstein does not
raise the question of whether the base level for that percentage
change was an historically high frequency period for that particular
problem.

E. INSTRUMENTATION

A more probing principle of evaluation is called instrumentation:
did the procedures for measuring the problem change at the same
time the new policy was adopted, as is often the case?46 In the
Gainesville case, this simply means assuring that the definition of
convenience stores remained the same before and after the new law,
at least for purposes of comparing the number of convenience store
robberies. While even asking such questions might be defined as
insulting, the answers can often be surprising.

F. APPLYING THE PRINCIPLES TO THE GAINESVILLE CASE

As it happens, answers to questions from the principles of eval-
uation became available shortly after Goldstein’s book went to
press. But the content of the answers is not the point. Far more
important is the fact that the book does not even raise questions
about the anecdotal evidence on problem-oriented policing. The

David Huizinga & Delbert S. Elliot, The Role of Arrest in Domestic Assault: The Omaha Police
Experiment, 28 CRIMINOLOGY 183 (1990); J. David Hirschel, et al., Charlotte Spouse Assault
key concept in science for understanding and verifying results, is a word that does not
even appear in the index to Goldstein’s book.
45 SMITH, supra note 39, at 65.
46 Id., at 63.
content of the answers about Gainesville merely demonstrates the importance of the more general point.

In a study funded under a small contract from the National Association of Convenience Stores—an admitted "interest group"—former Washington, D.C. police chief Jerry V. Wilson attempted in early 1990 to answer the questions raised above.\(^4\) Wilson holds no graduate degree and has no technical training in interrupted time series analysis, autoregressive integrated moving averages (ARIMA) modeling\(^4\), or other highly academic skills often used in before-after evaluation. But he was able to address each of the principles in a clear fashion, within the limitations of the data made available to him by the Gainesville Police and Alachua County Sheriff’s Office.

Wilson found that almost all possible rival hypotheses suggested by Campbell’s principles of evaluation had strong evidence to support them. More important, he found that the monthly number of convenience store robberies had dropped from seventeen to two, four months prior to the law becoming effective, with no change in monthly numbers after the two clerks were required. In other words, the essential facts of the claimed reduction were not before-and-after, as originally reported, but before-and-before. Thus we see the great need for respecting the first principle of evaluation (or any investigation): to establish the facts, quite apart from their interpretation, no matter what the problem-solvers have claimed.

The timing of the drop in convenience store robberies lends strong support for the \textit{history} hypothesis. The drop occurred in December 1986, about the same month in which the City of Gainesville activated a number of other mandatory convenience store security measures short of two clerks (such as restricted cash on hand, cameras for photographing thieves, and keeping signs off windows to increase visibility of the clerks). Also in that December, Alachua County Police told Wilson that three men suspected of being active convenience store robbers had been incarcerated. The men were suspected of committing robberies in both the City and the County, and both jurisdictions show the same sudden drop in robberies in the month the men were arrested. The County, however, adopted no new ordinances or convenience stores, and thus serves as a good control or comparison site. While there is admitted tension be-


tween the Gainesville Police chief and the Alachua County Sheriff, it is hard to imagine that the Sheriff’s staff would fabricate such dramatic statistical evidence just to make the Chief look bad.

Wilson found no evidence that the number of convenience stores had declined after the ordinance, so that there was apparently no change in the population of stores at risk. But the decline of robberies in that population preceded the two clerk law, and two strong alternate explanations highly correlated at the point in time when the drop did occur. This alone would tend to refute the conclusion that the robberies dropped as a result of the two clerk ordinance, although it is consistent with the theory that the drop was a result of some police work. The important issue, given the question of future legislation, is which aspect of police work.

The instability of the data was reported in the original Gainesville police study, but the regression to the mean was not. The period just prior to the enactment of the ordinance was, in fact, characterized by an unusually steep increase in the volume of convenience store robberies in Gainesville. This clearly supports the rival hypothesis of regression to the mean for at least a portion of the reduction in robberies. The annual numbers of such robberies\(^4\), without adjustment for changing definitions of convenience stores in police statistics, were fifty-two in 1981, twenty in 1982, twenty-seven in 1983, thirty in 1984, thirty-three in 1985, and sixty-one in 1986. The base year for calculating results of police work, then, had about twice the number of convenience store robberies as the annual mean of thirty-three for the preceding five years. The subsequent numbers were twenty-three in 1987 (when the two clerk law became effective in April), eighteen in 1988, and sixteen in 1989, for an average of nineteen, or about 33% lower than the mean before 1986. This is still impressive, but we must recall the many changes in the environment other than the requirement to employ two clerks at night. Moreover, Alachua County’s robberies of convenience stores dropped from fifty-three in 1986 to twenty in 1987. In both jurisdictions, most of the 1986 total had been in the months leading up to the capture of the three suspects.

Finally, Wilson did find difficulties with instrumentation that undermine the apparent 33% reduction from the 1981-85 period. After 1987, it is impossible to gauge the trends in Gainesville’s convenience store robberies compared to 1986 as the base year. The two-clerk law exempted nine of the fifty stores classified as

\(^4\) These figures come from Wilson, supra note 47, and from the Gainesville Police Department, supra note 22, which contains the figure of seventy-four robberies for 1986.
“convenience” by police in 1986, because they earned over half their sales from gasoline. The department could supply no statistics on robberies that included these nine stores after 1987. Wilson was unable to find any data on possible changes in the frequency of employee theft fraudulently reported as robbery.

In sum, Gainesville is hardly a “classic” example of an effective problem-solving process, as Goldstein claims.\(^{50}\) It is, rather, a classic example of the difficulties of interpreting the causes of changes in problems before and after a solution is implemented. Neither Gainesville in particular nor convenience store robbery are unique. Literally all of the examples of fully developed problem-solving in Goldstein’s book take the form of before-after comparisons, in one way or another, since that is the logic of his strategy. And to a greater or lesser degree, the principles of evaluation illustrated by the Gainesville case apply to all the other examples as well.

IV. HOW MUCH EVALUATION IS ENOUGH?

It is just as serious a matter to call a piece of police work a success as it is to call it a failure. As Goldstein himself observed in 1982, reflecting on his experiences in Madison\(^{51}\), it is a mistake to take statistical claims by police at face value—which is good advice for all governmental or corporate officials, for that matter. The Gainesville case shows why that advice should be in Goldstein’s book. What the case does not show is any clear standard for when to probe a claim of success, and when to accept it regardless of any possible rival hypotheses.

The question of how much rigor is necessary in evaluating the effects of public policies is clearly a difficult one, in which social science as a profession has a clearly acknowledged interest. Goldstein’s great contribution and the point of many of his examples, is to show that professional social scientists have no monopoly on the identification and diagnosis of problems or the invention of new solutions. That can and should be the basic part of police strategy. As Lindblom and Cohen pointed out the year Goldstein’s initial paper was published, social science is only one of many possible approaches to social problem solving, “because other forms of information and analysis—ordinary knowledge and causal analysis foremost among them—are often sufficient or better than” profes-

\(^{50}\) Goldstein, supra note 6, at 81. There is no doubt however, that he is correct in singling out as a model of analysis or diagnosis of a problem leading to a police response, whatever its deficiencies in assessing the effects of the response.

\(^{51}\) Herman Goldstein, Address at the Police Foundation Consultation on Police Order Maintenance Strategies (June 1982).
But the value of social science for solving problems is different from the value of social science principles for evaluating the effectiveness of a solution. Neither problem-solving nor solution-assessing are activities which require professional certification as a social scientist, but social science principles do have far greater value for the latter of the two activities. Police themselves can be trained in these principles of evaluation to perform more reliable peer assessments of results.

If problem-solving succeeds in becoming the predominant strategy of American policing, there may be far too many problems addressed for an independent review of every police result in any depth. Goldstein is right to seek a middle ground between no assessment at all, on the one hand, and complex or costly research, on the other. The vital question which he does not answer is what the middle ground should look like.

Failure to create a reliable system for assessing police results will mean a failure to institutionalize the importance of results themselves in American policing. If that happens, the basic goal of Goldstein's strategy will be displaced, even with widespread adoption of the process. The police may again become bus drivers too busy to stop for passengers. The only difference will be a different model bus.

One possible way to avoid this danger and to institutionalize better measurement of results is to train supervisors in principles of evaluation, supplemented by a specialized police evaluation unit for more complex assessments. The effort put into assessing results could then vary along several sliding scales. One scale might be the scope of the problem or its solution, with city-wide problems receiving more evaluation effort than solutions aimed at a single address or a single person. Another scale might be the level of police effort put into solving a problem; the more effort invested, the more intensive evaluation required. A third scale might be the on-going cost of the solution, either to the police agency or to some external organization. The more costly the solution—as in the case of convenience store clerks—the more probing the evaluation should be for rival hypotheses.

As Alfred North Whitehead once said, "Seek simplicity and distrust it." Goldstein's view of evaluating results of police work is

53 Quoted in Smith, supra note 39, at 315.
half right. While it is unnecessary to make every small scale police assignment a matter of complex or elaborate evaluation, it is absolutely necessary to subject big efforts to big assessments. The more important the issue, the broader the scope of persons affected by the police response, and the greater the cost of the response, the more simplicity of evaluation should be distrusted.

Colquhoun kept updating and expanding his *Treatise on Police* for a full seven editions. The enormous influence Goldstein’s treatise has already acquired will more than justify a revised edition. If future volumes can expand on the need for more rigorous assessment of police results by police themselves, Goldstein will complete his blueprint for building a new police institution. The result will stand far longer than the current version.

Lawrence W. Sherman
Professor of Criminology
University of Maryland at College Park
and
President, Crime Control Institute and Crime Control Research Corporation.

This book develops a critical examination of the politics of crime, the question of punishment in a just order, and reform in penology. In a rather short but clearly written, non-dogmatic style, Willem de Haan skillfully presents some clear-headed thinking about issues often circuitously addressed. Central to his efforts is a reconceptualization of punishment in terms of “redress,” and beyond that, a call for more creative, and even more idealistic theorizing.

De Haan first addresses the “crisis” in critical criminology. He traces the origins of radical/critical criminology and its search for a critical social theory. Pointing out the shortcomings of conventional criminological thought about the definition of crime and justification of punishment he notes that much of “realism”, whether Left or Right, is rooted in an enterprise constitutive of common-sense understandings of crime and punishment and hence supportive, inadvertently or advertently, of the consolidation of a conservative “law-and-order” mentality. Drawing inspiration from the abolitionist movement centered in Europe, he questions the wisdom of the Left in its tactics of simply replying to the conservative agenda of the Right. Necessary is transcendental activity, a praxis leading to societal change, rather than a reactive-negative politics. In fact, the Right has successfully captured the conceptual apparatus by formulating a common-sense answer to the compelling social problems of the age, including crime and punishment, while the Left has offered questionable solutions. De Haan, continuing to draw from the abolitionist movement, indicates that we should get away from defining troubles as “crime”; crime is a myth which well supports ever more pervasive methods of control. Instead, we should reconceptualize crime in an alternative discourse that centers on troubles faced by agents in a political economy. Perhaps, he suggests, crime should wither away as a concept and so too, theories of crime.

The Left often inadvertently contributes to the expansion of control-talk; for example, de Haan asks us to consider some recent social movements — environmental, feminist, and minority — and their generated “get-tough-with-the-violator” discourse. Clearly,
societal problems of major magnitude exist; the central question for de Haan becomes definitional and one concerning the appropriate societal reaction. More specifically, he argues, the current crisis rests more on the contemporary "fuzzy politics and morals" of the critical criminological movement. Coming from a noted critical criminologist, this self-reflective critique of the critical/radical criminological paradigm is welcome. All too often within "progressive" movements, in their haste to move from here to there, forms of domination are reinvented or reconstituted in new guise. The call for "reversing of the hierarchies," for example, a catchword incorporated from Jacques Derrida's postmodernist approach, too often means replacing one form of domination by another with self-serving ideological baggage and often resulting in further societal problems. This is yet another example of "fuzzy politics."

He then examines the often stated hypothesis that correlates unemployment rates with rising imprisonment rates. He shows that what has been overlooked is that the key factor seems to be the correlation between rising unemployment rates and increased duration of imprisonment, rather than just increased imprisonment rates. Looking at imprisonment rates and unemployment rates in the Netherlands, he disputes conventional critical criminological wisdom that argues that employment is the best predictor variable of imprisonment rates.

Next, de Haan suggests for serious thought Jurgen H. Habermas' much debated "ideal speech situation" as offering a potentially alternative method in reaching understanding in forums for dispute-settlements. He advocates a practical rationality devoid of instrumental means-ends rationality guided by the media of power, money, and exclusively efficiency-centered considerations. This thesis leads directly into the question of the necessity of punishment in a just order. Taking the position of a "moral rationalist" — that moral judgments can be rationally developed — de Haan advocates the use of reason rooted in deductive or inductive logic in context. Hence the appeal of Habermas' ideal speech situation. It is open, flexible to the unfolding activities of human activity, and allows each party to participate in a discourse replete with criticizable validity claims. Only in formalism with its underlying support of instrumental rationality are human agents stripped of their historically specific, multifarious nature. Thus, de Haan redefines the debate

---

1 See Jacques Derrida, Margins of Philosophy (Alan Bass trans. 1982).
2 This argument is particularly well-presented by Steven Box's much heralded findings reported in Recession, Crime and Punishment (1987).
3 See Jurgen H. Habermas, Legitimation Crisis 107-08 (1975).
concerning the question of the inevitability of punishment in a just order by reconceptualizing terms.

Accordingly, much theorizing by such notables as Alan Gewirth⁴ and Derek Phillips,⁵ he argues, are focused on sanction rather than punishment, and is hopelessly stuck with a unicausal model which stipulates punishment necessarily following crime. This leads inevitably to tautological arguments. This model just perpetuates suffering; it does not offer compensation for loss, according to de Haan. For de Haan, Habermas offers the potential for developing a substantive, rationally-based justification of moral principles rooted in symbolic communication, or in other words, practical rationality that is context specific.

In many ways, de Haan’s call does resurrect Max Weber's comparison of ideal-typical forms of legal thought such as formal rationality, formal irrationality, substantive rationality, and substantive irrationality.⁶ It was Weber who first noted the “irresolvable conflict” between legal principles rooted in substantive as opposed to formal rationality and justice. De Haan’s position could find a willing ally in Weber’s examination of substantive rational forms of legal thought and social control. This remains to be examined.

De Haan’s thesis also continues a recent trend by radical/critical theorists’ questioning whether punishment has a place in a just order. Jeffrey H. Reiman, for example, building on the John Rawl’s classic A Theory of Justice,⁷ offers what he calls rationally-derived first principles of justice which can be applied deductively.⁸ These moral principle embody the belief that “where people do cooperate to produce benefits that went into producing them... inequalities must work to maximize the share of everyone in society starting from the worst-off individual.” Reiman’s view, unlike de Haan’s, attempts to root a moral discourse materialistically in a political economy in his idea of a “labor theory of moral value.” Between de Haan and Reiman much important and assuredly useful debate will surely arise.

The next chapter examines the demise of popular justice tribunals in Cuba and the movement toward formalization. De Haan asks whether their experience can shed light on the development of moral first principles for movements in the United States. He offers

---

⁴ Alan Gewirth, Reason and Morality (1978).
⁵ Derek Phillips, Toward A Just Social Order (1986).
⁶ For a description of Max Weber and his work, see Anthony T. Kronman, Max Weber (1983).
much first hand experience as well as theoretical examination. He concludes that the Cuban socialist legality is useless in its potential applicability to a capitalist economy. He refrains from any outright condemnatory remarks, although, perhaps, one may take him to task and accuse his unwillingness to do so as a case of fuzzy politics. Certainly much good, as well as much questionable activity, has, and continues to take place in Cuba. To simply enter a conspiracy of silence concerning shortcomings (to be discussed only privately in Left circles), in my opinion, does a disservice to open and critical examination, and is counterproductive in developing visions of possible non-subjugating, alternative political economies and social formations.

In the last chapter, de Haan offers some insightful and refreshing suggestions about alternatives. Having already argued that practical discourse should be based on substantive rational principles, he offers the notion of "redress" to capture this alternative conceptualization to the formalism of crime and punishment discourse. After examining alternatives to legalistic definitions of crime — the Schwendingers' deductive approach on the one hand, and the feminist inductive approach on the other — he pleads that no solution to the definitional problem exists (p. 154). The real answer, again, is to reconceptualize terms. For de Haan, what needs to be instituted is a form of practical discourse under the principle of redress. It specifies a procedural justice rather than a formally substantive one. The elements of this discourse would include open-ended discussion about "unfortunate events," the encouragement of a diversity of reactions and analysis for each concrete situation, which in turn would encourage the proliferation of alternative meaning constructions. All this would be guided by a sensitivity to alternative dispute regulatory mechanisms which provide clues for alternative regulations of conflict. Decisions, therefore, should always be open-ended, subject to competing and criticizable validity claims.

Practical discourse, one rooted in context, would also develop, inductively, into general and principled conclusions (p. 159). It is only in this way that the complexities and multifarious nature of agency and accountability could be understood; on the one hand leading inductively to new principled conclusions, and on the other, acting as an open-ended guide for deductive lines of application.

---

To assure its feasibility, de Haan argues that we need (1) some minimal level of communicative competence by interlocutors and (2) some process by which potentially endless practical discourse could be brought to closure. As to the latter, some legal system still would be necessary to assure closure, fairness, and compliance (p. 167). This seemingly sticky dilemma of advocating, on the one hand, a practical discourse and, on the other, the necessity of a legal form rooted in legal rationality is answered by de Haan by his insistence for a search for the discovery of principles that are mutually compatible (p. 168). He seems to suggest that a continuously evolving system of inductive and deductive open-ended principles could address the question of generalizability as well as being sensitive to the uniqueness of contexts in producing a more human politics of redress.

What can be said of all this? De Haan has stimulated discussion for a wide audience. He demands that readers of his book become engaged. Responses will surely materialize from a broad political spectrum extending from the Right to Left, both praising and condemning his conclusions. However, he boldly encourages thinking in both directions: realism and idealism. The important point of all this is that we can have a happy marriage of the two. At a time where movements, both on the Left and the Right, are gaining momentum toward some form of “realism”, this is a courageous and admirable call. Certainly de Haan’s examination is not only critical of conventional criminology, but also critically reflects on the situation within his own paradigm. It’s about time! His suggestions are enlightening. We need, he tells us, a practical discourse rooted in context which can be given form in the notion of “redress.” This discourse would be open-ended. Inductively and deductively developed notions of moral principles could replace formal control-talk discourse based on purposive, mean-ends rational discourse which is guided by money, power, and exclusively efficiency considerations. His criticism of the Right in its articulation of a receptive, but questionable solution to the crime and imprisonment problem, as well as his criticism of the Left in its often fuzzy politics and morals, is a welcomed commentary to the often unconfronted shortcomings of particular paradigm-invoked thinking.

Dragan Milovanovic
Associate Professor
Department of Criminal Justice
Northeastern Illinois University