

EXPERIMENTS IN POLICE DISCRETION: SCIENTIFIC BOON OR DANGEROUS KNOWLEDGE?*

LAWRENCE W. SHERMAN†

I

INTRODUCTION

“If we would guide by the light of reason, we must let our minds be bold.” — Justice Brandeis, dissenting in *New State Ice Co. v. Liebmann*¹

Scientific experimentation holds the potential for revolutionizing police discretion.² It has already begun to erode the ignorance police have of the consequences of their actions after they leave an encounter. Science may soon be able to guide the discretionary allocation of scarce criminal justice resources to those offenders who are most deterrable, as well as to those who commit the most offenses. Science may also be able to tell police when *not* to act, if their actions would produce more harmful consequences than benefits.

Before rushing headlong into a brave new world of scientifically based police discretion, one must carefully consider its implications. Detailed knowledge of the consequences of police discretion in a wide range of situations could produce dangerous forms of inequity and discrimination. Moreover, that “knowledge” may be faulty in its application to particular cases. Yet, attempts to refine the knowledge to produce more accurate predictions in particular cases may produce even greater dangers of discrimination.

This article describes the problem that experiments in police discretion may solve, the experiments themselves, and the problems their findings may create. It articulates three familiar value models of police work that can be used to examine the uses or abuses of knowledge about the consequences of police discretion. Finally, it uses those models to examine the implications of (1) not knowing the consequences of police decisions, (2) knowing the *average*

Copyright © 1984 by Law and Contemporary Problems

* This paper was supported in part by Grant #80-IJ-CX-0042 from the National Institute of Justice, Crime Control Theory Program to the Police Foundation. Points of view or opinions stated in this document do not necessarily represent the official position of The U.S. Department of Justice, the Minneapolis Police Department, or the Police Foundation. The collaboration of Professor Richard A. Berk in the experiments described below is gratefully acknowledged.

† Vice President for Research, Police Foundation, and Professor of Criminology, The University of Maryland at College Park.

1. 285 U.S. 262, 311 (1932).

2. This paper defines “police” generically to include all publicly and privately paid agents of surveillance, conflict management, and apprehension who make decisions about placing people into the custody of governmental criminal justice systems. Cf. E. BITTNER, *THE FUNCTIONS OF THE POLICE IN MODERN SOCIETY* (1970) (using the word “police” to refer only to public officers).

effects of a police action on all kinds of people, and (3) knowing the *differential* effects of the same police action on different kinds of people in different circumstances.

II

THE LIMITS OF INTUITIVE POLICE KNOWLEDGE

What is the best way to handle a bar fight? What should police do when one neighbor threatens violence against another? What action should police take with juveniles suspected of vandalism or burglary? How should police deal with a man tampering with a car he does not own? When should police arrest, mediate, cajole, threaten, investigate, warn, laugh, or do nothing? Some of these techniques are no doubt more effective than others in different situations, but how should police match the response to the situation? Police make such judgments all the time; however, they may often err, measured by the short or long term consequences of their actions.

What do police officers know about the discretion they exercise? They say they know a great deal.³ They say that experience is the best teacher and that it helps to give them excellent judgment. Some police officers also contradict themselves, saying that every situation they confront is unique, that one cannot generalize, and that there are no predictable patterns of human behavior that can be discovered through scientific research.

No scientist should belittle the learning that comes from experience, however unsystematic.⁴ Experience, after all, is the foundation of empiricism and the antidote to superstition, religion, or dogma as a source of truth about the physical world.⁵ Direct experience, however, is better than indirect experience, and both are better than assumptions. The problem with experience as a basis for exercising police discretion is that it provides incomplete information with respect to each series of encounters. Police suffer a "funnel effect" of experience; as time passes, they have less and less knowledge about how their actions have affected a situation:

1. They have direct experience with people's responses to their behavior while officers and citizens are still physically present;
2. If some other officer gets a repeat call, involving the same parties, the officers who first responded may have indirect experience of what the effects of their actions may be after they leave the scene but during the same shift;
3. Generally, officers have only assumptions and no information about what effect their actions had on the people over the next day, week, month or year.

Police in small towns may know a good deal about what happens to all the people with whom they deal. Most police officers and most crimes, however, are concentrated in big cities, where the police and the public are largely

3. See generally Bayley & Bittner, *Learning the Skills of Policing*, LAW & CONTEMP. PROBS., Autumn 1984, at 35.

4. See C. LINDBLOM & D. COHEN, *USABLE KNOWLEDGE* 10-29 (1979) (comparing and relating "professional social inquiry," or scientific decisionmaking, to other ways of decisionmaking).

5. See D. BOORSTIN, *THE DISCOVERERS* 146-48 (1983) (describing how maps based on the experiences of sailors replaced maps based on theology).

unknown to each other. The urban police may deal with people in the most intimate settings concerning major events in their lives, and then never see those people again. That is not the fault of the individual officers. If fault exists, it is the consequence of the social organization of modern American police work.

The telephone and the patrol car have stripped police of their information base, shaping police work into a reactive mechanism for distributing legitimate coercive force.⁶ The police mission is defined as answering calls and being available to answer more calls. Police managers have created only an *input* information system, which tells police where to respond right away. They have failed to create a *feedback* information system which tells police what happens after they leave, or even after they make an arrest. In many cities, for example, police have no idea whether their arrests result in convictions or prison sentences. Police managers have failed also to create an *intelligence* system, which would provide contextual or background knowledge to help police better use their experience to predict the consequences of exercising their discretion in various ways.

The scope of police information systems limits intuitive police knowledge regarding the effects of their actions. Salespeople can develop intuitive knowledge about how to sell because they know when they have made the sale or not. Police officers cannot develop comparable knowledge because they never find out whether they "made a sale": whether the bar fight resulted in no subsequent acts of vengeance, whether the husband stopped beating his family, or whether the threat to attack a neighbor was carried out.

Compare the police to medieval masons or carpenters. Those craftsmen used discretion in constructing buildings. If the buildings later fell down, they probably heard about it. Compare police to early navigators. The navigators used discretion and guesswork to plot a course. If they landed on shoals, or in the wrong country, they received immediate feedback on the effects of their discretion. Any of the great crafts created without science had the same characteristic: immediate feedback on the aftermath (if not the causally related consequence) of a discretionary action. Police work has been called a craft,⁷ but such a term overstates the feedback police receive on the aftermath of their craftsmanship.

Police information systems not only fail to give feedback on the aftermath of police actions, but also fail to keep track of the "treatments," or police actions, taken by other officers. Police interacting with suspects and victims within an information vacuum is comparable to a doctor treating a patient without a medical chart. The officer at the scene has no idea whether another officer was there yesterday or last week, or what the officer did. The two officers may respond to the same problem in contradictory ways, canceling

6. Sherman, "Watching" and Crime Prevention, J. CONTEMP. STUD., Fall 1982, at 87, 89-90; see A. REISS, JR., THE POLICE AND THE PUBLIC 2-5 (1971). See generally E. BITTNER, *supra* note 2, at 36-48, 89-92.

7. J. WILSON, VARIETIES OF POLICE BEHAVIOR 283 (1968).

out any good effects or perhaps making matters worse. Just as different doctors, if they do not read a patient's chart, may prescribe different drugs that could interact to kill the patient, police can do great harm because they have no chart to read. Computer terminals in police cars may solve this problem in the future, if the terminals are ever programmed creatively to provide such information.

The absence of feedback, intelligence, and treatment information systems for police is hardly accidental. Many people prefer a society in which police are structured into systematic ignorance. Any debate on the merits of that ignorance, however, should consider the full range of its consequences. Moreover, the police themselves cannot credibly claim that experience is an adequate basis for their exercise of discretion when they are denied the information required to give them the necessary experience.

Suppose police acquired the necessary experience. Would that experience be sufficient? Would good feedback about the aftermath of police actions allow officers to learn to predict with some accuracy the effects of their actions? It probably would not. Given the complexity of human behavior (in contrast to, for example, the simple physics of carpentry), the effects of police actions are not likely to be determined only by the nature of the actions themselves. The consequences of police actions may depend on how those actions interacted with factors such as a suspect's age, sex, relationship to victim, education, employment, criminal history, church attendance, or even diet.⁸ It is unlikely that police officers, or anyone, could make sense of these complex interaction effects without systematic statistical controls. Just as trial lawyers have learned how to pick a good jury through experience, police may learn how to use discretion better if they obtained better feedback. Some trial lawyers, however, now supplement their intuitive conclusions with systematic jury research. Police, similarly, could well supplement their experience with statistical research.

In any case, police clearly lack reliable knowledge about the effects of their discretionary actions on suspects, victims, witnesses, and potential criminals. Whether better feedback could give police better knowledge is a moot question, since it seems unlikely they will develop better information systems in the near future. What seems more likely is that scientific experiments will provide police with more reliable predictions than individual experience could produce, even if better feedback systems were developed.

III

THE DEVELOPMENT OF DISCRETION RESEARCH

The contribution of science to police discretion has followed a classic pattern of development from description to explanation to prediction. The descriptive research began in the 1950's with the American Bar Foundation

8. See, e.g., G. NETTLER, *EXPLAINING CRIME* (1984).

(ABF) field studies of criminal justice in three states.⁹ The ABF research discovered that police were playing a judicial, rather than ministerial, role in deciding when to make arrests, and that arrest and detention were being used for purposes other than prosecution.¹⁰ The ABF discovery apparently shocked legal scholars,¹¹ but the police managed to avoid major repercussions by refusing to acknowledge the findings.

The police got into bigger trouble during the civil rights movement from the claim that police were biased against minorities. This social context directed the discretion research of the 1960's to explanations of *police* behavior based on suspect characteristics, rather than explanations of *suspect* behavior based on police actions. Observational studies¹² in the mid-1960's and early 1970's, showing blacks more likely to be arrested than whites (with a complex array of intervening variables), were replicated in different cities in 1977 with similar results.¹³ Quite apart from any scientific findings, activists had assumed for twenty years that discrimination existed. Nonetheless, the question of how to prevent discriminatory discretion remained unanswered, largely for lack of ideas about methods of controlling use of discretion by police.

By the 1980's, the scientific agenda on police discretion clearly had shifted to the consequences of discretion and away from its causes. In a major conceptual statement, Professor Herman Goldstein called for a problem-oriented approach to improving policing.¹⁴ Goldstein's argument laid the foundation for solving the problem of police ignorance regarding the consequences of their actions and for going beyond the limits of intuitive police knowledge. Goldstein's suggestion was to take specific crime and order problems, assess the various ways police could deal with them, and determine which ones work best. The clear implication of this proposal is use of the scientific experiment.

The classic scientific experiment can produce more conclusive findings about causal connections between two variables than can any other research method. Not only is it a better method than raw experience for developing predictions, but it is also better than most other methods of conducting formal research. As a recent report to Chief Justice Burger points out, experiments rule out both systematic and chance differences between groups receiving different treatment by the criminal justice system, so that any subsequent differences in their behavior will be caused (most likely) by the criminal

9. See Smith & Stason, *Preface* to W. LAFAVE, *ARREST: THE DECISION TO TAKE A SUSPECT INTO CUSTODY* at ix-xi (1965).

10. W. LAFAVE, *supra* note 9, at 437-38.

11. E.g., Goldstein, *Police Discretion Not to Invoke the Criminal Justice Process: Low Visibility Decisions in the Administration of Justice*, 69 *YALE L.J.* 543, 586-88 (1960) (urging full enforcement of all laws with legislative repeal of obsolete or unenforced laws).

12. These studies are reviewed in Sherman, *Causes of Police Behavior: The Current State of Quantitative Research*, 17 *J. RESEARCH CRIME & DELINQ.* 69 (1980); see also D. BLACK, *THE MANNERS AND CUSTOMS OF THE POLICE* (1980) (citing studies mentioned in text in describing police behavior).

13. Smith & Visher, *Street Level Justice: Situational Determinants of Police Arrest Decisions*, 29 *SOC. PROBS.* 167, 172 (1981).

14. Goldstein, *Improving Policing: A Problem-Oriented Approach*, 25 *CRIME & DELINQ.* 236 (1979).

justice actions.¹⁵ This causal link between groups' behavior and criminal justice actions is hard to determine merely by analyzing people treated differently for "natural" reasons. People who get longer sentences, for example, may commit more crimes upon release than those who get shorter sentences. Does that mean that longer sentences increase recidivism, or that more recidivistic people get longer sentences? Analysis of such natural differences cannot determine causation, but experiments can.

Experiments are especially powerful when they use random (i.e., lottery) assignment on large numbers of individuals. They are less powerful—and less properly called experiments—when they are conducted with a small number of people or without random assignment. Many problems in policing are unfortunately of this character. Assessing the effects of preventive patrol, for example, cannot be done at an individual level of analysis but only by neighborhood.¹⁶ There are not enough neighborhoods or resources available to most police agencies to select randomly from large numbers of neighborhoods. As a result, the reliability of any experiments conducted would be somewhat weakened.

Most police officer discretion (as opposed to managerial policy discretion), fortunately, is exercised in situational contexts with individuals. Random experiments are not only feasible for discovering the consequences of discretion, but are also more appropriate for studying discretion than for studying other police problems.

Two random experiments dealing with the discretion to arrest have been launched in the 1980's, both under my direction: the Minneapolis spouse assault experiment and the department store shoplifting experiment. The Minneapolis experiment was completed in 1983, and the department store experiment is scheduled for completion in 1985. Their design and findings illustrate the kind of knowledge that experiments can produce and the benefits and costs of having such knowledge.

The Minneapolis experiment¹⁷ assembled more than forty police officers who volunteered to give up their discretion for over a year in cases of simple domestic assault where the suspect was still present at the scene. Instead of deciding whether to arrest on the basis of their best judgment—or as studies in other cities suggest, on the basis of victim preference, suspect race, suspect demeanor, and size of the audience present¹⁸—the officers agreed to follow a random numbers formula (a lottery procedure) in deciding what to do. They had three options in the formula: arresting the suspect, "ordering" the

15. FEDERAL JUDICIAL CENTER ADVISORY COMM. ON EXPERIMENTATION IN THE LAW, *EXPERIMENTATION IN THE LAW* 17-18 (1981).

16. G. KELLING, T. PATE, D. DIECKMAN & C. BROWN, *THE KANSAS CITY PREVENTIVE PATROL EXPERIMENT* (1974).

17. See Sherman & Berk, *The Specific Deterrent Effects of Arrest for Domestic Assault*, 49 AM. SOC. REV. 261 (1984).

18. D. BLACK, *supra* note 12, at 85; R. Friedrich, *The Impact of Organizational, Individual and Situational Factors on Police Behavior* (1977) (unpublished Ph.D. dissertation, available at University of Michigan, Ann Arbor, Dept. of Political Science); Smith & Visser, *supra* note 13, at 168, 172-73.

suspect to leave for eight hours, or just talking (in a wide variety of ways) to the parties and leaving the scene. After each encounter, interviewers attempted to interview the victim every two weeks for a six-month follow-up period. The victims were asked whether subsequent violence or other problems had occurred since the police encounter. Official data were also tracked to see if any officers (whether or not they were in the experiment) had filled out a new offense report on or made an arrest in that household for domestic violence.

The results were striking. The arrested suspects, according to the victim interviews, were half as likely to have repeated their violence during the six-month follow-up as those who were ordered from the scene. Likewise, according to the official data, the arrested suspects were half as likely to have repeated their violence as the suspects who were only "advised." Both ways of measuring the consequences of police actions showed that arrest produced the lowest risk of new violence, even without other criminal justice sanctions. Since only four out of the 136 arrests produced any formal fine or other punishment imposed by a judge, the violence reduction can be attributed to the deterrent effect of arrest alone.¹⁹

The experiment was also interesting for what it did *not* find: that the effects of arrest were different for different kinds of people. It found that arrest had virtually the same deterrent effects for Native Americans, blacks, and whites, for employed and unemployed, and for suspects of different education and income levels. In statistical terms, the analysis found no interaction effects between police actions and suspect characteristics. Indeed, the only detectable interaction effect in the entire analysis involved an *officer* characteristic: whether the officer (according to the victim) took the time to listen to the victim's story before taking action. When the officers reportedly listened to the problems, the deterrent effect was enhanced; when they did not, the deterrent effect was greatly reduced.²⁰

The absence in the Minneapolis experiment of statistically detectable effects of interaction between police actions and suspect characteristics is not conclusive. The sample size of about 300, with only 150 complete interview sets, was too small, and the distribution of key variables too lopsided to provide a reasonable likelihood of detecting any differences in deterrability among kinds of individuals. In statistical terms, these methodological weaknesses create a risk of Type II error, or a bias in favor of the null hypothesis;²¹ the lack of adequate statistical power in a data set makes it possible to reject a hypothesis even when it should not be rejected and would not be rejected if sufficient power were available.²²

19. Sherman & Berk, *supra* note 17, at 270.

20. Sherman & Berk, *The Minneapolis Domestic Violence Experiment*, in POLICE FOUNDATION 1, 6 (1984).

21. H. BLALOCK, SOCIAL STATISTICS 112-16, 159 (1972).

22. *Id.* at 167.

Our department store shoplifting arrest experiment should avoid the problem of inadequate statistical power. Beginning in August of 1983, a major department store agreed to call the police to arrest every other person apprehended by store investigators for shoplifting. The experiment called for a continuation of this procedure until 1500 people had been apprehended. Both the arrested and the released shoplifters are to be followed up with department store and police records, and both will be interviewed six months to a year after the apprehension and asked how much more shoplifting they have done. The two groups will then be compared for the frequency and dollar value of their post-apprehension thefts.²³ The findings may help department stores determine the cost-effectiveness of calling the police and paying store investigators to go to court to testify against the shoplifters.

In addition, the findings may show something more. With a sample size of 1500, the department store experiment will have sufficient statistical power to detect interaction effects. The experiment should determine whether some kinds of people are more likely than others to be deterred from shoplifting as a result of being arrested. Differences between teenagers and adults, housewives and working women, churchgoers and nonchurchgoers, Catholics and Protestants, blacks and whites, and men and women should all be detectable in the analysis. Such differences may not exist. Because there are major differences among some such groups in their measured involvement in committing crime,²⁴ however, it seems unlikely that there would be no difference among them in deterrability. The commission of crime is by definition an absence of deterrence, or perhaps even a backfiring of attempts to deter. If labeling theory²⁵ is correct for at least some types of offenders, then the experiment may also show that arrest increases the frequency of shoplifting for such people.

If experiments in police discretion continue to create such knowledge we may soon have three classes of police-citizen encounters. The largest class would be the *ignorance* group, or those situations for which no experiments have yet been conducted. The second class would be the *equity* group, or those situations for which the available experiments show that one police action works best for everyone, regardless of personal characteristics. The third class of encounters would be the *differentiation* group, in which the available experiments show that different police actions are more or less effective for different kinds of people. How we view the policy implications of each of these three classes of encounters depends on the values and goals we hold for the police in a free society dedicated to equality under the law.

23. L. Sherman, *The Specific Deterrent Effects of Arrest: A Field Experiment* (1982) (unpublished proposal submitted to the National Institute of Justice).

24. Hindelang, *Race and Involvement in Common Law Personal Crimes*, 43 AM. SOC. REV. 93 (1978).

25. See the review of labeling theory in Sherman & Berk, *supra* note 17, at 261-62, 269.

IV

THREE PHILOSOPHICAL MODELS OF POLICE WORK

Professor Packer's classic analysis of the antinomy between the due process and crime control models of justice suggests the virtue of exploring value assumptions in public policy.²⁶ Discussions of police discretion rarely articulate such value premises, which nevertheless underlie many disagreements. Table 1 summarizes some important questions of value in police discretion and the answers to those questions implied by three different ways of thinking about discretion. These abstract models of analysis are all derived from various writings and discussions on these issues, so in one sense they have an empirical referent. They are nonetheless "ideal types" not necessarily found in this exact form in nature.

One model for thinking about the police is the formal, legalistic premise that police are bound by the letter of the law; they are mere ministers of the state charged with executing its edicts unthinkingly. This *ministerial* model contrasts sharply with the *police justice* model that police executives have articulated in recent years, according to which police are defined as highly trained experts capable of making complex decisions about how best to produce justice, just like prosecutors and judges. The concern with justice for its own sake differentiates this model from the *professional crime control* position, which uses doctors rather than lawyers for a model: the police should control disease (crime) the best way possible, rather than just dispense medicine (justice) regardless of its effectiveness.

The ministerial model originates in legislative conceptions of the police task. It is the basis for the various full enforcement laws and the penalties for an officer's failure to make an arrest when a crime is witnessed²⁷ and is the preference of one member of the ABF research group.²⁸ It is espoused by one commentator who has recently argued that police are morally obliged to enforce all the laws with full powers, if only to create pressure to reduce the scope of the criminal law.²⁹ For the most part, however, the ministerial model is a straw man. Police discretion has been so widely discussed for the past two decades that many people take it for granted, despite the fact that its exercise is an apparent violation of many statutes. Policymakers accept discretion uncritically, focusing only on the question of how to shape and control it.

Table 1 portrays the straw man in its full glory. The model values a virtual ban on decisions not to enforce the law if evidence legally sufficient for

26. H. PACKER, *THE LIMITS OF THE CRIMINAL SANCTION* 153-73 (1968).

27. *E.g.*, D.C. CODE ANN. § 4-142 (1981); *see also* H. GOLDSTEIN, *POLICING A FREE SOCIETY* 108 (1977) (commenting that most states require the police to enforce all laws and referring specifically to the D.C. statute). *But see* D. BLACK, *supra* note 12, at 90 (noting that police observed in 1966 in three cities, including Washington, made arrests in only 58% of the encounters with felony suspects and 44% of the encounters with misdemeanor suspects).

28. *See* Goldstein, *supra* note 11, at 586-87. *But cf.* W. LAFAVE, *supra* note 9, at 492-95 (recognizing that police use discretion and urging evaluation of how they do and should use it).

29. Interview with William Heffernan, John Jay College (March 1981).

enforcement is present. Like Packer's due process model,³⁰ its goal is justice and its presumption is the innocence of the suspect. Like Packer's crime control model,³¹ it is driven by a general preference aimed at deterrence, for the greatest possible likelihood of sanctioning each crime that occurs. It tolerates neither police judgments about a person's character nor any priorities related to scarce enforcement resources as a basis for deciding not to arrest. The law is the law, and that is it. First offenders, career criminals, sassy kids and polite grandfathers should all be treated equally by police according to the evidence available. Any judgments on these criteria should be made by the prosecutor, who is legally invested with discretionary power, not by the police. "Justice" in any sense is therefore removed from police hands and given to other officials to dispense.

TABLE 1
THREE PHILOSOPHICAL MODELS OF POLICE WORK

Value Questions	Model		
	Ministerial Justice	Police Justice	Professional Crime Control
1. Should police exercise discretion at all?	No	Yes	Yes
2. Should police seek justice for its own sake, rather than for crime control?	Yes	Yes	No
3. Should police presume guilt of most suspects?	No	Yes	Yes
4. Should the criminal sanction be applied to as much crime as possible?	Yes	Yes	No
5. Should police pursue specific, rather than general, deterrence?	No	No	Yes
6. Should police pursue crime control goals whatever the moral worth of the suspect?	Yes	No	Yes
7. Should scarce prosecution resources be focused only on the most deterrable offenders (or those most likely to be incapacitated)?	No	No	Yes

Police officers and administrators in the 1980's generally would not accept the ministerial model. Police thinkers ever since Vollmer³² have cultivated the notion of police work as a profession. There can be no professionalism, however, where there is no discretion to be guided by professional exper-

30. H. PACKER, *supra* note 26, at 164-67.

31. *See id.* at 158-60.

32. A. VOLLMER, *THE POLICE AND MODERN SOCIETY* (1936).

tise.³³ Acknowledging discretion is central to enhancing the prestige of the police, and courts have recognized officers' professional discretionary decisions.³⁴ The police justice model, however, has been more retributivist than utilitarian. As with Packer's crime control model, the police justice model is concerned with punishing criminals. The moral zeal of the police inclines them toward vengeance on behalf of victims and the society as a whole. Thus, advocates of the police justice model share the ministerial model's disregard of scarce criminal justice resources, faith in the wisdom of pushing arrest rates for all offenses as high as possible, and disregard for specific deterrence (or even such heresy as labeling theory)³⁵ considerations. Their conception of justice is primarily retribution or just deserts, but is based on the offender rather than on the offense. The police justice model parts company with the ministerial model on the individuation of vengeance (not crime control *per se*), which supports the officer's right to use his judgment of relative moral worth. Since most suspects presumably are guilty, and since police cannot possibly arrest them all, it is no less fitting for police than for judges to consider character in deciding what treatment a suspect "deserves." Insults to a police officer, like insults to a judge, are properly relevant facts to consider in deciding what action to take.

Curiously, this police justice model—which most police appear to espouse—is not a crime reduction model. It views policing much more as an end than as a means. Police arrest, harass, or move along some people and not others because those people merit such treatment in moral terms, not because such treatment makes society a better place or achieves the greatest good for the greatest number. Any deterrent effects of policing in this model, either specific or general, are merely by-products of justice as police define it.

A third conception of the goals of policing has been suggested by Herman Goldstein,³⁶ the American Bar Association project on criminal justice standards and goals,³⁷ and others.³⁸ This conception holds that police face a wide range of problems and should use the most effective means to solve them. Even if a problem includes violations of the law, the crime control model assumes that arrest and prosecution may not necessarily be the most desirable or effective way to deal with the problem. This model is pragmatic and utilitarian, using arrest only as a means to the end of crime control. There is no virtue to "justice" if it merely creates more crime.

The professional crime control model thus shares the police justice model's support for discretion and implicitly shares the working assumption

33. See Hughes, *Professions in THE PROFESSIONS IN AMERICA* 1, 2-3 (K. Lynn ed. 1965).

34. *E.g.*, *Foley v. Connelie*, 435 U.S. 291, 297-99 (1978).

35. "Labeling Theory" claims that the imposition of punishment enhances, rather than reduces, the likelihood of repeat offending. See, *e.g.*, E. LEMERT, *HUMAN DEVIANCE, SOCIAL PROBLEMS, AND SOCIAL CONTROL* (1972).

36. Goldstein, *supra* note 14, at 241-43.

37. AMERICAN BAR ASSOCIATION PROJECT ON STANDARDS FOR CRIMINAL JUSTICE, *STANDARDS RELATING TO THE URBAN POLICE FUNCTION* §§ 1.1(b), 2.2, 3.2 (1973).

38. Bayley & Bittner, *supra* note 3, at 35-36.

that most suspects are guilty (at least for the many minor offenses for which police use discretion). The two models differ on everything else. The crime control model is skeptical about the effectiveness of the criminal sanction in shaping behavior. It is more concerned with the future conduct of individual suspects than with the conduct of the hypothetical audience viewing police treatment of the suspect; that audience, however, is unlikely to learn how the suspect actually was treated.³⁹ The crime control model views individuation of treatment by "moral worth" as professional malpractice, just as different medical care according to doctors' judgments of patients' moral worth would be. The model does support individuation of treatment according to "what works." If criminal justice resources are scarce, it makes sense to save arrests for those who can be deterred or incapacitated. If an arrest has little chance of reducing individual risks of recidivism, then a crime control model does not require arrest.

There are admittedly more than three possible combinations of answers to the seven questions of value listed in Table 1. Some readers, for example, may prefer the police justice model without the presumption of guilt and moral worth. Yet, these three models may pose the most realistic policy options. The first is already the law de jure, the second the law de facto, and the third appears to have the strongest support from police experts.⁴⁰ It is not an idle exercise, then, to consider the implications of experiments in police discretion through the value lenses of these three models. Table 2 summarizes those implications.

TABLE 2
IMPLICATIONS OF KNOWLEDGE OF CONSEQUENCES OF POLICE
ACTIONS BY VALUE MODEL OF POLICE WORK

Model of Police Work	State of Knowledge		
	Ignorance	Equity	Differentiation
Ministerial Justice	Irrelevant	Irrelevant	Irrelevant
Police Justice	License for discrimination	Helpful, but not mandatory guide to action	Largely unusable, a threat to general deterrence; supportive of preferences based on moral worth
Professional Crime Control	Unacceptably dangerous	Mandatory guide to action	Helpful, perhaps mandatory guide to action except for constitutionally protected categories

39. Moreover, as a recent review of the deterrence literature argues, specific deterrent effects may comprise a large portion of what are presumed to be general deterrent effects, given the relatively small size and repeated sanctioning of the population of offenders. Lempert, *Organizing for Deterrence: Lessons from a Study of Child Support*, 16 *LAW & SOC'Y REV.* 513, 515 (1981-82).

40. See, e.g., *supra* notes 36-38 and accompanying text.

V

IMPLICATIONS OF IGNORANCE

Of the three classes of knowledge about police-citizen encounters,⁴¹ ignorance—an absence of experimental data on the consequences of police action—is the status quo and is likely to remain the largest category for some time to come. This state of affairs is not troublesome from the viewpoint of the ministerial model. It provides a convenient justification for the police justice model's discrimination according to moral worth. From the perspective of the professional crime control model, however, ignorance poses an ethically unacceptable risk of doing more harm than good.

The state of knowledge about consequences is irrelevant to the ministerial model: the law says what acts are crimes and all crimes should be followed by arrests until the legislature says otherwise. The ministerial model might grant that knowledge of consequences could be useful in guiding legislative decisions about how police should deal with various crime and order problems, but the model cannot incorporate knowledge into police decisions. Police are supposed to make only an evidentiary decision: is there probable cause to believe a person committed an offense sufficient to justify an arrest?

Ignorance is a convenient state of knowledge for the police justice model because it helps to legitimate police discrimination according to assessments of moral worth. Prosecutors and judges do not know the consequences of their decisions for either the specific deterrence of the offender or the general deterrence of the community, let alone the effects of such decisions on victims and witnesses. Nonetheless, they bear the mantle of professionalism to justify their making of such decisions. Their professionalism invests them with the right to determine what is just in a particular case, the right to make the punishment "fit" both the crime and the criminal. With no data to show that such retributivist actions may conflict with utilitarian considerations of crime control, the only embarrassment the lawyers suffer is from disparity in punishments administered within and between different courts. Police can avoid even that appearance of disparity by making low visibility decisions without a reviewable record.⁴² Without knowledge that would contradict their decisions, the police, as professional dispensers of justice, are justified in playing God.

The professional crime control model, in contrast, equates ignorance of police effects with ignorance of the effects of a new drug. Neither should be put into widespread practice without careful experimental testing. The fact that police discretion predates the rise of the scientific method seems irrelevant, for there is no "grandfather clause" in the obligation to do no harm. Leeching and ice baths predate scientific medicine, and they failed to survive scientific testing. Whether police are using similarly wrongheaded strategies can be answered only by research.

41. See *supra* p. 68.

42. See Goldstein, *supra* note 11, at 552-54.

Police interventions in citizens' lives are powerful treatments, potentially affecting life and death in both the short and long run. Like parenting, policing is a social intervention that seems so intuitive that anyone can do it well. Social interventions lack the mystery of chemical and surgical interventions. Yet, the commonsense view masks the complexity of the causal relationships involved.

To be sure, even in medicine there is debate about how much knowledge is required before a treatment should be used. Both police and doctors share the ethical obligation to try to do something about a problem, even if the something is merely likely, but not certain, to be helpful. It is this obligation that led one doctor to keep a young girl dying of encephalitis alive by giving her isoprinosine, an untested and unapproved drug.⁴³ It is this obligation also that led doctors to prescribe diethylstilbesterol (DES) to 900,000 pregnant women from 1955 to 1973, in order to lower their rate of miscarriages. Early, uncontrolled research had suggested the drug would have that benefit. Later, controlled research showed that the benefit was illusory and that children born of mothers taking the drug had higher rates of cancer.⁴⁴ The debate in medicine continues even though "[t]he randomized controlled clinical trial (RCT) has become the accepted standard for demonstrating the therapeutic effects in evaluating most forms of treatment."⁴⁵ Some doctors feel it is wrong to withhold potentially beneficial drugs from patients randomly assigned to a control group, especially where there are no known harmful side effects. Few doctors, however, would advocate using drugs that had not been tested by any kind of research.

The current state of ignorance in police discretion is worse than medical ignorance of the effects of untested drugs. Ignorance in police discretion allows dispensing totally untested treatments on the basis of race, age, sex, and demeanor. It may compound the discriminatory aspects of justice by producing more unintended long-term effects (like increased violence) within one racial group than within another. If, for example, the results of the Minneapolis spouse assault experiment may be generalized (which limitations in research design and sample size prohibit), the reported reluctance of police in some communities to arrest blacks in domestic assaults may have increased the risks of danger to black victims. If the police are to be professionals in crime control, they must not go on treating situations and people in ignorance of such effects.

43. *Doctors Debate Use of Controlled Studies to Test Effectiveness of New Treatments*, Wall St. J., Aug. 12, 1982, at 21, col. 4.

44. *Id.*

45. Spodick, *Randomized Controlled Clinical Trials: The Behavioral Case*, 247 J. AM. MED. A. 2258, 2258 (1982), quoted in *Doctors Debate Use of Controlled Studies to Test Effectiveness of New Treatments*, *supra* note 43.

VI

IMPLICATIONS OF EQUITY

Police research could, hypothetically, develop a body of knowledge consisting largely of equitable implications. Like the spouse assault experiment, other research could find that there is one treatment that works best on virtually everyone. This situation would also be similar to that in medicine, in which most treatments work for most people except for a small percentage of the population with adverse reactions to medicines (such as, for instance, those few people who are allergic to penicillin). One can imagine a set of equity findings: arrest works best for spouse assaults, mediation works best for bar fights, taking kids home works best for juvenile noise problems, but juvenile burglars should be diverted to pretrial restitution programs.

Once again, the ministerial model would interpret such findings as interesting for the legislatures and irrelevant to the police. If prosecuting of juvenile burglars only makes them commit offenses at a higher rate, then the prosecutor, not the police, should use that information.

The police justice model finds such information helpful, but hardly a binding guide to action. Minneapolis police officers who conducted the spouse assault experiment told the news media that the results of the research would not alter the way they handled domestic cases. "Every case is different. You can't generalize," went the familiar refrain.⁴⁶ This is not so much an antipositivist sentiment as an assertion of the right to retain discretion to vary their actions for reasons other than crime control. Some people deserve punishment and others may not, regardless of crime control. The information from such research may tip the balance in borderline cases, but it hardly changes the basic assumption with which police approach their cases.

The police justice model of police work will pose a major obstacle to immediate implementation of any new policies based on the results of research. The Minneapolis police chief has already issued an order calling for arrests in most cases of domestic assault. The prior practice had been to arrest only in rare circumstances when someone really "deserved" it. Moving an entire patrol force to change its discretion patterns may require much more than an order.

If police officers were socialized into accepting the professional crime control model of police work, however, they would readily accept the findings of such experiments as an obligatory guide to practice. If cancer treatment experiments consistently showed that chemotherapy produced higher survival rates than radiation for all kinds of people, oncologists would probably feel bound to prescribe chemotherapy. They could be sued for malpractice if they tried to vary the recommended treatment according to the race, age, sex, or "moral worth" of the patient (unless there was an interaction effect between those variables and the chemotherapy treatment). Similarly, as police "mal-

46. Meeting of domestic violence project police officers, Spring Hill Conference Center, Wayzata, Minn. (April 6, 1984).

practice" suits increase in number around the United States, it is conceivable that expert witnesses, citing the findings of research, could convince juries that police committed malpractice in failing to be guided by those findings. There have already been many suits against police for failing to arrest domestic assailants;⁴⁷ experimental evidence supporting the crime control benefits of making such arrests will make such cases even stronger. Indeed, one reason police may prefer the ignorance of the police justice model is that it reduces their vulnerability to such lawsuits.

Any police officer committed to the goal of crime control through specific deterrence would find it unconscionable to fail to follow consistent research findings. One may argue about the number of experiments in different settings needed to make the evidence persuasive, but once the evidence becomes persuasive, police should no sooner disregard it than they should help inmates escape from prison.

In many ways, equity is the least troublesome scenario. With equitable policy implications of research, police can control crime without running afoul of constitutional problems. The equity scenario, however, seems also the least likely to develop in the long run, despite the convenience of the medical analogy. The chemical makeup of the human body appears to be more homogeneous than the socio-genetic makeup of human behavior. Just as social behavior apparently varies more widely than the internal functioning of the human organism, the treatments necessary to affect that behavior may have to vary more as well in order to be effective.

VII

IMPLICATIONS OF DIFFERENTIATION

Perhaps the more likely scenario for the development of experiments in police discretion is an accumulation of findings of differentiation. If experiments are done with large enough sample sizes, the chances of finding different actions to be effective for different types of people seem to be rather good. One can imagine this set of findings: arrest works to deter employed male bar fighters under age thirty, but the best treatment for employed males over thirty is driving them away from the scene. Indeed, the complexity of such findings suggests an approach that is almost Orwellian in appearance. Nevertheless, that approach would put science to its most effective use in crime control.

The findings of a police actions experiment employing a large size sample can easily be transformed into software for microcomputers. Police in many cities already have mainframe computer terminals in their cars. By whatever hardware arrangements seem appropriate, police could easily use the software to obtain an advisory opinion on the basis of past research about what to do

47. See, e.g., *Thurman v. City of Torrington*, 595 F. Supp. 1521 (D. Conn. 1984) (§ 1983 action alleging violation of equal protection in that arrest decisions made by police gave less protection to women abused by husbands and boyfriends than to other women).

with a particular suspect. By asking the suspect questions that appear on the computer screen and punching in the answers, the officer could fill in the numbers to the preprogrammed equations predicting the likely post-encounter effects of police actions on suspect behavior. The equations could also include cost data to weigh against benefits, so that even if benefits could be predicted from one course of action, the cost-benefit ratio might lead the equation to flash on the screen a recommendation against taking that action.

The concept of computer-aided police discretion is a very likely result of the department store shoplifting experiment. A store security office offers ideal conditions for gathering information from the suspect (e.g., age, education, church attendance, community activities), entering the data, and reading a recommendation to arrest or not calculated from an equation based on the store's labor costs of prosecution. Given the wide discretion to arrest exercised in many other areas of this country's booming private policing industry, the computer-aided approach will initially be more common in that sector than among public police. The reason follows our three models: business is not in the business of meting out justice. The business of business is making money, and to avoid losing money its security directors pursue a clear goal of crime control.

To be sure, many store security officers share the police justice model and the urge to punish the offenders they catch. They describe their own use of discretion in the familiar terms of public police: sympathy for the elderly shoplifter and lenience for the very young, anger at the resistant suspect and a sense of challenge engendered by the "professional shoplifter."⁴⁸ Their arrest decisions are often viewed as a dispensing of justice. In contrast to public police officers, however, private security officers cannot afford such arrogation of power if their management disallows it. Without civil service or union contract protection, security officers can be fired at will. If management says to use a computer program to make arrest decisions, the officers will probably do so.⁴⁹

The ministerial model again finds all these developments irrelevant. The police justice model would find these developments most troublesome, since they seem to undermine general deterrence. Letting people go just because those individuals cannot be deterred would not make sense if it produced a net increase in total crime. If the perceived likelihood of arrest for a given offense were to decline with any of these decisionmaking aids, it could encourage previously deterred nonoffenders to begin offending. Given the primacy of general over specific deterrence in the police justice model, such a consequence would be unacceptable.

Such consequences would also be unacceptable to the professional crime

48. Interviews with 90 store security officers in a major midwestern department store (August, 1983).

49. Curiously, the ministerial model has no objection to this development in private policing. Since private police have the legal status of citizens deciding whether to report a crime, there is no legal impropriety to the discretion. Since there generally is no obligation to report crimes, citizens are free to use whatever criteria they choose to make that decision.

control model, but that model would impose a higher standard of proof before reaching a conclusion. The general deterrent effects of arrest decisions made for specific deterrence purposes is a question amenable to research. Randomly chosen department stores using such software could, for example, be monitored, and their loss rates could be compared with the loss rates of those stores assigned to a control group not using the software. If there were no differences, the hypothesized loss of general deterrence claim could be disproved. Since arrest-to-crime ratios are already so low for so many crimes, it seems unlikely (but possible) that restructuring arrest decisions according to research on specific deterrence would reduce whatever general deterrence effects the arrest now produces.

The professional crime control model welcomes any refinement of knowledge about what works. Differentiation of police actions by personal characteristics would make the use of police "tools" of action more precise and effective. Nevertheless, a professional crime control model cannot tolerate decisions based on constitutionally forbidden considerations. Under the fourteenth amendment, an equation for deciding whether public police should make an arrest must exclude race, national origin, and religion.⁵⁰ Private police probably would not be barred from using an equation with those variables, at least not under the current case law defining their legal status.⁵¹ Private police, however, would face an ethical problem in officially endorsing the use of factors that our nation has long deemed legally irrelevant.

No matter what model of police work is used, the inequity of police actions based on differential deterrability is ethically questionable. Even without the constitutionally protected factors in an equation, it seems unfair to punish people on the basis of their employment status, their frequency of church attendance, their marital status or home situation, or whether their parents were divorced. Many of these factors may even correlate with the constitutionally protected factors.

The unfairness may be worse. If the control theory of crime⁵² is correct, and if it interacts with sanctions so that more socially bonded people are more deterrable, then the unfairness will seem great indeed. Those who would be *most* highly recommended for arrest would be those most involved with church, family, community organizations, and the labor force. Those *least* likely to be recommended for arrest would be the marginal characters of low "moral worth" that police love to punish: the unemployed, unmarried, nonchurchgoing riffraff. One could characterize arresting respectable persons as doing them a favor if the arrest really deters them from further crime. It is unlikely, however, that they would be grateful for the favor.

The question of differentiation as social policy is increasingly common as scientific knowledge increases. The different pension benefits for males and

50. U.S. CONST. amend. XIV.; 42 U.S.C. § 1983 (1982).

51. See *Burdeau v. McDowell*, 256 U.S. 465 (1921).

52. See generally T. HIRSCHI, *CAUSES OF DELINQUENCY* 16-34 (1969) (the control theory sees crime as the result of a lack of a bond between an individual and society).

females based on different life expectancies is one example in which research suggested a policy with disparate effects, but the courts would not accept it.⁵³ Genetic screening for predispositions to occupational and other diseases is another example.⁵⁴ Such screening offers potentially enormous benefits by giving people early warnings of their predispositions to cardiovascular disease, cancer, and other killers. It also raises ethical questions about the uses to which such information should be put: whether employers should have a right to obtain genetic data related to occupational diseases, for example. Such developments may prompt attacks on science for learning too much, producing dangerous knowledge that allows managers and other powerful people to manipulate the lives of the weak.

An additional problem is that science will still have learned too little. At best, experiments can only point out *probabilities*—not *certainities*—of different individual reactions to police actions. The problem of false positives (people who are shown by the data probably to be deterrable but turn out not to be) and false negatives (people who are probably not deterrable but who actually could be) compound the inequities even more. Prediction in any individual case can always be in error. This unavoidable statistical property means that people would not only be arrested on the basis of their personal characteristics, but also on the basis of an imperfect prediction about the effect of those characteristics in shaping their reactions to arrest or nonarrest. This raises further questions about what level of false positives should be acceptable: none? five percent? fifteen? twenty-five? Each of these levels comprises a different value judgment about how much harm should be inflicted in order to accomplish good.

In that sense, the people in the pool of differentially treated suspects are responsible for putting themselves at risk of such treatment. The formulas did not (apparently) commit the offenses. These people are already vulnerable to arbitrariness and caprice. The question is whether experimental results would replace capriciousness with something better, or with dangerous knowledge that would create something even worse.

VIII

CONCLUSION

Do experiments in police discretion produce dangerous knowledge when they show differential effects of arrest, or do they provide a boon that allows police to avoid wasting resources and doing more harm than good? The experiments will probably provide both. Whether the benefits outweigh the harms will depend on exactly how such knowledge is used.

It is possible that the results of the experiments will be completely ignored. Despite the high level of press attention accompanying the publica-

53. See *Arizona Governing Comm. v. Norris*, 103 S. Ct. 3492 (1983).

54. See *New Gene Probes May Permit Early Predictions of Disease*, Wall St. J., Dec. 23, 1983, at 11, col. 1.

tion of these experiments, it seems likely that in the short run police will pay little attention to them. Those findings that are not too complex to be remembered easily—especially the equitable ones—will probably become part of the police academy and college curricula; in a generation police officers may accept the findings as common wisdom. Those findings that are too complex to remember without the help of a computer terminal and a software program will live or die depending on the policy decisions that are made about such technological approaches to police work. Private security may move slowly in that direction over the next decade, public police not until the next century.

Scholars conducting experiments in police discretion must face several questions now: Do we want these developments to occur? Should we help them to develop? And if they do, should we try to control their direction?

There are no morally satisfactory answers to these questions, for they present a three-way dilemma to anyone who adheres to a professional crime control model for police action. To do nothing to relieve the state of ignorance about the consequences of police action is unsatisfactory for two reasons. First, it helps justify the continuing discrimination in the exercise of police discretion. (Arguably there are ways to cure that problem other than creating knowledge about consequences, but no one has demonstrated any.) Second, a more compelling reason why the status quo is unsatisfactory is that police may be doing so much harm unknowingly. "The only alternative to experimenting with people is to fool around with people"⁵⁵ with unknown effect.

It would be nice if the experiments would all produce equitable policy implications, showing no interaction effect between police action and type of person affected. Scientists could even create this optimal solution by not searching for interaction effects. A conscious decision not to test for interactions, however, would be almost as dangerous as not testing at all. Suppose that arresting Hispanic men for spouse assault doubled the rate at which they subsequently committed homicide, while it cut the rate in half for black and Anglo males. It would be irresponsible for a scientist to say that arrest works best for everyone without at least looking for any such backfiring effects within subgroups.

The discovery of such findings (if they occur) is unsatisfactory because they imply a formal policy of unequal treatment according to personal characteristics. Unlike the current informal, low-visibility police practices which do the same thing, a formal policy makes a symbolic statement legitimating differential treatment as a principle; the reason is irrelevant.

The dilemma, however, cannot be avoided. Society must choose between two unsatisfactory practices. One is not doing such experiments, thereby helping to perpetuate particularistic assessments of moral worth as the prin-

55. PANEL ON RESEARCH ON REHABILITATIVE TECHNIQUES, NAT'L RESEARCH COUNCIL, *THE REHABILITATION OF CRIMINAL OFFENDERS* 71 (1979) (quoting Frederick Mosteller).

ciple of differentiation in police discretion. Whether or not police are ignorant of the consequences, they will continue to arrest some people and not others who are legally vulnerable to arrest. If only because of scarce resources, there will continue to be differential treatment of legally comparable individuals. Without any scientific data as an alternative, the whim and prejudice of an officer will guide that differentiation.

The other choice, and perhaps the preferable one, is to generate as much knowledge about differential effects as possible. This knowledge could contend with the prejudices now guiding police discretion and perhaps redirect them. It can have that effect immediately in private policing. If constitutionally protected classes are left out of equations for anything except violent crime (for which it might be worth a more careful balancing analysis of a victim's right to equal *impact* of law compared to a suspect's right to equal *treatment*), then there is little that is legally controversial about such uses of experimentally derived knowledge.

Some might argue that the ministerial approach is still the best solution. Create full enforcement and there will be no discrimination to counteract or replace. The law will be truly just and equitable in its application. The problem is that the law may not be equitable in its consequences. If the tools of science can demonstrate how law may be used more effectively to create a better society for all, can we not be bold enough to rethink the legal basis for the ministerial model? Or can we not be realistic enough to accept differentiation as inevitable, and that the only question is how differential treatment is to be guided?

In either case, the implication for the role of science is clear. The experiments should be done and the results put to the best use possible, despite the tragic consequences and inequities of treatment they may produce. Compared to the status quo, this choice provides a much better public policy.

