

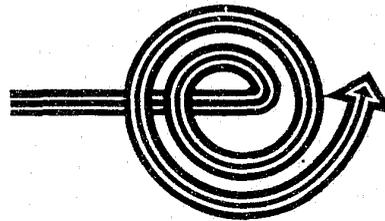


U. S. Department of Justice  
National Institute of Justice



# The First National Conference on Criminal Justice Evaluation

## Selected Papers



82918-  
82926

## About the National Institute of Justice

The National Institute of Justice is a research, development, and evaluation center within the U. S. Department of Justice. Established in 1979 by the Justice System Improvement Act, NIJ builds upon the foundation laid by the former National Institute of Law Enforcement and Criminal Justice, the first major Federal research program on crime and justice.

Carrying out the mandate assigned by Congress, the National Institute of Justice:

- Sponsors research and development to improve and strengthen the criminal justice system and related civil justice aspects, with a balanced program of basic and applied research.
- Evaluates the effectiveness of federally-funded justice improvement programs and identifies programs that promise to be successful if continued or repeated.
- Tests and demonstrates new and improved approaches to strengthen the justice system, and recommends actions that can be taken by Federal, State, and local governments and private organizations and individuals to achieve this goal.
- Disseminates information from research, demonstrations, evaluations, and special programs to Federal, State and local governments; and serves as an international clearinghouse of justice information.
- Trains criminal justice practitioners in research and evaluation findings, and assists the research community through fellowships and special seminars.

Authority for administering the Institute and awarding grants, contracts, and cooperative agreements is vested in the NIJ Director, in consultation with a 21-member Advisory Board. The Board recommends policies and priorities and advises on peer review procedures.

NIJ is authorized to support research and experimentation dealing with the full range of criminal justice issues and related civil justice matters. A portion of its resources goes to support work on these long-range priorities:

- Correlates of crime and determinants of criminal behavior
- Violent crime and the violent offender
- Community crime prevention
- Career criminals and habitual offenders
- Utilization and deployment of police resources
- Pretrial process: consistency, fairness, and delay reduction
- Sentencing
- Rehabilitation
- Deterrence
- Performance standards and measures for criminal justice

Reports of NIJ-sponsored studies are reviewed by Institute officials and staff. The views of outside experts knowledgeable in the report's subject area are also obtained. Publication indicates that the report meets the Institute's standards of quality, but it signifies no endorsement of conclusions or recommendations.

James L. Underwood  
*Acting Director*

U.S. Department of Justice  
National Institute of Justice

---

# The First National Conference on Criminal Justice Evaluation

## Selected Papers

Joel H. Garner  
and  
Victoria Jaycox

Editors

December 1981

U.S. Department of Justice  
National Institute of Justice

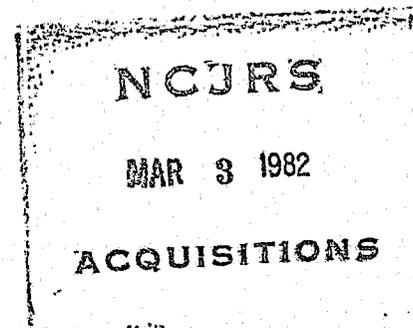
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official position or policies of the National Institute of Justice.

Permission to reproduce this copyrighted material has been granted by

Public Domain/N I J  
U.S. Dept. of Justice

to the National Criminal Justice Reference Service (NCJRS).

Further reproduction outside of the NCJRS system requires permission of the copyright owner.



**National Institute of Justice**  
James L. Underwood, *Acting Director*

These papers were among those presented at the First National Conference on Criminal Justice Evaluation held in Washington, D. C., February 22-24, 1977. Points of view are those of the authors and do not necessarily represent the position of the National Institute of Justice or the U. S. Department of Justice.

## FOREWORD

The first National Conference on Criminal Justice Evaluation was held in Washington, D.C., in February 1977. Over 1,400 individuals attended the 33 separate panels and presented, critiqued, or listened to 179 evaluation reports. While the Conference represented a culmination of several years of research and evaluation in criminal justice, most of the papers reported on work of quite recent vintage, demonstrating that the call for more and better evaluation in LEAA's 1973 reauthorizing legislation had not gone unheeded. The National Conference on Criminal Justice Evaluation revealed through the volume and the wealth of completed projects that criminal justice evaluation is a healthy and viable discipline.

The Conference was sponsored by the National Institute of Law Enforcement and Criminal Justice but the full participation of researchers and practitioners from other LEAA and DOJ offices, State and local government agencies, academic institutions, and private research firms truly made the Conference a product of the criminal justice community.

This volume of selected papers covers the full spectrum of topics and participants from the Conference. Some of their findings have already been incorporated into the knowledge base for improving criminal justice in the United States; some papers are truly innovative and provide significant methodological advances for future research and evaluation; other papers remain controversial and outside the accepted wisdom of criminal justice practices. This spectrum of perspective is what we hoped for in the Conference and what we at the National Institute will continue to encourage.

A complete list of the papers presented at the Conference is provided in the appendix.

CONTENTS

	<u>Page</u>
Foreword . . . . .	i
Keynote Address: Some Preliminary Reflections on Evaluation in Criminal Justice -- Sir Leon Radzinowicz . . . . .	1
SECTION I. POLICE EVALUATION	
[ An Evaluation of the Open Garage Door Burglary Program -- Ron Pennington . . . . .	53995 . . . . . 13
[ Police Roadblocks to Decriminalization: A Multiple-Time Series Analysis of Law Enforcement's Response to Changes in Public Drunkenness Statutes -- David E. Aaronson, Thomas Dienes, and Michael C. Musheno . . . . .	82919 31
[ The Challenge of Patrol Performance Evaluation in the Study of Policewomen -- Joyce L. Sichel . . . . .	44587 . . . . . 45
[ Patrol by Helicopter: An Evaluation -- Dr. John F. Schnelle and Robert Kirchner . . . . .	82920 . . . . . 59
SECTION II. EVALUATION IN THE COURTS	
[ Measuring the Effectiveness of the Bail Bond System as an Assurance of Trial Appearance -- Helen Reynolds . . . . .	53743 . . . . . 77
[ An Alternative to Court: An Evaluation of the Orange County (Florida) Bar Association Citizen Dispute Settlement Project -- Ross F. Conner and Ray Surette . . . . .	82921 . . . . . 91
[ Recidivism Rates of Diverted Juvenile Offenders -- Suzanne B. Lincoln, Kathie S. Teilman, Malcolm W. Klein, and Susan Labin . . . . .	53848 . . . . . 103
[ Voluntary Pre-Trial Diversion and the Question of Compliance: A Preliminary Evaluation -- Diane L. Gottheil . . . . .	53855 . . . . . 111
[ Prosecutorial Policy: Impact and Implementation -- Joan E. Jacoby . . . . .	53740 . . . . . 121

[ The Potential Impact of Information Systems in Evaluating Prosecution and Court Policy  
 -- Sidney H. Brounstein . . . . . 53246 . . . . . 137

SECTION III. EVALUATION IN CORRECTIONAL PROGRAMS

[ An Examination of Intensive Supervision as a Treatment Strategy for Probationers  
 -- Joseph H. Sasfy . . . . . 54737 . . . . . 155

[ Objective Selection Techniques as a Logical Prerequisite to the Evaluation of Work Release Programs  
 -- Douglas E. Scoven, Duane E. Brookhart, and J. B. Ruark . . . . . 53732 . . . . . 167

[ The Actualization and Impact of Team Classification in State Correctional Institutions  
 -- John R. Hepburn and Celesta A. Taylor . . . . . 53997 . . . . . 175

[ Residential Corrections Program in Minnesota: An Evaluation Report  
 -- Michael J. McMahon . . . . . 60207 . . . . . 193

[ Salt Lake City: Prototype Evaluation of Des Moines Replication  
 -- William M. Rhodes and Steven T. Seitz . . . . . 82922 . . . . . 205

SECTION IV. EVALUATING COMMUNITY PROGRAMS

[ Evaluating Citizen Crime Prevention Programs  
 -- Robert K. Yin . . . . . 53927 . . . . . 225

[ An Evaluation of the Seattle Community Crime Prevention Program  
 -- Kenneth E. Mathews, Jr. . . . . 82923 . . . . . 239

[ Atlanta's High Impact Anti-Crime Program: New Directions in Evaluating Project Performance  
 -- Samit Roy . . . . . 82924 . . . . . 247

[ Methadone Programs and Crime Reduction: A Comparison of New York and California Patients  
 -- Dale K. Sechrest . . . . . 62757 . . . . . 259

[ Street Layout and Residential Burglary  
 -- Julia Nutter, Douglas W. Frisbie, and Carol Bevis . . . . . 53724 . . . . . 269

SECTION V. THE METHODS OF EVALUATION

[ Recidivism and Research Design: Limitations of Experimental Control Research  
 -- Robert Martinson and Judith Wilks . . . . . 82925 . . . . . 279

[ The Three Faces of Evaluation: What Can Be Expected To Work  
 -- Herbert C. Quay . . . . . 44494 . . . . . 292

	<u>Page</u>
[ Measurement of Crime Pattern Modification: An Example, Gun Control -- Stuart Jay Deutsch . . . . .	5 3999 . . . . . 303
[ Experimental Design and Causal Analysis in Criminal Justice Evaluation -- James G. Heilman and Marshall T. Miller . . . . .	53929 . . . . . 315
[ Determinants of the Seriousness of Criminal Activity: The Misdemeanor-Felony Distinction -- Ann Dryden Witte and Peter Schmidt . . . . .	82906 . . . . . 325
[ Management by Evaluation: Putting Research to Work -- Dale Shears . . . . .	74173 . . . . . 343
[ Developmental History as a Method of Organizational Evaluation -- Jackwell Susman . . . . .	53164 . . . . . 351
[ Criminal Justice Information Needs of Illinois Local Government Officials -- S. Burkett Milner and Barbara E. Ferrara . . . . .	82267 . . . . . 361

APPENDIX

KEYNOTE ADDRESS: SOME PRELIMINARY REFLECTIONS ON EVALUATION  
IN CRIMINAL MATTERS

Sir Leon Radzinowicz  
Cambridge, England

Distinguished Company,

I am greatly honored to have been invited to join you on so important an occasion and to have been asked to address you. I am also delighted to acknowledge the important part played in the production of this paper by my close collaborator, Miss Joan F. S. King, Senior Assistant in Research in the Cambridge Institute of Criminology.

This is an impressive gathering. At least 150 reports; 34 panels; more than 1,000 participants. What strikes me especially is its remarkable geographical spread: literally from all parts of this great Republic. The exceptionally wide range of topics and interests--all focused on the central theme of evaluation. And last, but not least, we have the existing combination of well-established criminological and penological authorities with numerous young men and women all embarking upon the thorny path of criminological research and reassessment. Nor can I ignore to emphasize the sober, critical, and yet not purely negative, bent of virtually all the papers submitted to the Conference.

In view of the particular occasion and the accumulated wealth of material, you cannot expect from me more than a sharing with you of some preliminary reflections on evaluation in criminal matters against an historical and comparative background. That is what I shall endeavor to do within the next half hour.

To look around this Conference is to see, on all sides, that blending of curiosity, social conscience, and business acumen that is so pre-eminently American. To read the reports submitted to it is to see, as on a map, the directions in which the process of evaluation has been moving. It can indeed be said that there remains no part of the system of criminal justice that escapes this probing scrutiny.

Evaluation has been applied to all aspects of the work of the police from their work on the streets and the earliest tentative checking of possible suspects, to the evidence they may give in court. It has invaded the darker recesses of plea bargaining, and brought to light the implications of refusing bail. It has penetrated some of the reticence that once surrounded sentencing and has even been knocking upon the closed doors of that last stronghold of privacy, the jury room. It has left a devastating trail all through the penal system, from probation via the penal institutions to parole. It has been applied not only to processes but to persons: the policeman and the probation

officer are finding themselves and their work as much subject to evaluation as the system within which they function. Even the anonymous citizen has been pressed to reveal his attitudes to criminal law, the extent to which he has committed or suffered crime, the way he would rate the gravity of various crimes, and his views on the workings of the criminal justice system. Last of all, evaluators themselves have been coming under the microscope, their research scrupulously analyzed and reworked, themselves evaluated as ruthlessly as they have evaluated others.

It is tempting, indeed almost natural, to think of this as a purely modern phenomenon: In one sense it is. But in another it owes much to its ancestors, some of whom I think it well, however briefly, to recall.

That sturdy old iconoclast, Jeremy Bentham, made no bones about his touchstone when faced with established institutions, however venerable: "What use are they?" If they failed to pass that test, they were out, as far as he was concerned--and in many cases as far as posterity has been concerned as well. They were also out if they appeared to him to be manifestly unjust. In his old age, Lord Brougham said of Bentham, "The age of Law Reform and the age of Jeremy Bentham are one and the same . . . . He it was who first made the mighty step of trying the whole provisions of our jurisprudence by the test of expediency . . . their adaptation to the circumstances of society, to the wants of men and the promotion of human happiness . . . . None ever before Mr. Bentham took in the whole departments of legislation. None before him can be said to have treated it as a science, and by so treating, made it one."

Bentham's scrutiny extended to all branches of criminal justice--the substance of criminal law, the ramifications of procedure, the nature, degrees, and varieties of punishment. Subsequent schools of criminal law and criminology have laid varying emphasis on different aspects of this very wide perspective.

The classical school of criminal law, whose influence pervaded Europe and invaded the United States in the nineteenth century, was primarily concerned with the logic of justice, proportion, and procedures. It rested upon a series of assumptions, idealistic but, alas, unproven. The classicists believed they could devise a just, rational, predictable system of procedures and penalties which would maintain itself by its manifest virtues. They believed they could back this up by virtual certainty of detection: a prerequisite if the system was to be either just or effective. And they believed that the combination of certainty with proportion and moderation would deter both offenders who suffered punishment and any others who might be tempted.

Once a good run of criminal statistics had been established, the first tool of empirical, as distinct from philosophical, evaluation was in men's hands. And it undermined two of the basic classical assumptions. In England, for example, penalties had been brought into more reasonable proportion to offenses, and the police had begun to be organized into disciplined and efficient forces. Yet, Edwin Chadwick, one of the architects of police reform, was also one of the first to point out the continuing impunity among persistent criminals, and the failures of the penal system to prevent recidivism even among those who were caught. A system of deterrence by certainty of detection and punishment sounded fine on paper, but it was not working out in practice.

The positivists, when they came to the fore toward the end of the nineteenth century, were quick to seize upon these failures, as well as upon the inhuman rigidities associated with the classical insistence upon proportion in punishment. To demonstrate its bankruptcy, they quoted statistics of recidivism not only from England but from Italy, France, and Germany. They stressed experiment and observation as the keys to evaluation. Yet they, too, had their ideological assumptions and, given a free hand, would have made them the basis for whole systems of criminal justice. They assumed that the battle for procedural protections had been fought and won once and for all; that fairness in prosecution and punishment could be taken for granted. They assumed that a system of criminal justice could be made more humane and effective by classifying criminals on the basis of their propensities, their "dangerousness," and that this, rather than the crime before the court, should decide the measures to be used in dealing with them. It was the positivists who elevated the indeterminate sentence to such a lofty pedestal, seeing it as the ideal means of protecting society and, where possible, reforming the criminal.

Meanwhile, the Marxist criminologists were protesting that it was impossible to achieve either justice or protection within a capitalist society, whatever the means that might be used. To them, evaluation of a system of criminal justice was subsidiary to evaluation of the kind of society it was designed to support. At that stage, early in the present century, they had one great advantage. There was as yet no existing regime which had even attempted to mold itself on the precepts of Marx. Like the early enthusiastic classicists, pursuing their liberal dream, they lived in a glass house, which, being purely ideological, offered no targets for stones.

The capitalist countries, encumbered with real systems and the real defects and criticisms to which that exposed them, enjoyed no such advantage. They were in many ways, strongly self-critical. It was English investigators who so ruthlessly analyzed and condemned the English prison system at the end of the nineteenth century. It was American sociologists who explored and denounced, in turn, the crimes of big business, the corruptions of police, the festering of cities. But they nevertheless rejected the wholesale Marxist condemnations, convinced that it was possible, for example, to frame laws which would bring to book the wealthiest predators, to check as well as to expose police corruption, to counteract the evils of city slums. Standards of criminal justice could be raised with rising standards of society.

Neither the assumptions of the Marxists or of their opponents have survived subsequent events and subsequent scrutiny. Criminologists have again proved themselves highly efficient undertakers, burying many of these hopes under mounds of statistics. There have been the revelations of the prevalence of hidden crime, of the obduracy of prisons in the face of attempts to transform or redirect them, of sickening recurrences of corruption among those entrusted with enforcing the law. And there has been lack of evidence that the numerous experiments directed at reforming offenders have had any significant impact upon general rates of recidivism. Over the last 20 years this process of systematic disillusionment has deepened and widened, engulfing all aspects of criminal justice and, into the bargain, a great deal of earlier research into their effectiveness.

But we must preserve a sense of proportion. To quote a felicitous remark by Professor Arnold S. Trebach: "We seem to have too many critics and too few

play writers." We all think we know what is meant by evaluation--otherwise we should not be here. Yet this enigmatic and complex concept eludes any simple definition. I have so often quoted a certain cynical comment about criminal statistics that I have come to believe I invented it myself. I cannot resist adapting it to our topic today: "Evaluation is like a bikini. What it reveals is suggestive. But what it hides is vital." On top of that, the subject becomes ever more sophisticated. Nothing would seem more definitive than death but it too has now become difficult to evaluate. We are offered not one but half-a-dozen competing criteria. Perhaps we should count ourselves fortunate that at last there must come a point where everyone agrees that the poor fellow is dead.

The two major senses of the word evaluation are not always clearly distinguished. It is not always realized that evaluation in one sense usually implies, however unconsciously, evaluation in the other.

Both the Oxford Dictionary and Webster's give priority to the more mundane of the two kinds of meaning. To evaluate is "to work out the value of something" and value is "a fair equivalent or return--material or monetary worth." In other words, to evaluate is to find out whether we are getting our money's worth in what may be seen as objective terms. This is very much the function sense of "cost benefit analysis" and stresses the idea of finding objective facts as means of measurement. Inevitably we seize upon the most accessible and apparently objective index of achievement--such as rates of detection by the police, or of recidivism among those who have been through the penal system--only to find that, in practice, it is as full of distortions and loopholes as a broken net . . . .

But Webster also defines the word evaluate as "to determine the significance and worth of a thing, usually by careful appraisal or study." The Oxford Dictionary also defines "value," in the ethical sense as "that which is worthy of esteem for its own sake, that which has intrinsic worth." A classical example of this kind of evaluation is the well-known statement made by the young Winston Churchill when Home Secretary in 1910. May I quote it again:

"The mood and temper of the public in regard to the treatment of crime and criminals is one of the most unfailing tests of the civilisation of any country. A calm, dispassionate recognition of the rights of the accused, and even the convicted, criminal against the State - a constant heart searching by all charged with the duty of punishment - a desire and eagerness to rehabilitate in the world of industry those who have paid their due in the coinage of punishment; tireless efforts towards the discovery of curative and regenerative processes; unfailing faith that there is a treasure, if you can only find it, in the heart of every man."

The second meaning is subtler, but no less important. To evaluate is "to determine the significance or worth of a thing," and value is something "worthy of esteem for its own sake," something of intrinsic worth. These are subjective judgments, but they refer to the very basis of our evaluations; the things we think it worthwhile to measure and to compare, and our decisions about the criteria of "success," "effectiveness," or "failure." Their

influence upon what we look for, what we find, and what we do with our findings, is often taken for granted to the point of being overlooked.

Not only the broad themes of evaluation, but the ways that they are approached, the aspects examined, the kinds of questions asked and answers found, are heavily influenced by evaluative judgments. You can evaluate a scheme in many different ways: in terms of whether it is workable; of whether it has any effect in reducing crime or recidivism; of whether it is being carried out as its promoters intended; of whether it achieves their declared objectives; of whether it has unwanted side effects; of what impact it has on other parts of the system of criminal justice; of its acceptability to public opinion.

Of course, it is not only evaluation that is a complex matter. Criminal justice is even more so, multiple in its nature, its purpose, its functions. It is not like a business, with an easily measurable material objective. It is at least as complicated, its elements as intangible, as for example, a system of education. And it raises equally strong feelings, both among those who want to maintain traditional systems and those who want to revolutionize them.

Moreover, evaluation in criminal justice, as in education, is subject to a mass of technical problems: the extent of hidden crime; the limited reliability of criminal statistics; the varying standards and criteria of failure in parole or probation; the constant change in social and economic situations; the shifting objectives, practices, and personnel in the execution of projects, and in those to whom they are directed; the impact of research itself on the processes and attitudes it seeks to examine.

Then there are the pressures of our situation today. Running all through the reports of this conference are the recurring themes of rising crime and growing costs. And there is, alas, no reason in sight to expect that either will decline. That brings a mounting urgency to the theme of evaluation. We want to know what we are getting for our money in the sphere of criminal justice. But cost-benefit considerations apply also to evaluative studies themselves. Certainly we spend huge sums on our police, our courts, our prisons. Certainly we need to know whether we are spending them usefully. But evaluation does not come as a Christmas gift. If we want first-class research officers, sound materials, careful analysis, adequate validation, these also cost a great deal of money.

It is as important to be selective in evaluation as in sweeping away old institutions or initiating new ones. But who should do the selection? Should it be the politicians or administrators? Should it be those directly involved in the field? Should it be public opinion? Should it be the evaluators? Each has a claim to a hearing. To ignore any of them may invite misunderstanding and eventual frustration. Yet none is immune from bees in the bonnet, from bias or self-interest. Sometimes the focus may be dictated by more sinister motives: there may be pressure for evaluation and change in one direction to distract attention from far greater evils in another. Glancing at the small print in the London Times a few days ago, I saw that a Roman senator has asked the Italian government to introduce paper bedsheets in all prisons to stop inmates from escaping by tying their bedlinen into ropes. The careful evaluation of such a proposal could still further defer evaluation of the rottenness of Italian prisons, by now a world secret. We are very far, as yet, from the

stage at which we can base firm and final policy decisions upon what emerges from evaluative research.

Since the criminal justice system has multiple objectives, we cannot sweep away any section of it solely as evidence that it is not achieving one of them. Prisons, for example, may not rehabilitate: at least in some instance it is essential that they continue to contain. Evaluation is a matter of degree: for example, if the police are catching no more than one criminal in three--or even less--this does not mean that they are totally ineffective: to withdraw all support from them would be to move into a radically new dimension of crime and disorder.

Evaluation in the sense of cost-benefit analysis is deeply dependent upon the more subtle kinds of evaluation. How do you weigh the costs and benefits of investment in one part of the system against those of investment in another? How do you weigh investment in criminal justice as such against investment in other social goals? Yet in the face of all these cautions it cannot be said too emphatically that evaluative studies must be continued, refined, and developed. Disappointment with crude findings must be used--as they are in the far larger, established physical sciences--as stepping stones to deeper and more accurate understanding of the complex web of human relationships called the criminal justice system. Even in what I may dare to call their infancy, they have done much to sweep away pretenses, brought more sobriety, realism, and self-examination into the work of those who gave the awesome responsibilities of "disposal," where law breakers are concerned (a sinister enough word in all conscience). And evaluation has made it far less easy than it used to be for demagogues to get away with sweeping assertions and the kind of promises that mean nothing and cannot be kept.

If you take evaluation in its widest sense, you find it covering virtually everything--all individual factors in crime, all social factors, all aspects of political and social life. It is like the days when we were so absorbed in finding "causes of crime" or ways of "predicting" it.

It is necessary to distinguish between quantitative and qualitative assessment; they may sometimes be combined, but sometimes they are separate and different views may result if a question is approached by way of quantitative or qualitative evaluation.

Some effects are visible, some are not. Yet the nonvisible may be at least as important as the visible. For example, if the police are very effective, they may reduce disorder virtually to nil: the tendency is then to assume that that is a normal state of affairs (just as we assume it is normal not to be burgled every time we go out). We begin to worry about visible effects only when police efficiency is low and crime and disorder high. Evaluation should not be carried out by the people to introduce innovations or implement the established system. Even so, bias among evaluators may be as potent and as prejudicial as bias among those who launch new experiments or defend established traditions. The subjects chosen for evaluation, the ways investigations are designed, strangely affect the nature of the findings.

May I add a note of warning. It might be possible to devise an evaluative study of the relation between bedsheets and escapes. It might even lead to a useful reduction in absconding. It would pass the test in technical terms.

Yet other criteria are needed to decide priorities for evaluative research. Some means are surely needed to steer scarce resources of skill, of time and money, toward projects near the center of criminological and penological concern, investigations that promise to throw light on the kind of problem we most need to tackle.

It is often said that evaluative studies have practical objectives, the hope of influencing future decisions. But the translation of research findings into action is by no means a clear and simple process. It is seldom possible to base a policy decision upon the results of an isolated study. Occasionally, success or disaster may be so clearly demonstrated that no further investigation is necessary. But generally two, three, or more investigations will be required before a decision can justifiably be based upon them.

There is no room for the facile assumption that a project which works in one setting, with one group of practitioners and clients, can be translated as it stands to a different time or place--let alone to a different country--or to other kinds of practitioners or offenders. Evaluative studies lend themselves less than most to generalization.

There is the need to keep in mind the degree of success that can reasonably be hoped for, the degree of failure that must inevitably be accepted, in a field as intrinsically discouraging as that of criminal justice. Even a small indication of achievement can have a disproportionate effect in relieving pressures, raising morale, and allaying public disquiet.

Because criminal justice has multiple objectives and functions, it cannot be concluded that an institution or project which is failing to achieve one of them is necessarily valueless and should be swept away. A great deal of evaluative effort and talent has been devoted to demonstrating that prisons fail to rehabilitate prisoners. Admittedly many prison administrators have in the past colluded in presenting rehabilitation as the primary purpose of their regimes, and the public have wanted to believe them. But prisons have other functions, more ancient, and more indispensable. In the last resort they are designed to contain and restrain. No amount of evaluation in other terms will make them expendable in that most basic of functions.

To reach decisions about changes in criminal justice, we must take account of the system as a whole. In this sense too, not one, but a series of evaluations is needed. Before deciding that something is expendable in one part of the continuum, we must consider its impact upon the rest. To abolish it may impose intolerable strains on other parts. And before sweeping away one section as inefficient, costly, or even harmful, we must consider whether the alternatives may be even less efficient, more expensive, or damaging, than what we have and what we complain of. Likewise, before throwing our resources into something new, we must consider whether we are thereby undermining something already established which could still be of service and which may, indeed, include vital safeguards built in by experience over many years.

If I sound very conservative to you, you must allow me that I have lived long enough to have watched the rise and fall of many fashions and many extremes in criminal justice. I plead only for balance, and a measure of sober economies in mingling the best of the new with the best of the old. In the last resort we came back to the intangibles of evaluation, the areas where no

amount of objective calculation will give us the answers, where we have to decide on the basis of "intrinsic worth"--what weight we will attach, for example, to the protection of the community from serious crime and the protection of the criminal from serious injustice.

Yet I have no doubt at all that, despite all its complications, all its flaws, and all its limitations, the development of evaluation must continue. Only a completely stagnant society, which has given up hope of anything, can afford to neglect it. Even the motley assortment of authoritarian regimes now tightening their grips around the world have their methods of evaluation, of weighing up the costs and benefits to their regimes of the various ways of handling their criminals and their delinquents. There must, for example, be calculation about the most effective way of countering the dangers presented by dissidents, not only in the dark corridors of the public prosecutors and at police headquarters, but also at the very highest political level. Who should be imprisoned, tortured, or consigned to a psychiatric hospital? Who should be exiled? How can we evaluate the impact of these alternatives, in different cases, upon public opinion at home, upon world opinion abroad? But for evaluation in the sense we know it, an opening up and examination of the processes of criminal justice and the penal system, we are forced to depend upon the testimony of those who have been their victims. We hear from the Solzhenitsyns', not from the institutes of criminology of the totalitarian states.

Evaluation in an open society is a very different matter. In fact, I would say that genuine evaluation is possible only in an open society. It is indeed one of its essential political implications: an open society wants to know what is going on, from the Oval Office in the White House to a local jail in Mississippi. In this field, and not for the first time, American criminological practice has given the lead to the world. Because the United States has been free of the rigid frameworks and clichés of the European schools of criminal law and criminology, it has been free to enter fresh fields of research. Because of its essentially questioning outlook, at once idealistic and utilitarian, it has persistently demanded to know what has come out of its development and experiments. Because it has been faced with a persistent and growing phenomenon of crime, in spite of early optimism about the possibilities of control, it has gone on to develop new approaches. I see that Mr. Caplan has spoken of evaluation as a "new science." I would venture to disagree: it is more like a new and intriguing chapter in a very long story, a refinement and expansion of a kind of assessment that has long been an art and is now becoming a science as well. Moreover, it is a development to which the nation has devoted a generous share of resources.

Evaluation opens doors in the system of criminal justice, not only for administrators and field-workers, but for politicians and legislators. It has done much to sweep away pretenses, brought more sobriety and realism into the outlook of all who have a say in dealing with crime and criminals. Evaluative studies, as well as the persistence of crime, have brought home the lesson that the options are few, and that the possibilities of influence are very limited. The programs of politicians, small and great, will always include some allusions to the elimination of corruption and the control of crime. But we are by now far away, I am glad to say, from the fervent crusades of the political exponents of so-called "law and order" and of those who believed that crime would be absorbed and rendered negligible by the benefits of the "Great Society."

It was at once sobering and reassuring to see during the latest election, that although President Ford had in mind definite ways of dealing with certain classes of criminals, he abstained from emphasizing it too persistently. And his opponent, President Carter, barely touched upon penal problems. That should not be taken to mean that he does not realize the importance of the problems for the country as a whole; but only that he did not regard them as election issues. To borrow the title of a fashionable song--"Promises, Promises, Promises"--are mercifully losing their grip upon penal policy. The studies of evaluation have certainly played an important role in producing this beneficial change in attitude.

But I have come here not to teach, but to learn. It is time for me to stop. I am looking forward eagerly to watching the work of the various panels. Still more I look forward to following what emerges from it all in the years to come, to seeing sound fruit from the seeds being scattered here.

SECTION I  
POLICE EVALUATION

Preceding page blank

AN EVALUATION OF THE OPEN GARAGE  
DOOR BURGLARY PROGRAM

Ron Pennington  
Research Associate  
St. Louis County Police Department

INTRODUCTION

Police departments have historically either used a preventive patrol-oriented strategy or a target-hardening strategy to control the incidence of crime in their jurisdiction. A patrol-oriented strategy is usually based on the assumption that an increased police presence will deter crime and increase apprehensions. The Kansas City Preventive Patrol experiment, which tested the effect that three different patrolling levels had both on the occurrence of crime and on the community's attitudes about crime, constitutes one of the more well-known experiments of this type.<sup>1</sup> On the other hand, a target-hardening strategy will normally use nonpatrol means to implement a program, e.g., a police-community relations drive to recruit citizens into a home security program. Operation Identification programs, which encourage citizens to mark any moveable and valuable piece of property with an engraved identifying code and which attempt to deter potential burglars by putting a program decal on the door or window of the program participant, are usually administered by police-community relations bureaus.<sup>2</sup>

Program evaluation of patrol-oriented and target-hardening programs are primarily designed to test the utility of the program's concept by showing its effect on a wide variety of goals. In the Kansas City Preventive Patrol experiment, impact measures included data on citizen attitudes about crime and the police as well as data about the occurrence rates of crime, e.g., robbery, assault, rape, homicide, burglary, auto theft, vandalism, purse snatching, and larceny. The Operation Identification program was evaluated by the following criteria: reduction in residential burglaries, citizen fears about crime, recruitment to the program, police-community benefits, and property return rates.

In those situations where a police department has designed a program to attack a specific crime problem, elements of both program strategies have been used. The anti-subway robbery program, implemented by the New York City Police Department in 1965, illustrates this point. The New York City Police Department increased the number of men patrolling its subways by 150 percent in 1965 and the program soon had the effect of reducing subway offenses from a high level of 7,000 crimes in 1964 to a low level of 5,000 crimes in 1965. However, by 1968 field interrogations of arrested robbers and an analysis from crime statistics suggested that the program may have had a partial displacement effect, i.e., more bus robberies were being committed in lieu of subway robberies. In order to deal with this problem, New York City officials

implemented in 1969 an exact bus fare program which had the effect of reducing monthly bus robberies by 98 percent. In summary, what started out as a patrol-oriented program soon incorporated elements of a target-hardening program and in both instances, the specific nature of the crime dictated the program strategy to be used.<sup>3</sup>

This study assesses a crime prevention program that also used the elements of a patrol-oriented and a target-hardening strategy and which was implemented by the St. Louis County Police Department from April 1, 1976 to September 30, 1976. The St. Louis County Police Department primarily patrols the unincorporated areas of St. Louis County. The entire county is legally separated from the City of St. Louis and has within its boundary 951,671 people, 510 square miles, and 94 self-governing municipalities. The unincorporated area of St. Louis County has about one-third of the county's total population (348,431 people) and about two-thirds of its land area (325 square miles). The department is divided into five precinct areas and two of these precincts were selected as the site for the pilot program. According to the 1970 Census, residents in this test area tend to range from the lower middle to middle income bracket, are almost exclusively white, and have a significant juvenile population below 18 years of age (i.e., 34 percent).

The pilot program was designed to reduce unlawful entry garage burglaries<sup>4</sup> and home burglaries. It was based on the following assumption: if homeowners who leave their garage doors open were informed by the department that such negligent behavior was strongly associated with the commission of an unlawful entry garage burglary, then these types of crimes, which are primarily committed by juvenile offenders, could be prevented.<sup>5</sup> Patrol officers were told to write down the address of any home where an open garage door was spotted and where no resident appeared to be home. Lists of these addresses were forwarded to headquarters and a letter was sent to the resident. The letter stated that open garage doors provided burglars with an excellent opportunity to commit a theft from a garage or from a home that was attached to a garage. In those situations where a resident was observed with a garage door open and where program statistics indicated that a previous letter had already been mailed to the resident, a subsequent and more strongly worded letter was sent out. Additional brochure information, which outlined certain preventive techniques that the resident could undertake in order to help prevent home burglaries, was also included. No other contact was made with the resident after this second letter was sent out.

Several research findings substantiated the need for an antigarage burglary program. First, data showed that a greater percentage of the burglaries committed in areas patrolled by the department were garage burglaries than in the rest of St. Louis County. Twenty percent of all burglaries that were committed in areas patrolled by the department during 1974 and 1975 were garage burglaries and only 13 percent of all burglaries that were committed in the rest of the county during the same time period were garage burglaries. Second, the number of garage burglaries increased by 8 percent (from 1,074 in 1974 to 1,162 in 1975); while at the same time, the average monetary value of property stolen from a garage burglary increased by 33 percent (from \$160 in 1974 to \$213 in 1975). Third, data also showed that garage burglaries were seasonal crimes. Sixty-seven percent of all garage burglaries that were committed in areas patrolled by the department during 1974 and 1975 occurred from the months of April to September. Finally, a study that randomly selected 95

garage burglary reports written in 1974 revealed that at least 65 percent of these reports had a known open garage door means of entry. This particular finding was also supported by another study which showed that 86 percent of all garage burglaries committed in areas patrolled by the department during 1974 and 1975 were associated with an unlocked means of entry. In summary, the data showed that garage burglaries were: (a) a problem that was becoming worse in terms of the actual number of crimes committed and in terms of the average monetary value of property stolen, (b) a crime that was strongly related to the warmer months of the year, and (c) a crime that may not have occurred if the victim had taken the simple precautionary measure of closing the garage door.

#### AN ANALYSIS OF THE PROGRAM'S IMPACT

At the end of the program period, analysis of the data showed that unlawful entry garage burglaries decreased by 32 percent in the test area from a preprogram period (April to September 1975) to the program period (April to September 1976). However, home burglaries increased by 7 percent from the pre-program period to the program period. Since the program had no apparent effect on home burglaries, they were eliminated from the study. Table 1 summarizes these findings.

TABLE 1.--The distribution of unlawful entry garage burglaries and home burglaries in the test area by time periods

Time periods	Unlawful entry garage burglaries	Home burglaries	Total
Preprogram (April-September 1975)	218	415	633
Program period (April-September 1976)	147	444	591
Total	365	859	1,224
Percent change	- 32%	+ 7%	- 7%

This study will use a policy analysis perspective to help determine whether the 32 percent decrease in unlawful entry garage burglaries could be attributed to the effectiveness of the program. According to James Q. Wilson, a policy analysis perspective tests for the effect that a certain short-term policy innovation has on a specific crime.<sup>6</sup> However, one of the difficulties with this type of approach is that researchers do not always have the necessary planning time to build into a new policy the program features that would create a true experimental situation.<sup>7</sup> Usually, program evaluators will try to find some way to randomize the introduction of the program stimulus. However, whenever randomization is not possible, the researcher may use a quasi-experimental analysis. According to Donald Campbell, this type of analysis provides the

researcher with the ability to "introduce something like experimental design into his scheduling of data collection procedures."<sup>8</sup> Carol Weiss also argues that quasi-experimental designs should be used when the conditions for a true experiment do not exist.<sup>9</sup>

Eighty-four subprecinct areas in the test zone (known as COGIS blocks)<sup>10</sup> were used as the unit of analysis in a test that attempted to determine if the program had an effect on unlawful entry garage burglaries. Two types of variables were computed at this level: letter saturation levels and changes in unlawful entry garage burglaries from the preprogram period to the program period.

In order to measure the level of letter saturation for each subprecinct area, the total number of initial letters and the total number of secondary letters that were mailed out during the program period were divided by the number of houses in each COGIS block. Housing data were selected as the base measure because they were the most valid measure of potential risk for this type of crime.<sup>11</sup> Although criminal justice students have habitually used a per capita base for these types of measures, such a practice for this study would clearly be inappropriate because structures and not people constituted the environmental risk encountered by the program.<sup>12</sup>

The second variable classified COGIS areas by whether they experienced an increase, a decrease, or no change in the number of unlawful entry garage burglaries that were committed from the preprogram period. In order to reduce the uncertainty about the relationship between true crime and reported crimes, the data were treated ordinally. According to Jones, this technique is valid so long as there is a positive relationship between the two concepts, i.e., when the true crime rate increases (decreases), the reported crime rate will increase (decrease).<sup>13</sup> This analysis also assumes that the program did not have a "Hawthorne effect" on the victims who normally report (or do not report) garage burglaries to the police.

Table 2 presents a contingency table that shows how the test zones' 84 COGIS areas are distributed between the saturation level of initial letters and the change in unlawful entry garage burglaries. The first variable consists of three categories:

- (a) a low saturation level (less than 9.2 initial letters per 100 homes),
- (b) a middle saturation level (between 9.3 and 21.4 initial letters per 100 homes), and
- (c) an upper saturation level (more than 21.5 initial letters per 100 homes).

The program supporting hypothesis is: those COGIS blocks that experienced higher saturation levels of initial letters should tend to be more strongly associated with a decrease in unlawful entry garage burglaries than those COGIS blocks that experienced lower saturation levels of initial letters.

The data in table 2 show that the saturation levels of initial letters were not strongly related to the change in unlawful entry garage burglaries. Only 50 percent of those COGIS blocks that experienced an upper saturation

TABLE 2.--Saturation levels of initial letters by  
changes in unlawful entry garage  
burglaries within the test zone

Change in unlawful entry garage burglary	Saturation levels			Total
	Low level	Middle level	Upper level	
Increase	10 (37%)	10 (34%)	5 (18%)	25
No change	4 (15%)	4 (14%)	9 (32%)	17
Decrease	13 (48%)	15 (52%)	14 (50%)	42
Total	27 (100%)	29 (100%)	28 (100%)	84

level of initial letters had a decrease in unlawful entry garage burglaries. Fifty-two percent of the COGIS blocks in the middle saturation range and 48 percent of those COGIS blocks in the lower saturation range had a decrease in unlawful entry garage burglaries.

When the open garage door program was implemented by the department, it was felt that a second letter might provide an additional stimulus to those homeowners who continued to leave their garage doors open. Consequently, it was possible that a relationship between saturation levels of secondary letters and unlawful entry garage burglaries might exist even though no relationship was found for initial letters. Saturation levels of secondary letters classified COGIS blocks into the following categories:

- (a) no saturation level,
- (b) a low saturation level (less than 5.0 secondary letters per 100 homes), and
- (c) an upper saturation level (greater than 5.0 letters per 100 homes).

The following hypothesis was tested: those COGIS blocks that experienced higher saturation levels of secondary letters should tend to be more strongly associated with a decrease in unlawful entry garage burglaries than those COGIS blocks that experienced lower saturation levels of secondary letters.

The data in table 3 show that saturation levels of secondary letters are not related to changes in unlawful entry garage burglaries. Only 49 percent of those COGIS blocks that experienced an upper saturation level of secondary

TABLE 3.--Saturation levels of secondary letters by  
changes in unlawful entry garage  
burglaries within the test area

Change in unlawful entry garage burglary	Saturation levels			Total
	No level	Low level	Upper level	
Increase	5 (50%)	10 (27%)	10 (27%)	25
No change	1 (10%)	7 (19%)	9 (24%)	17
Decrease	4 (40%)	20 (54%)	18 (49%)	42
Total	10 (100%)	37 (100%)	37 (100%)	84

letters had a decrease in unlawful entry garage burglaries. On the other hand, 54 percent of those COGIS blocks that experienced a low saturation level of secondary letters had a decrease in unlawful entry garage burglaries.

#### CHECKS FOR INTERNAL VALIDITY THREATS

According to Campbell, one of the overriding virtues of quasi-experimental designs in a nontesting environment is that they control for alternative explanations for why a program did (or did not) have an impact. These explanations are called by Campbell internal validity threats. Two of these threats have particular relevance to this study. They are:

- (a) instrumentation (a shifting of the measuring instrument independent of any change in the phenomenon measured) and
- (b) regression (the atypical occurrence of an exceptionally large number of unlawful entry garage burglaries during the pretest period; thereby causing a regression toward a general trend line that would have predicted fewer unlawful entry garage burglaries during the program period).<sup>14</sup>

The problem of instrumentation (or instrument decay) actually entails questions about the validity and reliability of crime statistics. According to Skogan, a validity problem in crime statistics occurs when "a researcher's procedures may not be measuring the object of analysis or the resulting figures may be artifacts of the measuring process" and a reliability question in crime statistics will "gauge the ability of police patrol teams to classify the same sort of events in the same manner."<sup>15</sup> The problem of instrumentation will occur whenever a validity or a reliability problem threatens a study's findings to the extent that the program's impact (or lack of impact) can be attributed to

a shift in the measuring instrument. The study was confronted with both types of measurement problems.

The problem of hidden crime, i.e., the difference between true crime and reported crime, constitutes the most serious validity threat to any evaluation study using crime statistics. Ostrom notes that many criminal acts are never reported to the police for various reasons.<sup>16</sup> Specifically germane to this study, one governmental victimization survey of eight large cities found that 55 percent of unlawful entry burglaries were never reported to the police.<sup>17</sup> However, Maltz notes that unreported crime is a serious problem in program evaluations only when there is evidence that the program might have an effect on reporting rates.<sup>18</sup>

The only effective test for this measurement problem would have been to implement an expensive victimization survey before and after the program was started. However, a validity test of sorts was made by predicting that a non-treated control area would have the same percentage decrease of unlawful entry garage burglaries as was experienced in the test area. The same percentage decrease in both areas would support earlier findings that the program had no impact in the test area since the control area's decrease could be attributed to reasons other than the effect of the program. Because any valid measurement of a variable will tend to consistently predict the same outcome with a fairly high degree of accuracy, such a test would also provide limited evidence that reporting rates were not significantly affected by the introduction of the program.

Table 4 shows the number of unlawful entry garage burglaries that occurred in the test area and the control area during the preprogram and program period. The control area for this table included the three precinct areas patrolled by the department which did not experience the introduction of the program. The data in table 4 show that both areas had about the same percentage decrease of unlawful entry garage burglaries, i.e., a 32 percent decrease in the test area and a 34 percent decrease in the control area. Consequently, the data in this table support the previous finding of this study that the program had no discernible impact on the occurrence rate of unlawful entry garage burglaries in the test area.

Because a crime prevention program may affect how a police officer may perceive a program-related crime, crime statistics may not be reliable. According to Ostrom, the researcher may not be aware of the variations in the reporting practices within a police department.<sup>19</sup> This problem becomes critical to an evaluation study whenever the program's apparent impact (or lack of impact) can be attributed to a change in police reporting practices in the test area. In essence, this situation is a problem of instrumentation.

Information from radio dispatched calls for service, which are computerized by the department, was used to indicate whether more (or fewer) crime incident calls were recorded in the department's crime statistics. Specifically, the percentage of all larceny calls whose final disposition were recorded as a "report taken" was plotted by each month of the experimental period in the test and control area. Larceny incident calls were selected because garage thefts are always dispatched and recorded as larcenies.<sup>20</sup>

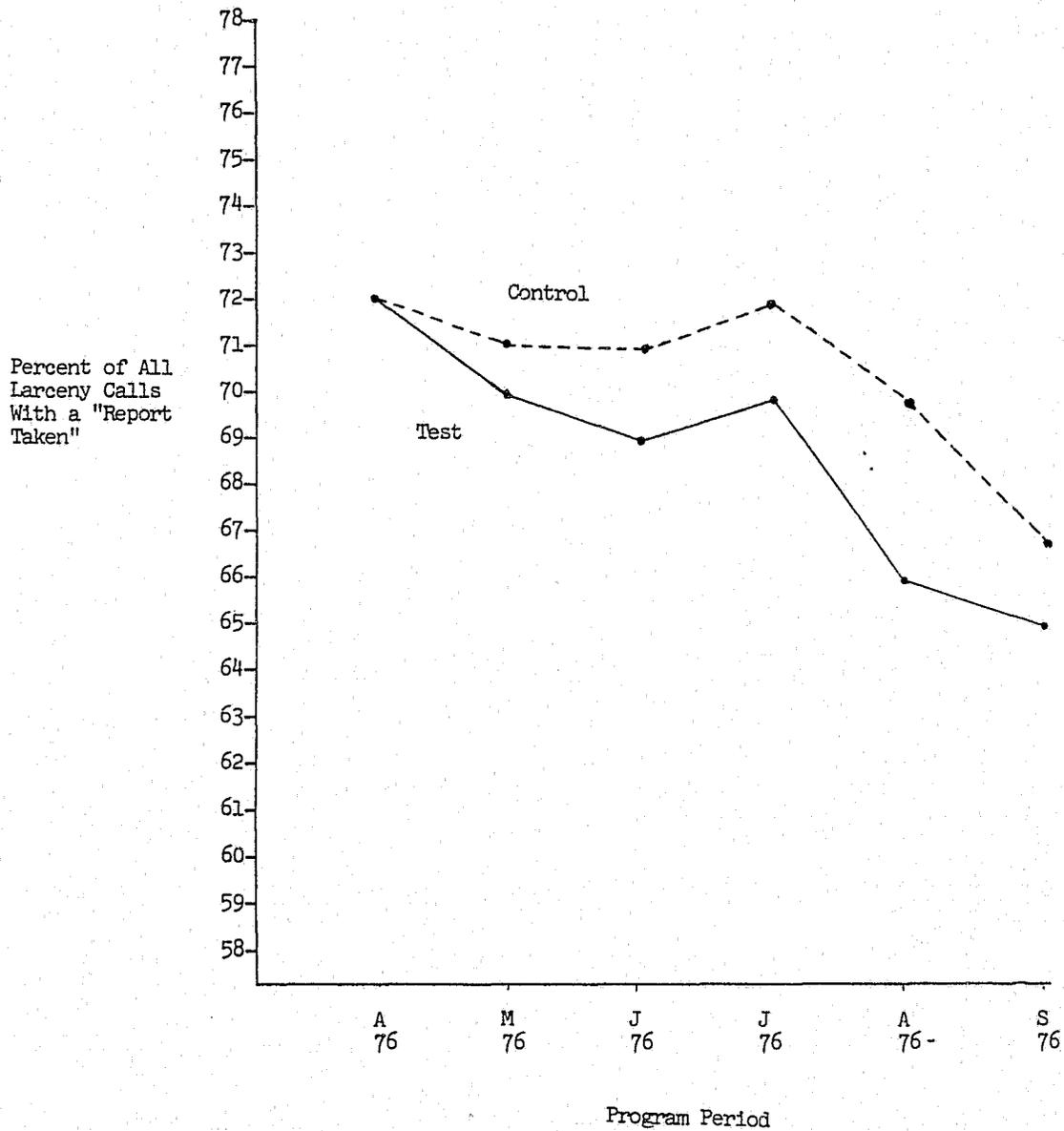
TABLE 4.--The number of unlawful entry garage burglaries that were committed in the test area and the control area during the pretest and program period

Time	Test area	Control area	Both areas
Preprogram (April-Sept. 1975)	218 ( 60%)	432 ( 60%)	650 ( 60%)
Program period (April-Sept. 1976)	147 ( 40%)	286 ( 40%)	433 ( 40%)
Total	365 (100%)	718 (100%)	1,083 (100%)
% of reduced crime	-32%	-34%	-33%

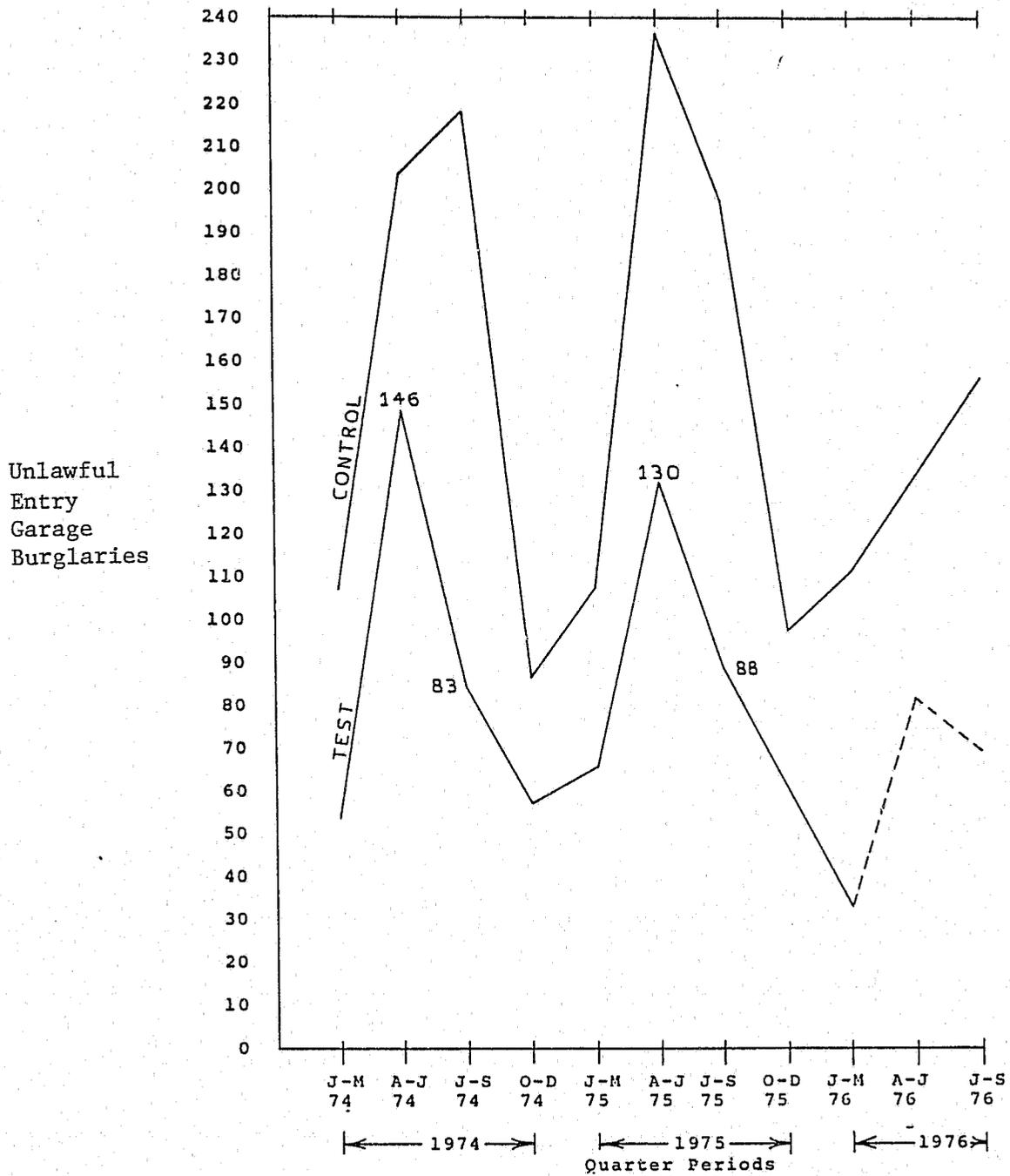
Graph 1 shows that at the initial outset of the program, the control area and the test area had about the same larceny reporting percentage. There was no percentage difference between the two areas in April 1976, a 1 percent difference between the areas in May 1976, and a 2 percent difference between the areas in June and July 1976. By August, the margin of difference between the areas had grown to 4 percent; but it declined to 2 percent in September 1976. In summary, for all months after April 1976, the test area tended to have a lower larceny reporting rate than the control area, suggesting that the program might have had some slight depressant effect on the reporting rate in the test area. Because the difference between the two areas' reporting rates was small and since the lower rate in the test area would have indicated more crime, it was concluded that any possible program effect on police reporting practices probably did not have a confounding influence on the previous findings of this study.

The second internal validity threat which presented a serious challenge to the study's findings was regression. According to Campbell, an interrupted time series test is the most effective way to determine whether this threat has occurred in an experiment.<sup>21</sup> Graph 2 is a time series analysis that shows the number of unlawful entry garage burglaries committed in the test and the control areas during quarterly preprogram and program periods that go back to January 1974. The dashed lines represent the test areas during the experimental program period. In essence, the graph shows that a regression effect did not occur during the program period because the quarterly 1975 preprogram periods (April-June 1975 and July-September 1975) were not greater than the same quarterly periods in 1974. The graph shows, for example, that there were 146 unlawful entry garage burglaries during the April-June 1974 period compared to the 130 unlawful entry garage burglaries during the April-June 1975 period and that there were 83 unlawful entry garage burglaries during the July-September 1974 period compared to the 88 unlawful entry garage burglaries during the July-September 1975 period.

GRAPH 1.--Percent of all larceny calls with a "report taken" disposition by monthly experimental periods for the test and control areas



GRAPH 2.--Unlawful entry garage burglaries committed in the test and control areas by quarterly periods



In order to provide more information about the program's impact, the control area was also plotted in the graph, thereby expanding the analysis from a simple time series test to a multiple time series test. This additional feature proved useful because it clearly illustrated the seasonal characteristic of the study's crime. Generally, there were fewer unlawful entry garage burglaries during the colder months of the year (October to February) and there were more of them during the warmer months of the year (April to September). However, the graph also clearly shows that the magnitude of these warmer month increases started to decline in 1975, 1 full year before the program was implemented. Consequently, the data suggest that although unlawful entry garage burglaries were at an all time low during the program period as compared to previous warmer month periods, this trend may have simply been a part of a larger trend which started in 1975 and which has simply continued independently of any impact from the program.

#### CONCLUSION: EXPLANATIONS OF PROGRAM FAILURE

Two explanations may account for the failure of the program to have an impact. First, the program may not have reached a very high saturation level, i.e., only a small percent of negligent homeowners were reached by letters. Second, it was possible that the letters simply may not have evoked the necessary motivational force which would have changed the negligent behavioral patterns of homeowners. In other words, the normal homeowner who received a letter from the department simply ignored it. Data, which were available for the first explanation but which were not available for the second explanation, indicated that there was not enough supporting evidence for the first explanation. Consequently, the credibility of the second explanation was enhanced even though no direct test could be implemented.

One of the principal reasons for a program to fail is that the degree of change, which is necessary for the program to have an impact, is not reached. Freeman and Bernstein call this policy problem "process evaluation," and they maintain that it is an integral part of any evaluation study.<sup>22</sup> For example, the Kansas City preventive patrol experiment was critiqued by Richard Larson because reactive beat areas (areas receiving no preventive patrols) probably did not conform to the conditions the researchers sought to introduce. Consequently, the study's failure to find statistical relationships between varying police visibility levels and the study's numerous dependent variables was not particularly surprising.<sup>23</sup>

Data in table 5 show that the garage burglary program was successfully implemented in the test area. According to the table, 17.2 percent of all homes in both test precincts received at least one letter and 5.6 percent of all homes in both test precincts received a second letter. Aware of the fact that not all homes have garages and that not all garage owners leave their doors open, the real saturation scores are probably much higher than the scores recorded here. In addition, it should be noted that most of the letters were mailed during the first 3 months of the program. From an evaluation point of view, this was desirable because the program's impact should have occurred, at the latest, during the last 3 months of the program.

Although there is no supporting evidence to argue for the acceptance of the second explanation, i.e., the program's failure to motivate a behavior change on the part of the homeowner, one is forced to conclude by a process of

TABLE 5.--The number of initial and secondary letters mailed related to homes in each precinct by time periods

Program time periods	Precinct area					
	Precinct A <sup>1</sup>		Precinct B <sup>2</sup>		Both Precincts	
	Initial	Secondary	Initial	Secondary	Initial	Secondary
First three program months (April-June 1976)	3,012 <sup>3</sup> (13.7) <sup>4</sup>	802 (3.6)	2,461 (16.8)	658 (4.5)	5,473 (14.9)	1,460 (4.0)
Second three program months (July-Sept. 1976)	213 ( 1.0)	176 ( .8)	622 ( 4.2)	403 (2.7)	835 ( 2.3)	579 (1.6)
Total	3,225	978	3,083	1,061	6,308	2,039

<sup>1</sup>Precinct A has 21,936 homes according to the 1970 Census.

<sup>2</sup>Precinct B has 14,658 homes according to the 1970 Census.

<sup>3</sup>Number of letters mailed.

<sup>4</sup>Number of letters per 100 homes in the precinct.

Source: Housing statistics were collected from the owner total column on page 8 through 95, U.S. Department of Commerce, Bureau of the Census, Block Statistics St. Louis, Mo.--III. Urbanized Area, 1970 Census of Housing.

elimination that this explanation is probably the most plausible. However, three types of tests could have been implemented to determine whether this explanation could be more strongly supported. First, a pretest and posttest sample survey, which would have determined how often garage owners closed their doors, could have been conducted in the test and the control area. Second, a more unobtrusive method could have been designed which would have measured how many garage doors were left open in randomly selected areas throughout different periods of the experiment. Finally, some of the sightings made by the police during the program could have been treated as a control group to the extent that no letters would have been mailed to the resident. Consequently, a longevity study could have been implemented in order to determine if initial or secondary letter recipients tended to be victimized less than those homeowners who were spotted with open garage doors but who never received a letter.

In summary, three reasons underscore why these tests were never implemented. They were:

- (a) experimental requirements were not seriously considered before the program was implemented,
- (b) the additional tests would have increased the costs of a pilot program which was already becoming too expensive, and
- (c) program designers never thought that it would be desirable to determine why a program might fail.

Because of these reasons, a definitive answer about why the program had failed could not be rendered. However, through careful use of the data available, it was possible to determine that (a) the program had failed and (b) that the failure could not be attributed to a pseudo-statistical effect originating from the data. Given the uncontrolled environment in which the quasi-experimentalist must conduct a study, these two findings are not insignificant.

## NOTES

1. For a rather detailed evaluation study of this program, see Police Foundation, The Kansas City Preventive Patrol Experiment: A Technical Report (Washington, D.C., 1974).
2. See, The Institute for Program Evaluation: Summary of the Assessment of Operation Identification Effectiveness and Plans for Evaluating a Single Project: Phase I Evaluation of Operation Identification, prepared for National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice, 1975.
3. Jan M. Chaiken, Michael W. Lawless, and Keith A. Stevenson, "The Impact of Police Activity on Subway Crime," Journal of Urban Analysis, II, 2 (1974), 173-205.
4. An unlawful entry garage burglary was defined in this study as an illegal entry into a garage through an unlawful means of entry for the purpose of committing a theft. A home burglary was defined in this study as an illegal entry into a house which was a permanently fixed structure through either a forcible, unlawful, or attempted forcible means of entry for the purpose of committing a theft. Both of these definitions are consistent with the definitional criteria outlined by the Uniform Crime Reporting System. See: United States Department of Justice, Federal Bureau of Investigation, Uniform Crime Reporting Handbook: How to Prepare Uniform Crime Reports (Washington, D.C.: Government Printing Office, 1976), 22-27.
5. A reduction in home burglaries was a goal to the extent that a garage was attached to a home. In other words, it was reasoned that an easy entrance into a garage might lead to the burglary of a home that was attached to the garage. Since not every home was attached to a garage, it was predicted that the program would have more impact on garage burglaries than on home burglaries.
6. James Q. Wilson, Crime and Criminologists in Crime and Criminal Justice, ed. by Michael A. Mulkey (Lexington, Mass.: Lexington Books, D. C. Heath and Company, 1975), 13.
7. William S. Harrar and D. Lee Bawden, "The Use of Experimentation in Policy Formulation and Evaluation," Urban Affairs Quarterly VII, 4 (June 1972), 423-424.
8. Donald T. Campbell and Julian C. Stanley, Experimental and Quasi-Experimental Design for Research (Chicago, Ill.: Rand McNally & Company, 1963), 34.
9. Carol Weiss, Evaluation Research: Methods of Assessing Program Effectiveness (Englewood Cliffs, N.J.: Prentice Hall, Inc., 1972), 67-73.

10. COGIS blocks are police-reporting areas for the department. All offense reports and radio dispatch reports are geo-coded by COGIS block number. It should also be noted that COGIS blocks can be aggregated up to conform with Census tract boundaries, i.e., COGIS blocks are subdivided Census tracts.
11. Although it is true that a housing indicator is a weak measure of the number of garages in a COGIS block, it was nevertheless the best indicator that was available for this study. In addition, it could also be argued that 1970 housing data no longer reliably reflect today's true housing stock. However, building activity has not radically changed the residential makeup of the test area during this 6-year period.
12. Boggs criticizes this particular practice by forcibly arguing that the risk or target group, to which the crime is directed against, should be used as the base measure for any crime occurrence rate. By taking into account what she calls "environmental opportunities," the researcher may upgrade the validity of his indicators. See, Sarah L. Boggs, "Urban Crime Patterns," *American Sociological Review*, XXX (December 1965), 889-901.
13. E. T. Jones, "Evaluating Everyday Policies: Police Activity and Crime Incidence," *Urban Affairs Quarterly* VIII, 3 (March 1973), 271.
14. See Donald T. Campbell and Julian C. Stanley, "The Connecticut Crackdown on Speeding: Time Series Data in Quasi-Experimental Analysis," *Law & Society Review*, III, 1 (August 1968), 39.
15. Wesley G. Skogan, "Comparing Measures of Crime: Police Statistics and Estimates of Citizen Victimization in American Cities," *American Statistical Association Proceedings of the Social Statistical Section* (1974), 44.
16. Elinor Ostrom, "Institutional Arrangements and the Measurement of Policy Consequences, Applications to Evaluating Police Performance," *Urban Affairs Quarterly*, VI, 4 (June 1971), 458.
17. U.S. Department of Justice, Law Enforcement Assistance Administration, National Criminal Justice Information and Statistical Service, *An Analysis of Victimization Survey Results from the Eight Impact Cities*, (Washington, D.C.: Government Printing Office, 1974), 391.
18. Michael D. Maltz, *Evaluations of Crime Patrol Programs*, Report to U.S. Department of Justice, Law Enforcement Assistance Administration, National Criminal Justice Information and Statistical Service, Washington, D.C. (April 1972), 29.
19. Ostrom, "Institutional Arrangement," 459.
20. Under Missouri statutes, garage thefts are classified as larcenies. Consequently, they are dispatched and recorded by the department as larcenies. However, the department's Central Records Bureau reclassifies all garage larcenies to burglaries whenever inputting the information into the UCR system.
21. Campbell and Stanley, "The Connecticut Crackdown," 42.

22. Howard E. Freeman and Ilene N. Bernstein, Evaluation Research and Public Policies in Policy Studies and the Social Sciences, ed. by Stuart S. Nagel (Lexington, Mass.: Lexington Books, D. C. Heath and Company, 1975), 12.
23. Richard A. Larson, "What Happened to Patrol Operations in Kansas City? A Review of the Kansas City Preventive Experiment," Journal of Criminal Justice, III (1975), 267-297.

## BIBLIOGRAPHY

- Boggs, Sarah L. "Urban Crime Patterns," American Sociological Review, Volume XXX, Number 6 (December 1975), 889-908.
- Campbell, Donald T. and Stanley, Julian C. Experimental and Quasi-Experimental Design for Research, Chicago, Ill.: Rand McNally & Company, 1963.
- \_\_\_\_\_ and Ross, H. Lawrence. "The Connecticut Crackdown on Speeding: Time Series Data in Quasi-Experimental Analysis," Law & Society Review, Volume III, Number 1 (August 1968), 33-53.
- Chaiken, Jan M., Lawless, Michael W., and Stevenson, Keith A. "The Impact of Police Activity on Subway Crime," Journal of Urban Analysis, Volume II, Number 2 (1974), 173-205.
- Freeman, Howard E. and Bernstein, Ilene N. Evaluation Research and Public Policies. Policy Studies and the Social Sciences. Edited by Stuart S. Nagel. Lexington, Mass.: Lexington Books, D. C. Heath and Company, 1975.
- Harrar, William S. and Bowden, D. Lee. "The Use of Experimentation in Policy Formation and Evaluation," Urban Affairs Quarterly, Volume VII, Number 4 (June 1972), 419-430.
- Jones, E. Terrence. "Evaluating Everyday Policies, Police Activity, and Crime Incidence," Urban Affairs Quarterly, Volume VIII, Number 3 (March 1973), 267-279.
- \_\_\_\_\_. Conducting Political Research, New York: Harper & Row Publishers, 1971.
- Larson, Richard A. "What Happened to Patrol Operations in Kansas City? A Review of the Kansas City Preventive Patrol Experiment," Journal of Criminal Justice, Volume III (1975), 267-297.
- Maltz, Michael D. Evaluations of Crime Patrol Programs. Report to U.S. Department of Justice, Law Enforcement Assistance Administration, National Institute of Law Enforcement and Criminal Justice, Washington, D.C., April 1972.
- Ostrom, Elinor. "Institutional Arrangements and the Measurement of Policy Consequences, Applications to Evaluating Police Performance," Urban Affairs Quarterly, Volume VI, Number 4 (June 1971), 447-475.
- Police Foundation. The Kansas City Preventive Patrol Experiment, A Technical Report, Washington, D.C., 1974.
- Skogan, Wesley G. "Comparing Measures of Crime: Police Statistics and Survey Estimates of Citizen Victimization in American Cities," American Statistical Association Proceedings of the Social Statistics Section, Washington, D.C., 1974.

The Institute for Public Program Analysis. Summary of the Assessment of Operation Identification's Effectiveness and Plans for Evaluating a Single Project: Phase I Evaluation of Operation Identification. Prepared for National Institute for Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice, May 1975.

U.S. Department of Justice, Federal Bureau of Investigation. Uniform Crime Reporting Handbook: How to Prepare Uniform Crime Reports. Washington, D.C.: Government Printing Office, 1976.

U.S. Department of Justice, Law Enforcement Assistance Administration, National Criminal Justice Information and Statistical Service. An Analysis of Victimization Survey Results from the Eight Impact Cities. Washington, D.C.: Government Printing Office, 1974.

Weiss, Carol H. Evaluation Research. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1972.

Wilson, James Q. Crime and Criminologists. Crime and Criminal Justice. Edited by Michael A. Mulkey. Lexington, Mass.: Lexington Books, D. C. Heath and Company, 1975.

82919

POLICE ROADBLOCKS TO DECRIMINALIZATION: A MULTIPLE-TIME-SERIES ANALYSIS OF LAW ENFORCEMENT'S RESPONSE TO CHANGES IN PUBLIC DRUNKENNESS STATUTES\*

Dr. C. Thomas Dienes  
Professor of Law  
Law School  
George Washington University

Dr. David E. Aaronson  
Professor of Law  
Washington College of Law  
The American University

Dr. Michael C. Musheno  
Associate Professor of Criminal Justice and Public Affairs  
Arizona State University

INTRODUCTION

Decriminalization as an alternative strategy for handling public drunkenness took hold in the 1960's and early 1970's. The regional and national forces that coalesced around this issue<sup>1</sup> as reform-oriented, policy subsystems<sup>2</sup> focused on both the illegitimacy and impracticability of municipal court deliberation for solving this social and public health problem. In states where legislative and judicial mandates calling for decriminalization were eventually passed, reformers gave little attention to the potential reaction of the police to such a change. They simply assumed that the police would continue to serve as a viable intake agent for public inebriates under the "new" public health model of detoxification and treatment.

This article empirically evaluates the impact of decriminalization on police departments' performance in Washington, D.C. and Minneapolis, Minnesota. We question the facile assumption of routine police support for this task. Specifically, we hypothesize that there will be a statistically significant decline in the number of public inebriates formally handled by the police in the manner designated by the "law in the books."

The conceptual basis for this hypothesis is derived from the literature on organization theory as well as studies focusing on police behavior. First, given the removal of the criminal sanction, the intake of public inebriates falls outside the parameters of what both police officers and the command structure of police departments consider proper and important tasks.<sup>3</sup> Also, the loss of the criminal sanction eliminates a critical organizational incentive that elicits patrol officers' cooperation to carry out this often messy and time-consuming job.<sup>4</sup> Thus, given the broad discretionary powers available to implementing agencies<sup>5</sup> and their respective street-level bureaucrats,<sup>6</sup> we

\* We acknowledge the financial assistance of NILE-LEAA Grant Number 74NI-99-0055.

would argue that any newly formulated tasks that run counter to the "organizational" and "self-interests" of bureaucracies (and their respective field officers) have very limited potential for full implementation (unless, of course, these new mandates carry with them explicit incentives designed to elicit organizational support at both the administrative and street levels).

Last, police intake of inebriates under a public health mandate requires the cooperation of two different public service bureaucracies that diverge in both their organizational structure and value orientation. Such a fragmented authority structure is an additional impediment to goal achievement.<sup>7</sup>

As for our research design, this paper is part of the growing body of literature which merges the common threads of empirical impact analysis and public policy analysis.<sup>8</sup> Thus, this "policy impact study" empirically evaluates the impact of state judicial and legislative mandates on agencies' responses to these directives.<sup>9</sup> We contribute, then, to both the breaking of the "upper court bias" associated with public law research<sup>10</sup> and policy analysis literature's increased focus on empirically assessing public agencies' interpretation of the law.<sup>11</sup>

Specifically, Washington, D.C. and Minneapolis, Minnesota have experienced three legal phases in the handling of public inebriates: (1) a criminal phase, (2) a transitional phase, and (3) a public health phase. In both jurisdictions, the transitional phase is marked by appellate court decisions which ruled on the chronic skid-row inebriate's status in the criminal justice system.<sup>12</sup> Decriminalization and the emergence of the public health phase derived from broader legislative mandates that required the establishment of new institutions for servicing the public inebriate population.<sup>13</sup>

#### DESIGN AND DATA COLLECTION

To empirically test the impact of decriminalization, we carried out an "interrupted time-series quasi-experiment"<sup>14</sup> based on a "stratified multiple-group-multiple I design"<sup>15</sup> (see figure 1). Specifically, we have collected monthly public drunkenness arrest rates (predecriminalization) and monthly rates of police deliveries to detoxification facilities (postdecriminalization) for two experimental cities: (1) Washington, D.C. (a high arrest jurisdiction)<sup>16</sup> and (2) Minneapolis, Minnesota (a moderate arrest jurisdiction).<sup>17</sup> Also, we have collected the available monthly arrest data for two control cities where decriminalization has not been implemented: Houston, Texas (a high arrest jurisdiction) and San Francisco, California (a moderate arrest jurisdiction).

These selections closely meet the criteria of what scholars often point to as critical ingredients for a strong design. The ". . . design is more valid the more heterogeneous each set of states is within itself and the more similar the two sets of states when each set is viewed as a whole."<sup>18</sup>

As many scholars carrying out time-series analysis well know, a laborious effort is often required in the search for relevant and reliable data that also provide enough observations to allow sophisticated analysis.<sup>19</sup> In that we were collecting data from four different municipalities, we were unable to collect an equivalent number of monthly observations for each jurisdiction, nor is

FIGURE 1.--Stratified multiple-group-multiple I design

---

Type A (D.C.--high arrest):	... 0 0 0 0 I <sub>1</sub> 0 0 0 0 ...
Type B (Minn.--moderate arrest):	... 0 0 0 0 I <sub>1</sub> 0 0 0 0 ...
Control A (Houston--high arrest):	... 0 0 0 0 I <sub>2</sub> 0 0 0 0 ...
Control B (S.F.--moderate arrest):	... 0 0 0 0 I <sub>2</sub> 0 0 0 0 ...

---

I<sub>1</sub>: decriminalization of public drunkenness.

I<sub>2</sub>: no decriminalization of public drunkenness.

the time sequence the same for each jurisdiction. Also, the date of decriminalization (I<sub>1</sub>) is different in the experimental jurisdictions.

Graphs 1 through 4 depict these differences and also indicate the decision rules arrived at concerning the placement of the intervention line (I<sub>1</sub> or I<sub>2</sub>) for each jurisdiction. The intervention line drawn for each of the decriminalized jurisdictions (I<sub>1</sub>) was based on two criteria:

- (1) the date that decriminalization took effect in each jurisdiction, and
- (2) the date that the public health facility (i.e., the detox facility) opened to receive clients.

In Minneapolis, the Alcoholism Receiving Center opened on the same date decriminalization became effective--July 1, 1971. Thus, for Minneapolis, we designate this date as the point of intervention. While decriminalization became effective on August 1, 1968, in Washington, D.C., the Detoxification Center was not fully operational until November 1, 1968. For Washington, D.C., then, we designate November 1, 1968, as the point of intervention.

We based the decision rule for drawing the intervention lines in the control jurisdictions (I<sub>2</sub>) on the following considerations:

- (1) a review of the number of observations that were available before and after decriminalization for the experimental jurisdictions;
- (2) a desire to match and therefore control for potential seasonal patterns emerging from police behavior in the experimental and control jurisdictions; and
- (3) an attempt to maximize the overlay of observations among the jurisdictions.

A composite of these decision rules and their influence on the overall design is depicted in figure 2.

FIGURE 2.--Distribution of observations

Wash., D.C. :	:	0 <sub>-34</sub> . . . . .	0 <sub>-1</sub> I <sub>1</sub> 0 <sub>+1</sub> . . . . .	0 <sub>+74</sub>
Minneapolis, Minn.:	:	0 <sub>-66</sub> . . . . .	0 <sub>-1</sub> I <sub>1</sub> 0 <sub>+1</sub> . . . . .	0 <sub>+38</sub>
Houston, Tex. :	:	*0 <sub>-18</sub> . . . . .	0 <sub>-1</sub> I <sub>2</sub> 0 <sub>+1</sub> . . . . .	0 <sub>+36</sub>
S.F., Calif. :	:	0 <sub>-18</sub> . . . . .	0 <sub>-1</sub> I <sub>2</sub> 0 <sub>+1</sub> . . . . .	0 <sub>+34</sub>

\* The 36 observations after I<sub>2</sub> (no decriminalization) are not continuous. Twenty-four monthly observations (1972, 1973) were unavailable.

FINDINGS

The data provide considerable support for our decriminalization hypothesis. Specifically, in Washington, D.C., the estimated change in level is a reduction of 76.4 police intakes per month which is significantly different from zero.<sup>20</sup> In Minneapolis, the impact of decriminalization on police intakes is more dramatic. Here, the estimated change in level is an even greater reduction of 263.2 police intakes per month.<sup>21</sup> Simple analysis of the data from our control jurisdictions (i.e., visual scanning)<sup>22</sup> shows that no similar effect takes place in police departments where criminal sanctions against public drunkenness remain intact (see graphs 3 and 4).

Does this mean, then, that one effect of decriminalization is increased neglect of the public inebriate population? Rather than concluding from the above analysis that inebriates are being left on the street at a significantly higher rate since decriminalization, we also investigated a series of alternative dispositions and control factors that could not be analyzed under the stratified multiple-group-multiple I design. As we will show below, our investigation of these factors points to the importance of "micro analysis" in tracing the impact of legal mandates on administrative agencies.

For each experimental jurisdiction (see figures 3 and 4), we analyzed whether a change in the recidivism rate (pre-, postdecriminalization) and/or a change in the size of the drinking population (pre-, postdecriminalization) might explain the apparent reduction in police pick-ups.

As we noted above, the reform legislation in both jurisdictions allows for self-admissions to the Detox facilities and grants the police two additional options for handling public inebriates--take the person home or deliver the individual to a facility equipped to handle alcoholism (e.g., hospital). Also, the Minnesota legislation explicitly sanctions civil pick-up of public inebriates. Thus, the Hennepin County Alcoholism Receiving Center staffs a Civil Pick-up Van designed to reduce pressure on the Minneapolis Police Department in the downtown section of the city (First Precinct) where the street inebriate problems are most acute.<sup>23</sup> Finally, in addition to these approved actions, we investigated whether the police are using nonapproved options for

processing public inebriates (i.e., misdemeanor charges: disorderly, vagrancy) in both jurisdictions.

FIGURE 3.--Micro analysis framework: Washington, D.C.

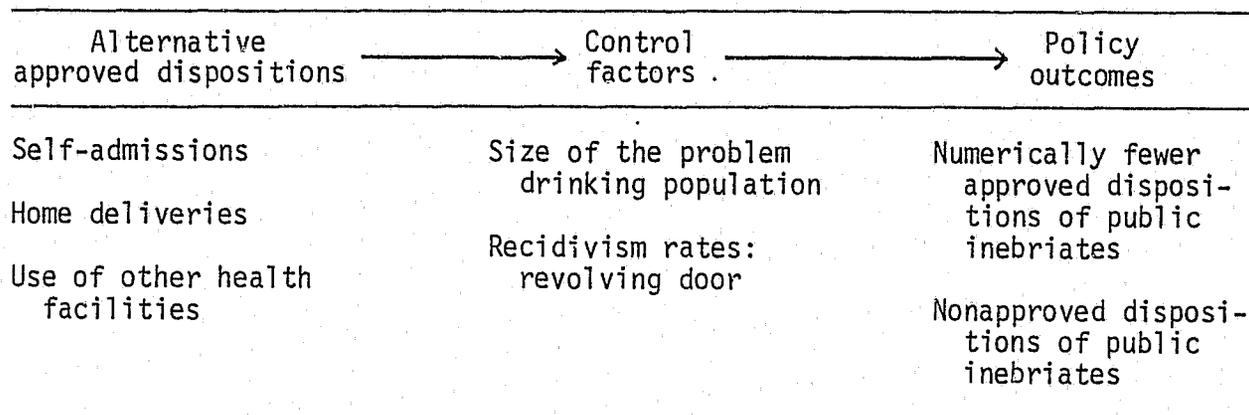
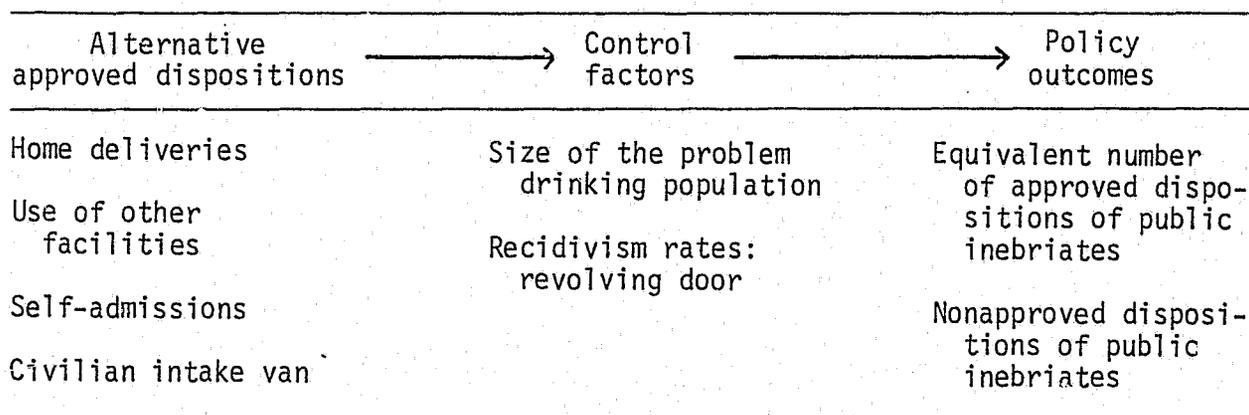


FIGURE 4.--Micro analysis framework: Minneapolis, Minn.



In Washington, D.C., we expected no significant alteration in our original finding of a significant decline in the number of public inebriates formally handled by the public system. While a series of alternative dispositions existed in the legislation, we detected no administrative initiative on the part of public health or police personnel to implement any of these options. Indeed, our exhaustive evaluation of these alternatives revealed no public health or law enforcement recordkeeping for these options, and no report of any sizable use of these options corresponding to the emergence of the post-reform era.<sup>24</sup>

As for control factors, the size of the problem-drinking population in Washington, D.C. has shown a yearly increase ever since such estimates have been calculated by the public health community in the District (i.e., 1960).<sup>25</sup>

Thus, given the absence of any decrease in the size of the problem-drinking population, there is no reason to expect any decrease in the public inebriate population that corresponds to decriminalization.

Last, our unit of analysis for the foregoing research has been "rate of intake" without consideration given to the number of individuals that are picked up in each period. Thus, one could postulate that as many individuals are being picked up in the post-ARA period as were in the pre-ARA period with the only difference being the lower rate of recidivism in the latter period. While this is unlikely due to the 72-hour restriction on involuntary commitment under decriminalization, we calculated estimates of the number of individuals that the police processed in 4 pre-ARA years (1964, 1966, 1967, 1968)<sup>26</sup> and compared these findings with the yearly recidivism rates for the D.C. Detoxification Center (1969-1973).<sup>27</sup> As shown in tables 1 and 2, the recidivism rates are uniformly higher in the post-ARA era, and therefore the revolving door argument fails to explain the discrepancy in police intake between the two periods.

Finally, we researched the possibility that the police and the courts are processing public inebriates for criminal offenses in the post-ARA period. In order to test this rival hypothesis, interviews were conducted with court personnel to determine whether such a practice was occurring and if so, to find out what offenses were being used for this purpose. All of those interviewed asserted that public inebriates are no longer being processed by the courts and, in addition, many pointed out that the primary factor responsible for reducing the case backlog in the Criminal Division of the Superior Court has been the removal of public drunkenness as a criminal offense. Some further suggested that because such charges as disorderly conduct and vagrancy were often attached to public drunkenness charges in the pre-ARA period, the criminal justice system has seen a reduction of these offenses in the post-ARA era.

We obtained official police statistics to probe these assertions, and to consider the possibility that other charges (principally disorderly conduct and vagrancy) were being used to process public drunks through the criminal justice system in the post-ARA period. As indicated in graphs 5 and 6, official arrest statistics from the Metropolitan Police Department establish that disorderly conduct and vagrancy charges have decreased substantially in the post-ARA period. The sharp increase in disorderly conduct arrests in fiscal year 1971 is most likely attributable to police actions regarding antiwar demonstrations, as over 9,000 of the arrests took place in May 1971, the month of the "May Day Demonstrations" in Washington, D.C. Thus, the official statistics and the information derived from the interviews strongly suggest that other crimes are not being used to any significant extent to process public drunks.

Unlike Washington, D.C., we expected our analysis of alternative dispositions in Minneapolis to reveal a significant nonpolice network of public inebriate intake that compensates for reduced police involvement. The public health establishment promotes self-admissions and their initiation of the civilian intake van assures a flow of public inebriates to Detox who frequent the "honky-tonk" areas of downtown Minneapolis.

As in the District of Columbia, neither of the control factors explains the discrepancy in pick-up between the two periods. The problem-drinking

TABLE 1.--Estimation of recidivism rate for individuals  
arrested by police, 1964, 1966, 1967, 1968

Year	Rate of arrest <sup>a</sup>	Court sample recidivism rate <sup>b</sup>	Estimation of indivs. arrested <sup>c</sup>
1964	44,107	1.58	27,916
1966	42,189	2.59	16,289
1967	31,860	1.48	21,527
1968	14,354	1.23	11,670

<sup>a</sup>Based on official statistics, Metropolitan Police Department, which are compiled on a FY basis. A rough conversion, using 50 percent of each FY has been made to bring this data into congruity with the court data.

<sup>b</sup>Based on sample of arrested individuals, D.C. Court of General Sessions Index, by calendar year.

<sup>c</sup>Rate of arrest divided by court sample recidivism rate.

TABLE 2.--Recidivism rate for individuals delivered to  
Detox, calendar 1969, 1970, 1971, 1972, 1973

Year	Rate of admissions <sup>a</sup>	Recidivism	Individuals admitted <sup>a</sup>
1969	11,695	3.03	3,856
1970	14,293	3.32	4,310
1971	14,845	3.15	4,707
1972	12,465	2.87	4,345
1973	10,436	2.68	3,893

<sup>a</sup>Official statistics of the Men's Detoxification Center.

population has slightly increased since decriminalization,<sup>28</sup> and as shown in table 3, public drunkenness recidivism rates are higher in the decriminalized era.

TABLE 3.--Comparison of public drunkenness recidivism rates between criminal and decriminalized

Year	# of individuals	Estimated recidivism
1967 <sup>a</sup>	N = 145	3.79
1970 <sup>a</sup>	N = 176	3.94
1972 <sup>b</sup>	N = 176	4.71
1974 <sup>b</sup>	N = 151	5.03

<sup>a</sup>Based on official arrest records, Minneapolis Police Department, Bureau of Identifications.

<sup>b</sup>Based on official records, Alcoholism Receiving Center, Department of MH/MR/CD.

While we found no significant use of home deliveries or other health facilities by the Minneapolis police officers,<sup>29</sup> our investigation of alternative routes of disposition initiated by the Alcoholism Receiving Center (ARC) produced significant findings. Unlike other public health facilities that rely almost totally on police departments for the delivery of public inebriates to their doors, ARC's staff has aggressively sought out other means of attracting clients to their center.<sup>30</sup> The development of the Civil Pick-Up Service was designed to reduce pressure on the Minneapolis Police Department in the downtown section of the city (First Precinct) where street inebriate problems are most acute.<sup>31</sup> Also, they have made an effort to encourage self-admissions of problem drinkers from more stable socioeconomic backgrounds through advertising and by working closely with businesses and government agencies.<sup>32</sup> Perhaps, then, such overall involvement by the public health community significantly compensates for the reduction in police attention to this problem.

Graph 7 shows that the combined public health initiatives of civilian pick-up and encouragement of self-admissions do indeed compensate for the decrease in police intakes.<sup>33</sup> Prior to the existence of the Civil Pick-up Service, ". . . the Minneapolis Police Department accounted for 40% of the total admissions from 4:00 pm to 12:00 pm."<sup>34</sup> After the implementation of this option, ". . . the Pick-Up Team transported almost 50% of the total admissions to the Center and 80% of police and team admissions combined. . ."<sup>35</sup> for the same hours.

In fact, statistics collected by ARC show that the use of this option has increased total admissions while further reducing police involvement. For example, in June through August of 1974, ". . . the total number of admissions

to the Center increased 17% (from 2,299 to 2,689) while police referrals were reduced from 844 to 480 admissions."<sup>36</sup> Based on total admissions for the first 8 months of 1974, Civil Pick-Up admissions increased from 19 percent to 27 percent while police admissions were reduced from 23 percent to 17 percent.<sup>37</sup>

But are the police more fully involved in the intake of public inebriates through the use of minor criminal offenses in the decriminalized period? Public health officials have felt that since decriminalization the police have been picking up a considerable number of public inebriates, arresting them for disorderly conduct, and releasing them before court appearance is required.<sup>38</sup>

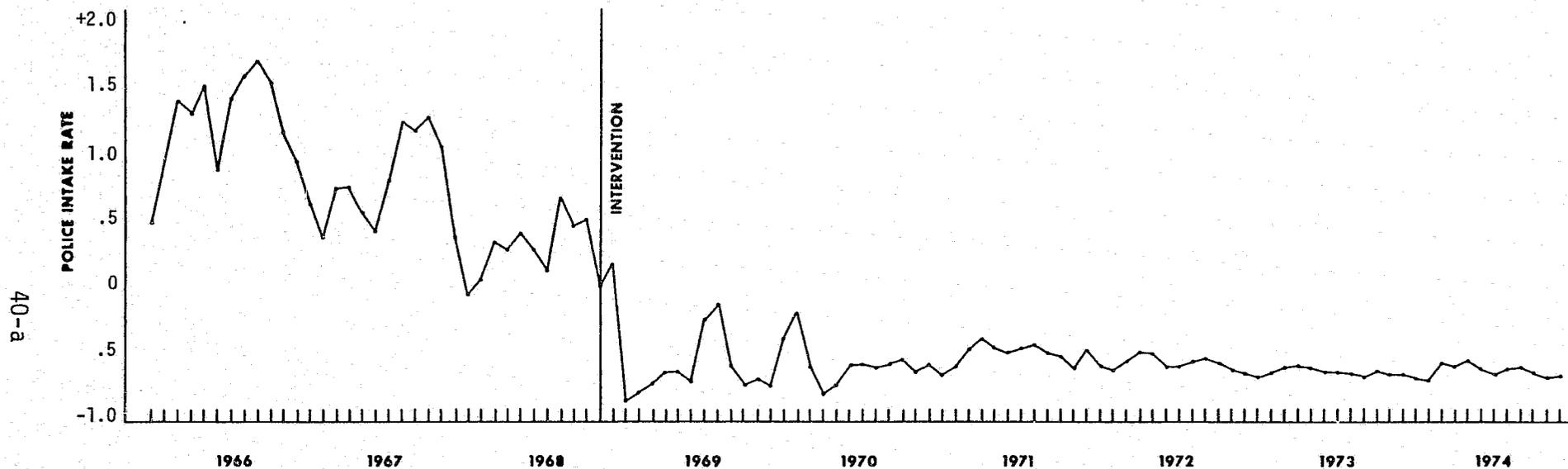
We obtained official police statistics from the Minneapolis Police Department to probe this assertion, focusing on disorderly conduct and vagrancy. The findings displayed in graphs 8 and 9 strongly indicate that the police are utilizing disorderly conduct to illegitimately arrest public inebriates. While vagrancy has shown a steady decline since 1960, the use of disorderly conduct has significantly increased<sup>39</sup> since decriminalization. From 1960 to 1966, the yearly average for disorderly arrests was 697 while during the transitional period,<sup>40</sup> this average increased to 1,167. Since decriminalization (1971-1975) the yearly average has jumped to 1,875. Thus, probably in response to the problem of keeping the streets clear of public inebriates,<sup>41</sup> and due to the overcrowding at the Alcoholism Receiving Center, the Metropolitan Police Department has become increasingly dependent on disorderly conduct as a reliable means of disposition.

## CONCLUSION

Our multiple-time-series analysis does confirm a statistically significant decline in the number of public inebriates formally handled by the police in the manner designated by the "law in the books." This finding does raise serious doubts about the use of police to carry out decriminalization policy. However, our comparative analysis does not necessarily lead to a conclusion that more public inebriates are being left on the street since decriminalization. As revealed in our micro analyses of the experimental jurisdictions, special ameliorative administrative action on the part of the public health community (e.g., use of civilian intake van, encouragement of self-admissions) does compensate for reduced police attention. Also, we found that in a jurisdiction that expects the streets to be kept clear of public inebriates, the police may find avenues of dispositions (e.g., the use of disorderly) that are less than legal.

From a methodological perspective, our study demonstrates the strengths of the interrupted time-series quasi-experiment for testing the impact of legal mandates on agencies responsible for implementation. We also establish the importance of "micro analysis" for tracing a series of plausible rival hypotheses and alternative administrative dispositions that cannot be controlled for in comparative analysis. In short, policy-impact analysis requires a design that can both reveal broad trends and scrutinize the unique responses of individual jurisdictions. Such an approach is especially critical if one hopes to use research results to prompt refinement of policy decisions.<sup>42</sup>

GRAPH 1.--Monthly police intake rates for public intoxication:<sup>a</sup>  
Washington, D.C.<sup>b</sup>

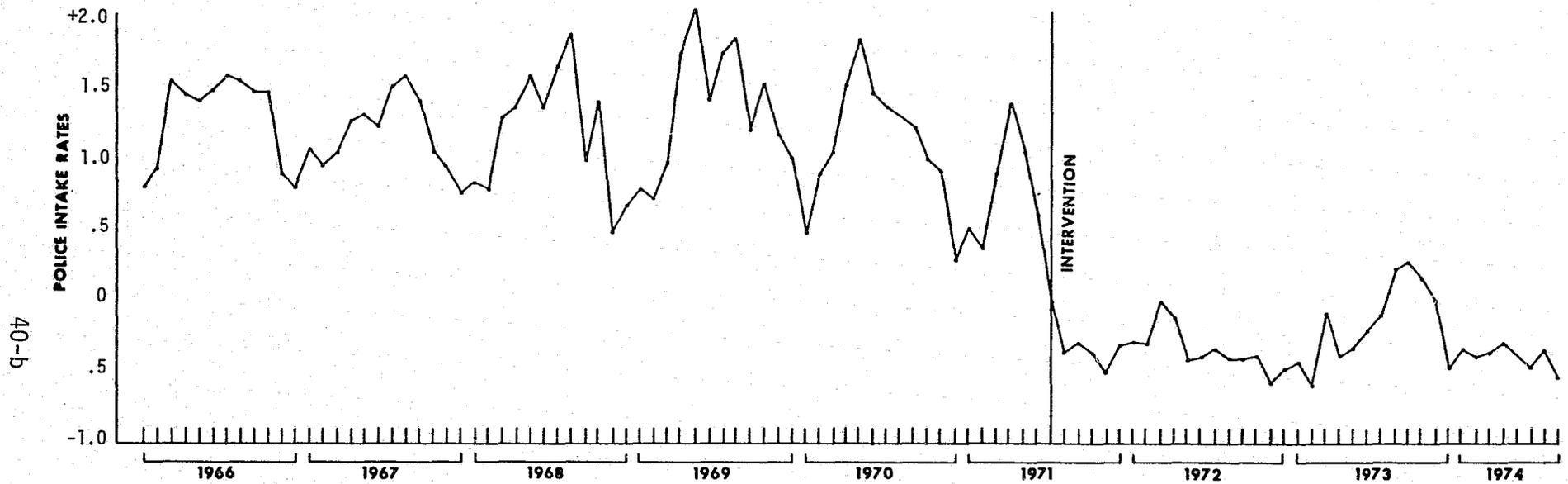


<sup>a</sup>Based on official statistics of Metropolitan Police Department, Washington, D.C. and official records of the D.C. Detoxification Center.

<sup>b</sup>Point of intervention--November 1, 1968.

\*Arrests and deliveries to Detox.

GRAPH 2.--Monthly police intake rates for public intoxication.<sup>a</sup>  
Minneapolis, Minnesota<sup>b</sup>

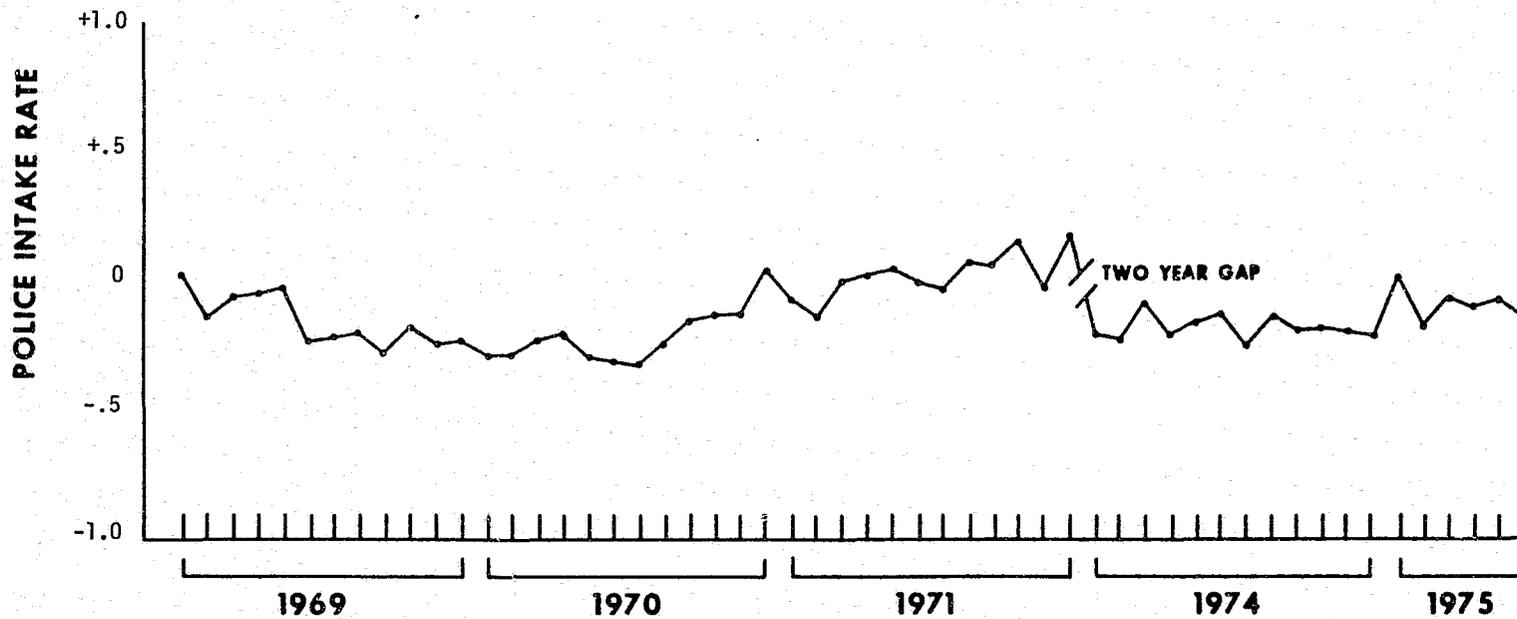


<sup>a</sup>Based on official statistics of Minneapolis Police Department, Minneapolis, Minnesota, and monthly intake statistics, Alcoholism Receiving Center.

<sup>b</sup>Point of intervention--July 1, 1971.

\*Arrests and deliveries to Detox.

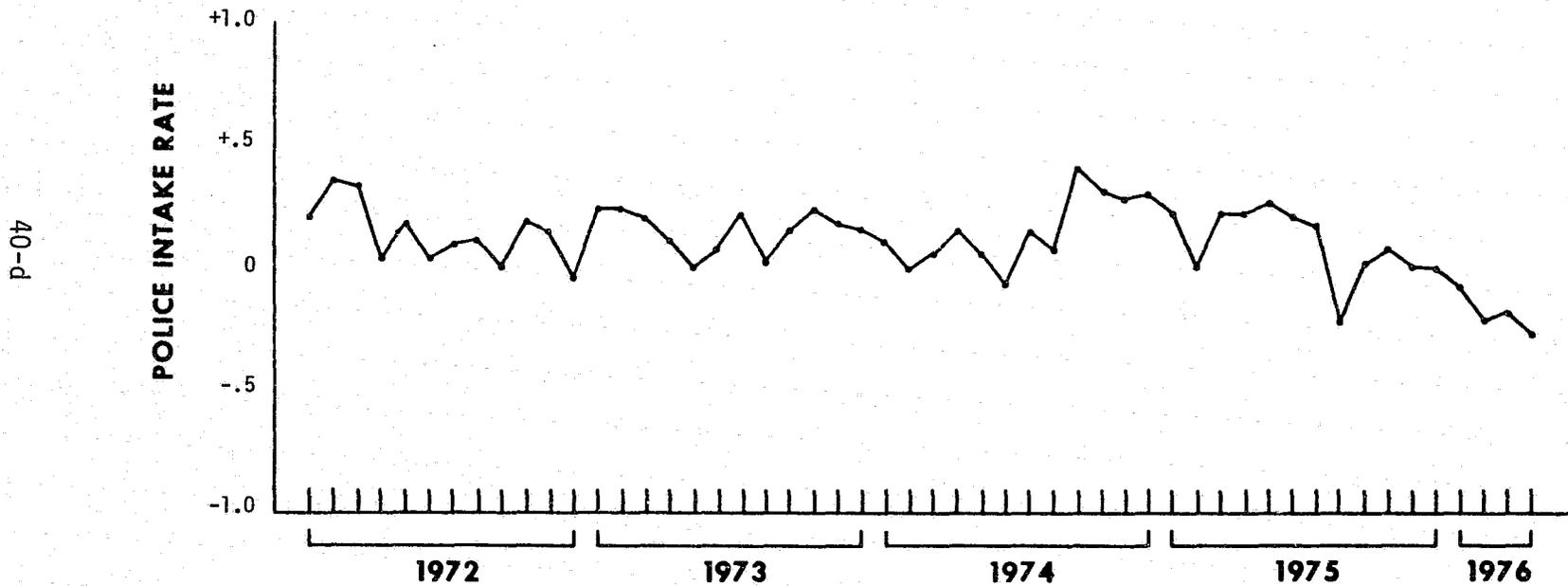
GRAPH 3.--Monthly police arrests for public intoxication:<sup>a</sup>  
Houston, Texas<sup>b</sup>



<sup>a</sup>Based on official statistics of Houston Police Department, Houston, Texas.

<sup>b</sup>Point of intervention--July 1, 1970.

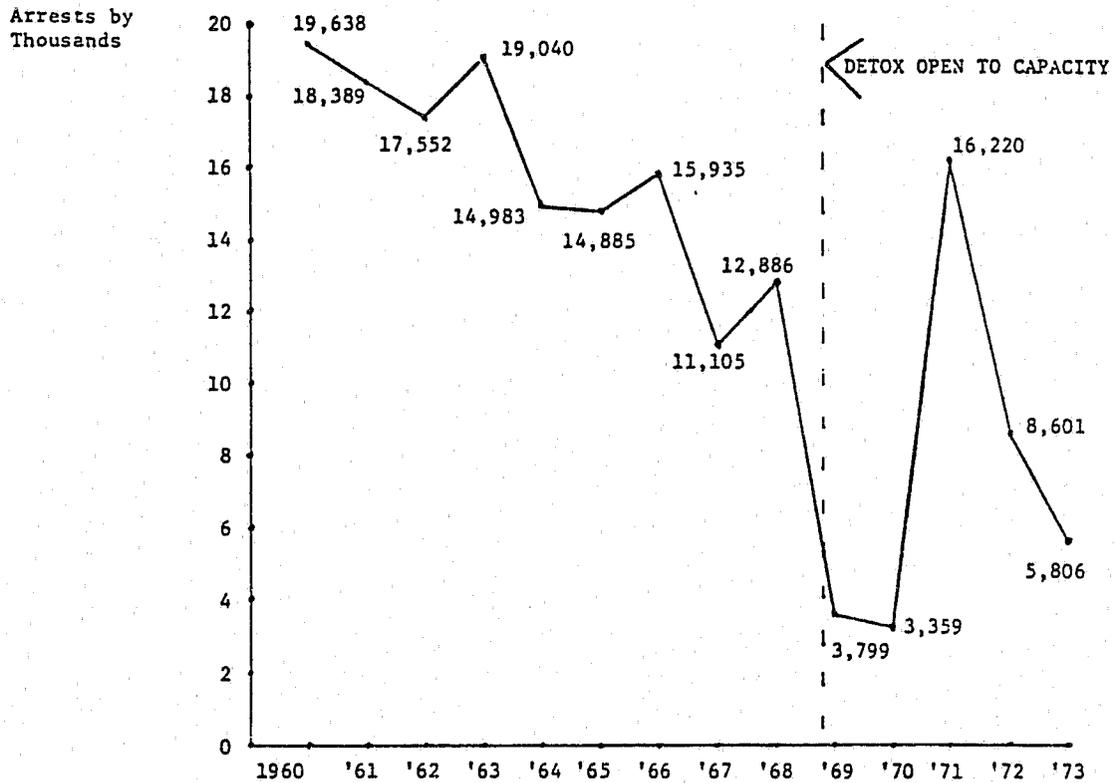
GRAPH 4.--Monthly police arrests for public intoxication.<sup>a</sup>  
San Francisco, California<sup>b</sup>



<sup>a</sup>Based on official statistics of San Francisco Police Department, San Francisco, California.

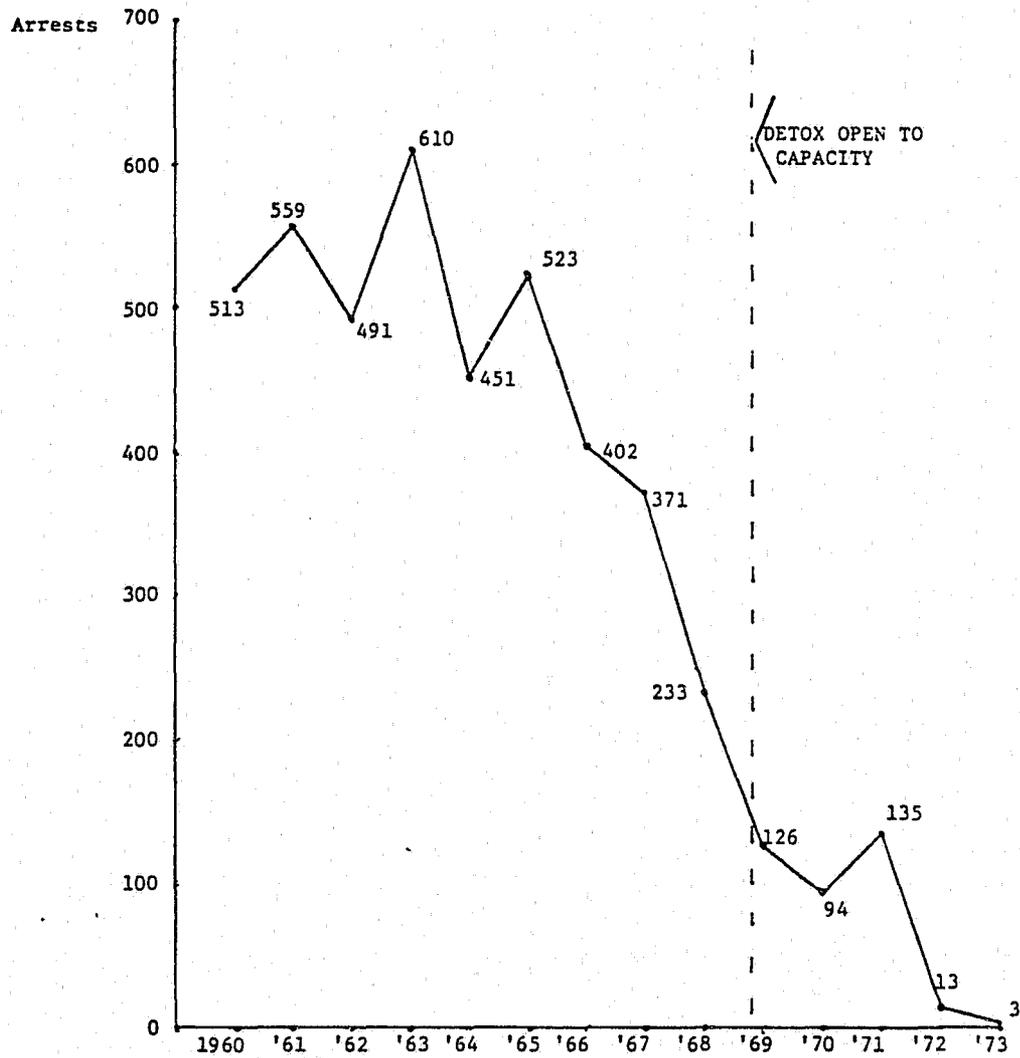
<sup>b</sup>Point of intervention--July 1, 1973.

GRAPH 5.--Disorderly conduct arrests,<sup>a</sup> District of Columbia,  
fiscal years 1960-1973



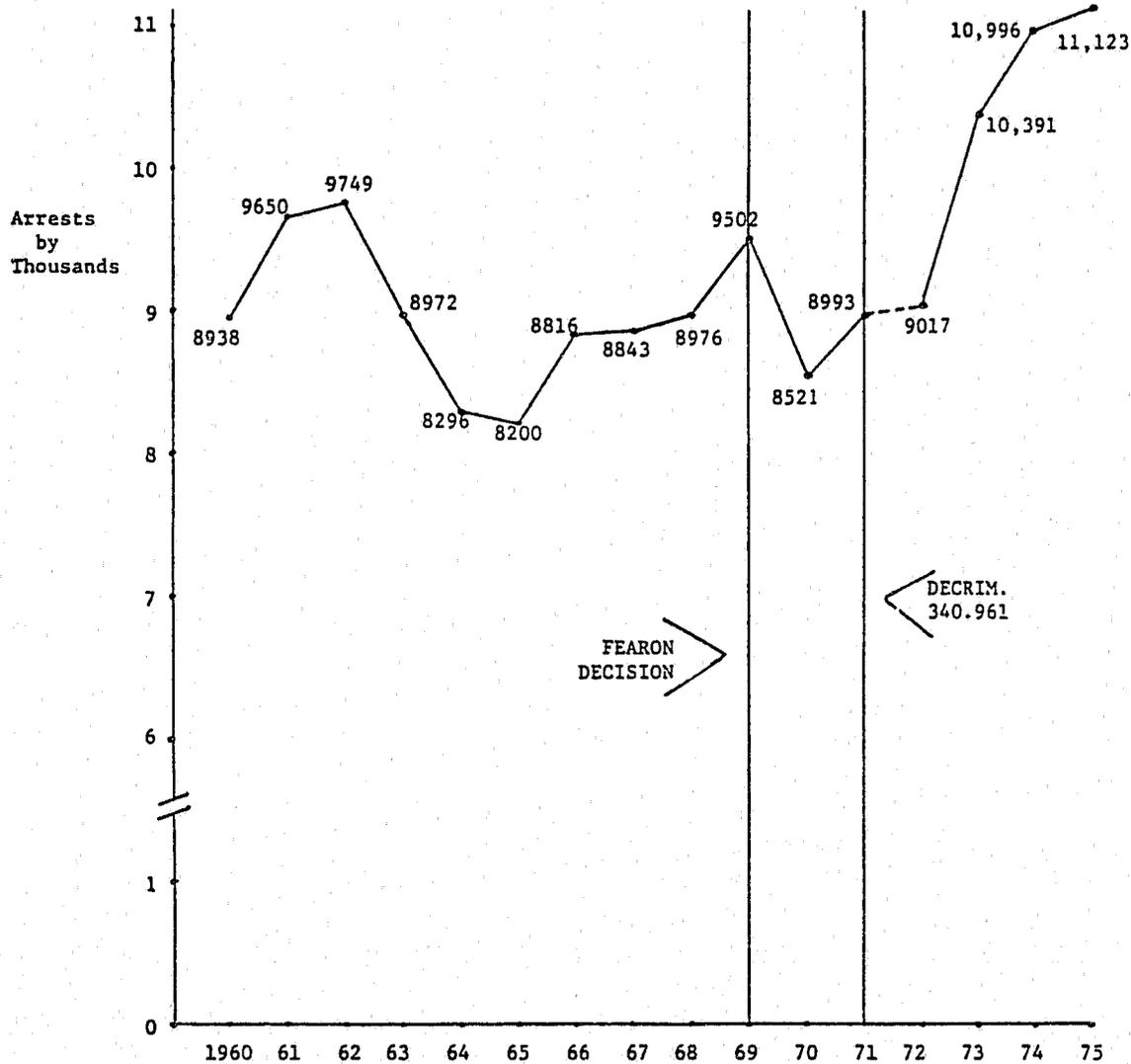
<sup>a</sup>Figures are official statistics of Metropolitan Police Department, Washington, D.C. Annual Reports, 1960-1973.

GRAPH 6.--Vagrancy arrests,<sup>a</sup> District of Columbia,  
fiscal years 1960-1973



<sup>a</sup> Figures are official statistics of Metropolitan Police Department, Washington, D.C. Annual Reports, 1960-1973.

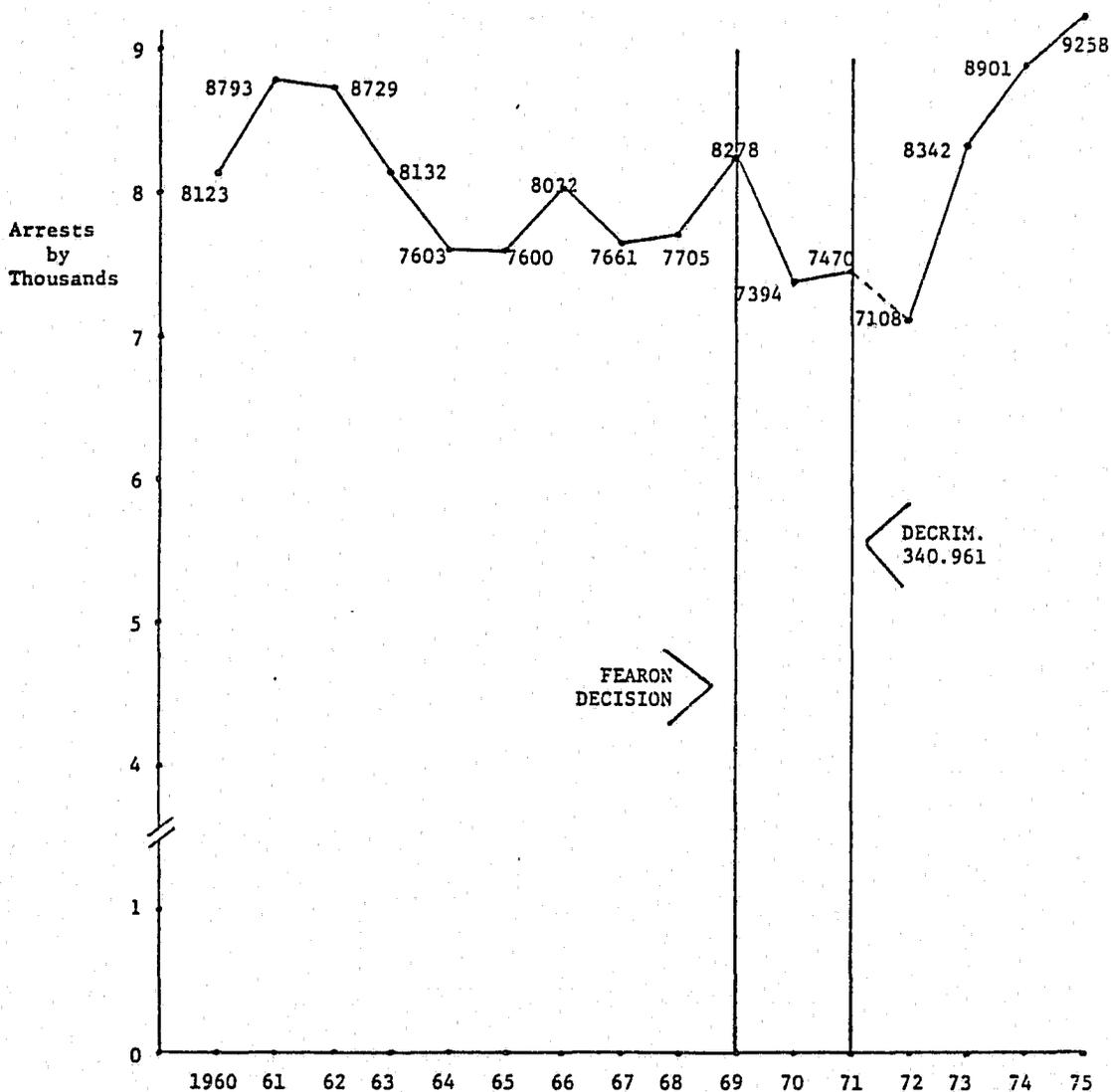
GRAPH 7. --Public drunkenness arrests, disorderly arrests, vagrancy arrests<sup>a</sup>, and all admissions<sup>b</sup> to the Alcoholism Receiving Center, Minneapolis, Minnesota, 1960-1975



<sup>a</sup> Figures are total yearly arrests, Official Statistics of Minneapolis Police Department, Annual Reports, 1960-1975.

<sup>b</sup> Figures are all police deliveries, civil pick-ups, self-admissions, and other means of intake, from Monthly Intake Comparison Statistics, Alcoholism Receiving Center, 1971-1975.

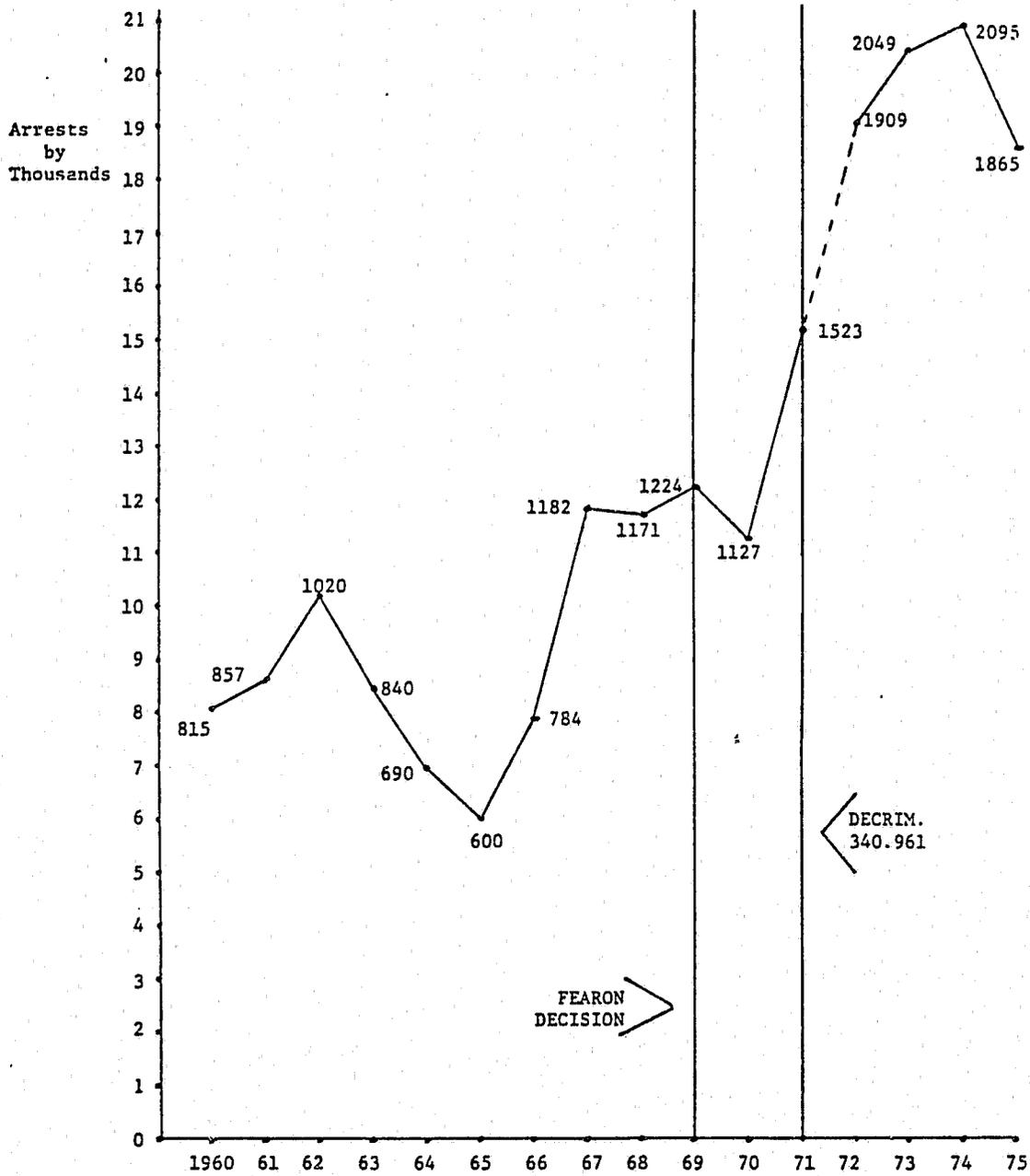
GRAPH 8.--Public drunkenness arrests<sup>a</sup> and  
all referrals to Alcoholism Receiving Center<sup>b</sup>,  
Minneapolis, Minnesota, 1960-1975



<sup>a</sup> Figures are total drunkenness arrests, Official Statistics of Minneapolis Police Department, Annual Reports, 1960-1975.

<sup>b</sup> Figures are all police deliveries, civil pick-ups, self-admissions, and other means of intake, from Monthly Intake Comparison Statistics, Alcoholism Receiving Center, 1971-1975.

GRAPH 9.--Disorderly conduct and vagrancy arrests combined,  
Minneapolis, Minnesota, 1960-1975



<sup>a</sup> Figures are yearly statistics, Official Statistics of the Minneapolis Police Department, Annual Reports, 1960-1975.

## NOTES

1. In the mid-1960's, three prestigious commissions (the U.S. and District of Columbia's Crime Commissions and the cooperative Commission on the Study of Alcoholism) rejected the criminal approach to public drunkenness and recommended the substitution of a public health approach. In 1969, the American Bar Association and the American Medical Association collaborated on model legislation for divesting public intoxication of its criminal status. In 1971, the National Conference of Commissioners on Uniform State Laws drafted model legislation for decriminalization--the Uniform Alcoholism and Intoxification Treatment Act. In Washington, D.C., the Washington Area Council on Alcoholism and Drug Abuse worked toward decriminalization throughout the 1960's and in Minneapolis, Minnesota, a similar group worked as members of the Minnesota Council on Alcohol Problems.
2. For an elaboration on the role of "policy subsystems" in the policy formulation process, see, e.g., A. L. Fritschler, *Smoking and Politics* (1969); J. L. Freeman, *The Political Process* (1965).
3. See, e.g., D. C. Perry, *Police in the Metropolis* 23-57 (1975).
4. Departments have often given credit for such arrests much in the same way they award credit for making other misdemeanor and traffic arrests. Former Police Chief of Washington, D.C., Jerry V. Wilson, discusses the importance of this incentive. See J. V. Wilson, *Executive Control of Policies for Police Handling of Public Inebriates*, 1975 (unpublished paper from the American University).
5. K. C. Davis, *Police Discretion* (1975).
6. J. Q. Wilson, *Varieties of Police Behavior: Management of Law and Order in Eight Communities* 85-89 (1971).
7. See J. Levine, M. Musheno, and D. Palumbo, "The Limits of Rational Choice in Evaluating Criminal Justice Policy" in *Policy Studies and the Social Sciences* 94-99 (S. Nagel ed. 1975).
8. For a discussion of these common threads, see T. Dye, *Understanding Public Policy* 291-296 (1972).
9. Studies of this genre include D. T. Campbell & H. L. Ross, "The Connecticut Crackdown on Speeding: Time-Series Analysis Data in Quasi-Experimental Analysis," 3 *Law & Society Rev.* 33 (1968); G. V. Glass, "Analysis of Data on the Connecticut Speeding Crackdown as a Time Series Quasi-Experiment," 3 *Law & Society Rev.* 55 (1968); G. V. Glass, G. Tiao, & T. Maguire, "The 1960 Revision of German Divorce Laws: Analysis of Data as a Time-Series Quasi-Experiment," 5 *Law & Society Rev.* 539 (1971); F. E. Zimring, "Firearms and Federal Law: The Gun Control Act of 1968," 4 *J. of Legal Studies* 133 (1975); and H. L. Ross, "The Scandinavian Myth: The Effectiveness of Drinking-and-Driving Legislation in Sweden and Norway" 4 *J. of Legal Studies* 258 (1975).

10. For an early study that contributed to the expansion of public law research beyond the workings of the Supreme Court, see K. Dolbeare, *Trial Courts in Urban Politics* (1967).
11. For other similar works, see N. Milner, "Comparative Analysis of Patterns of Compliance with Supreme Court Decisions: Miranda and the Police in Four Communities," 5 *Law & Society Rev.* 119 (1970); E. Ostrom, et al., "Community Organization and the Provision of Police Services," *Sage Professional Papers on Administrative and Policy Studies* 1 (1973); R. Medalie, et al., "Custodial Police Interrogation in Our Nation's Capital: An Attempt to Implement Miranda," 66 *Mich. L. Rev.* 1347 (1968).
12. *Easter v. District of Columbia*, 361 F.2d 50 (D.C. Cir. 1966), and *State v. Fearon*, 166 N.W.2d 720 (1969).
13. In Washington, D.C., the D.C. Alcoholic Rehabilitation Act of 1976, P.L. 90-452, 82 Stat. 618 (1968) retained the police as the legal instrument for removing intoxicated persons from the streets, but the MPD was to pick up "patients" under a public health provision which reads:

"Except as provided in subsection (b) of this section, any person who is intoxicated in public: (1) may be taken or sent to his home or to a public or private health facility; (2) if not taken or sent to his home or such facility under paragraph one shall be taken to a detoxification center."

In Minnesota, the legislature ended the criminal processing of public drunkenness by repealing 340.96 and passing 340.961. The latter provision provided that drunkenness was not a crime, and repealed the municipal ordinances prohibiting public intoxication. As of July 1, 1971, this enactment left law enforcement personnel with only the provisions of the Hospitalization and Commitment Act when encountering a drunken person in public (section 253A.04):

"(a) take the person into 'custody' and transport him to a facility equipped to treat alcoholism and provide for emergency care or treatment (72 hour limit to involuntary treatment); or

"(b) take the person home if he is not endangering himself, other people or property; or

"(c) leave the person where he is found."

14. D. T. Campbell and J. C. Stanley, *Experimental and Quasi-Experimental Designs for Research* (1966).
15. B. V. Glass, V. L. Willson, and J. M. Gottman, *Design and Analysis of Time-Series Experiments* (1975).
16. By "high arrest jurisdiction," we mean a jurisdiction whose police department has given high priority to the public drunkenness offense by making a large number of arrests over time.

17. By "moderate arrest jurisdiction," we mean a jurisdiction whose police department has given only limited priority to the public drunkenness offense by making a relatively low number of arrests over time.
18. Lempert, "Strategies of Research Design in the Legal Impact Study: The Control of Plausible Rival Hypotheses," 1 Law and Society Rev. 121 (1966).
19. Observation requirements for sophisticated analysis are discussed. See n. 15, supra.
20. Fortunately, Professor V. Glass of the University of Colorado has developed a computer program, CORREL, which computes autocorrelations and partial autocorrelations for raw data. CORREL also includes a seasonal option for identifying cyclic series. He applied his program to our data for Washington, D.C. and Minneapolis, Minnesota. The data were analyzed as a  $p=0$ ,  $d=1$ ,  $q=1$  (integrated moving averages) with a seasonal component (cycle = 12). For Washington, D.C., this analysis produced a  $T=3.20$ , significant at .001 with 106 degrees of freedom.
21.  $T = -4.84$ , significant at .001 with 102 degrees of freedom.
22. Professor Glass advised and we concurred that visual scanning of the control jurisdictions' data in graphs 3 and 4 adequately establishes that no similar effect is taking place in these criminal jurisdictions.
23. The "law on the books" in Minnesota does grant broad discretionary powers to the police by adding a final approved option--"leave the person where he is found." Hospitalization and Commitment Act, section 253A.04.
24. See D. Aaronson, C. T. Dienes, and M. Musheno, Final Report--Project on Public Inebriation, 1976 (unpublished Law Enforcement Assistance Administration Grant Report #74NI-99-0055).
25. Based on Jellinek Formula as calculated and reported by Dr. D. Mindlin, Director of Adams Mill Alcoholism Center, Washington, D.C.
26. Since police have no record of the number of individuals they processed for this charge in the pre-ARA period, court records (The D.C. Court of General Sessions Index) listing cases for each calendar year in alphabetical order by individual name were used. More individuals with multiple arrests would be processed in the courts while the more affluent single offenders would forfeit their collateral rather than be exposed to the court process. Therefore, this bias of the estimate runs counterproductive to our research hypothesis.
27. Post-ARA population statistics exist on the number of individuals admitted to Detox for each post-ARA year.
28. Mr. Robert Olander, Research Sociologist for the Department of Mental Health, Mental Retardation, and Chemical Dependency, applied the standard Jellinek Formula to the mean of the yearly census figures of Hennepin County's adult population from 1965 to 1970 as a way of estimating the size of the potential problem-drinking population during the criminal era. He found a yearly average of 37,346 potential problem drinkers for this

- period. For the decriminalized era (1971-1975), he found a yearly average of 38,390. This finding is strengthened by the fact that between 1971 and 1975, Hennepin County registered a slight decrease in population.
29. See D. Aaronson, C. T. Dienes, and M. Musheno, Final Report--Project on Public Inebriates, 1976 (unpublished Law Enforcement Assistance Administration Grant Report #74NI-99-0055).
  30. Interview with Mr. Leonard Boche, Director, Department of MH/MR/CD, June 3, 1975.
  31. Hennepin County Alcoholism Receiving Center, "The Public Inebriate: An Innovative Approach to the Transporting of Clients to a Detoxification Center," 1975 (unpublished paper presented to North American Congress on Alcohol and Drug Problems).
  32. Interview with Mr. Paul Thorne, Director, Alcoholism Receiving Center, June 5, 1975.
  33. Only yearly data are available:  $T = .16$ ,  $df = 11 + 5 - 2 = 14$ ,  $p = N.S.$  Thus, there is no significant difference in pick-ups between the two periods when one adds the intakes generated by the efforts of the Alcoholism Receiving Center's staff.
  34. "The Public Inebriate: An Innovative Approach to the Transporting of Clients to a Detoxification Center," 1.
  35. Id., 2.
  36. Id., 4.
  37. Id., 4.
  38. Interview with Mr. Leonard Boche, June 3, 1975.
  39.  $T = 2.61$ ;  $df = 14$ ;  $p = .02$ .
  40. Transitional Period: Pre-Court Screening to Decriminalization: 1967-1970.
  41. Urban renewal has displaced the chronic skid-row population and therefore, they are more often seen in and around the thriving downtown commercial and business district. Their appearance in this area increases community pressure on the police and the civilian intake van to keep the streets clear of public inebriates. Based on interviews with Mrs. Meredith Hart, League of Women Voters, Minneapolis, Minnesota, July 3, 1975, and Sgt. Robert Havenstein, Planning and Research, Minneapolis Police Department, Minneapolis, Minnesota, June 3, 1975.
  42. Such a goal is proper for policy analysis research. See, e.g., J. S. Coleman, "Problems of Conceptualization and Measurement in Studying Policy Impacts," in Public Policy Evaluation 21-26 (K. Dolbeare ed. 1975).

## THE CHALLENGE OF PATROL PERFORMANCE EVALUATION IN THE STUDY OF POLICEWOMEN

Joyce L. Sichel, Ph.D.  
Vera Institute of Justice

Women on Patrol: A pilot study of patrol performance in New York City has been completed recently by the Vera Institute of Justice (Vera) and the New York City Police Department (NYPD). In this effort to gain additional knowledge about women in policing, the research staff had to confront the issue of evaluating patrol performance. Without adequate performance indicators, it is inappropriate to report on police patrol performance, whether by male or female officers; further, the need for unemotional, objective comparison of the sexes demands performance indicators which are reliable and unbiased. As the Chief of Operations in the Miami, Florida Police Department reported recently to the International Association of Chiefs of Police, ". . . the more we question the utilization of policewomen, the more we must question the validity of the criteria used to determine the performance of our male officers."<sup>1</sup>

It was the goal of Vera-NYPD study staff to select the best measuring tools available, and to develop new ones as needed in those areas where it was thought that the performance of men and women might differ--for example, physical activity, violence, and human relations.

### MEASURES WHICH WERE AVAILABLE

Traditionally, measures of patrol outcome have been used to assess patrol officers' performance. Especially popular have been the so-called "productivity" measures, such as number of arrests made, summonses issued, days worked. Among the advantages of these output counts are low cost, simplicity of collection, intuitive appeal, and face validity because of their concreteness. Their disadvantages include their questionable relevance to quality in patrolling and the measures' dependence on such arbitrary external factors as political and administrative pressures and changes in recordkeeping methods.

Also traditional have been rating scales for the judgment of police performance. For example, a study of women State Traffic Officers hired by the California Highway Patrol<sup>2</sup> relied almost entirely on ratings by trainers and supervisors for the measurement of field performance. Another case of reliance on supervisors' ratings was in the development of a police selection test by Educational Testing Service.<sup>3</sup> ETS staff used ratings as the performance criterion against which the test was validated. However, they found that supervisors showed rating biases according to the race and education of the officer being rated. In later work with firemen, they attempted to assess and correct for such bias through adjusted ratings.

While rating bias against officers because of their sex was not shown in the Vera-NYPD study, some police supervisors probably show bias in their rating of female officers; any such bias could be assessed and biased ratings corrected in a manner similar to ETS's by those municipalities depending upon ratings for their measure of performance.

Other research with rating scales has shown a strong tendency for supervisors to rate most officers close to the mean, avoiding extremes. As Gary Marx of M.I.T.<sup>4</sup> pointed out, the ratings may ". . . become empty rituals where almost everyone's performance is rated as satisfactory." Also, supervisors tend to rate all of an officer's attributes similarly so that ratings of specific qualities may not be valid.<sup>5</sup> An effort to develop well-defined, reliable scales for the assessment of police performance has been made by Frank Landy at Pennsylvania State University.<sup>6</sup> He uses behavioral anchors and provides careful instructions to the raters, who are either peers or supervisors of the officers; he finds they are able to use the scales in a reliable fashion.

The rating method at its best is highly appealing because even elaborate scaling is relatively economical and simple to use. However, its abiding problem is its extreme subjectivity, and when women and minorities are to be evaluated, it is essential that even the most refined scaling methods take rater bias into account.

An indirect technique for measuring performance has been test scores. Many researchers have collected and analyzed scores on physical fitness, fire-arms proficiency, driving, and academic subjects related to policing. However, unless these tests are directly related to activities which are either frequent or important to the patrol job, their value as performance indicators is unclear.

The New York State Trooper selection program<sup>7</sup> represents one effort to make testing highly relevant to performance. Potential for performing as a trooper is assessed through written items and through trial of physical ability. The tests were constructed by applying the "job element method" developed in the field of personnel management. General requirements which were judged by peers and supervisors to be specific to the trooper's job were converted into specific testing items of both written and active nature. For example, changing a car tire was used as an active test for ability to use physical resources. Such testing has the benefit of avoiding discriminatory selection criteria which might be irrelevant to performance, such as height and weight. The New York State system instead asks whether an applicant can see a person over the roof of a standard sedan and crawl through a 3' x 3' space. This kind of system also allows unequal weighting to be given to aspects of job performance; for example, being able to distinguish different automobile colors can be weighted more or less heavily than being able to see over a car roof.

The major problem with this kind of system is, of course, the quality of the transition from performance concept to test item. Moving from abstract definitions of job skills to concrete operations for their measurement is always difficult and sometimes much is distorted or lost in the translation. However, this is true not only in constructing test items as a means of performance measurement, but also for developing specific indicators to count, inventory, or observe in other modes of performance evaluation.

The observation of actual patrol is a performance evaluation method which has been less common.<sup>8</sup> Albert Reiss was an early user of observation for research evaluation of patrol in Boston, Chicago, and Washington, D.C. The Washington, D.C. policewomen study, Policewomen on Patrol, also employed observers to ride in the patrol cars to watch men and women performing patrol. Lewis Sherman in St. Louis and Harriet Connolly and Judith Greenwald in New York City also applied the technique to studies of policewomen. These studies have generally developed standardized observation materials, rather than relying on informal or anecdotal reports. The content of the materials has tended to be subjective, relying on observer judgment of performance--such as degree of officer aggressiveness--for performance measures.

A more objective observation instrument was developed and employed by Cruse and Rubin in Miami, Florida.<sup>9</sup> Five-point scales were used by observers to describe the degree to which an officer showed controlling, counseling, assisting, sympathizing, threatening, suspicion, and humor.

Another less traditional mode which has been used is performance evaluation by the clients of police service. The Kansas City Preventive Patrol Experiment<sup>10</sup> solicited citizen opinions about police service. So did the studies of policewomen in Washington, D.C. and St. Louis.

Finally, arrest quality measurement has recently emerged in counterpoint to arrest productivity counts. Studies of policewomen have looked at the ratio of convictions to arrests, reasoning that this taps the officer's legal knowledge and judgment in making the arrest, wording the complaint, and giving testimony in court. Also, the rate at which arrests are dismissed in the initial screening by a district attorney has been used as an indicator of poor arrest quality. However, these disposition measures have been criticized<sup>11</sup> for being susceptible to many influences outside of police control, and so they are probably most valuable for comparing arrests made at approximately the same time and within the same jurisdiction.

#### THE APPROACH ADOPTED FOR PATROL PERFORMANCE MEASUREMENT

It was apparent to the Vera-NYPD staff that many existing patrol performance approaches and measures were appropriate and useful in a study of policewomen on patrol. It was also clear, however, that there was a need to refine and expand patrol performance measurement. The present study attempted to combine, enhance, and invert performance measures as needed to develop a comprehensive picture of women's patrol performance. The approach finally developed was an observational system which stressed "controlling" activity as a pivotal aspect of patrol performance.

Evaluating the Process of Patrol through an Observational System. Gary Marx has pointed out that the process of policing--the means for reaching patrol outcomes--has been a neglected dimension in performance evaluation. There have only been a few efforts in this direction; these have been the research projects which have used observation as a tool for performance measurement. The Vera-NYPD study decided to build upon these first-hand observation efforts to focus on a description of how men and women carry out the patrol process. Of the various methods tried, this proved to be the most productive for performance assessment.

On-the-job activity was recorded and classified by specially trained observers, both police personnel and civilians, using structured observation forms with an original numerical behavior-coding system to describe the types and frequencies of behaviors being performed by the officers under study. The forms developed are part of the appendix to this paper. This observation system allowed answers to many questions of importance to the study of women on patrol. For example, how active the women were in comparison to male officers could be measured by counting the number of behavior codes recorded. Whether the kinds of activity the female officers showed on patrol differed from male officers could be assessed by the frequencies of various behaviors coded. Through this kind of analysis, for example, it was learned that the N.Y.C. women performed many kinds of activities as frequently as male officers, including service activities for citizens, and requesting help from other officers; women did less searching of premises and vehicles, and were less apt to "pick up" street incidents and "back up" patrol cars already on jobs. The system yielded additional data revealing that women officers did more sympathizing and comforting of citizens than did men.

It was also possible to observe that men and women did not differ in their self-control on patrol. In an attempt to be as objective as possible, observers were instructed to use physical indicators for emotional arousal. The criteria for an officer being emotionally "heightened" included his or her voice being much louder than normal or having hands that were shaking. For an officer to be recorded as "out of control" that officer either used exaggerated physical gestures or shouted very loudly or became very red in the face. Stressing specific features which observers should note led to good consistency in use of the observation forms, and to a higher degree of interobserver agreement than was present when more impressionistic judgments were called for.

Another example of the kind of question which the observation system could address was what kinds of strenuous physical activity men and women officers did while on patrol. A separate observation form was used to record the occasions when any officer on the scene did something strenuous. Who participated and did not, as well as what they did, was recorded for tabulation and analysis. Unfortunately, findings were limited by the rarity of these events.

A New Patrol Measure: Controlling. Probably the most important way in which the present study's observation system contributed to performance measurement was in its ability to trace the incidence and progression of "control-seeking" behavior in officer-citizen encounters.<sup>12</sup> Control-seeking may be defined as the attempt to influence another person or persons to take a particular action. Control-seeking behavior usually take the form of verbal assertions, but is sometimes physical and occasionally violent; in this study, observers were instructed to recognize control-seeking activity as one of a set of specific behaviors ranging from ordering, requesting, and threatening to applying physical restraints and firing a weapon. A full list of the behaviors defined as control-seeking is presented in figure 1.

The control-seeking approach circumvents problems associated with measuring performance in violent situations, an area of particular concern to those involved in evaluating the performance of policewomen. Control-seeking efforts occur with far greater frequency than do violent events. Specific behaviors rather than subjective impressions mark situations where tensions are heightened; for example, a citizen or officer may be observed to utter a threat,

FIGURE 1.--Behaviors defined as control-seeking

---

001=orders	043=by head
002=requests	044=by other body part
003=recommends	<u>*GRABS*</u>
004=reasons	045=hand/arm
005=makes deal	046=clothing
006=humors/flatters	047=hair
007=shames	048=leg
<u>*THREATENS*</u>	049=head
008=official action	050=other body part
009=physical force	051=another person's weapon
010=use of weapon	<u>*HOLDS*</u>
011=unspecified	052=hand/arm
	053=clothing
012=attempts to frisk	054=hair
013=attempts to search person	055=leg
014=attempts to search premises	056=head
015=announces arrest	057=other body part
016=attempts to serve D.A.T.	
017=attempts to serve summons	058=slaps
018=attempts to handcuff	059=punches
019=attempts to place in car	060=wrestles
	061=kicks
020=chases on foot	062=knees
021=chases in car	063=flips
	064=hurls
022=stands above	065=hurls object at
023=confronts eye-to-eye	066=douses
024=positions body to block	067=bites
025=orders by gesture	068=shakes
026=taps for attention	069=twists arm
027=prods by light touch	<u>*GRABS FOR*</u>
028=leads	070=baton
029=leads by hand	071=gun
	072=knife
030=rushes toward	073=other weapon
031=pounces on	<u>*SHOWS*</u>
032=pins to ground/wall	074=baton
033=sits on	075=gun
034=stands on	076=released gun
<u>*PUSHES*</u>	077=knife
035=with hand	078=other weapon
036=with foot	<u>*PREPARES TO USE*</u>
037=with body	079=baton
038=with baton/gun	080=gun as baton
<u>*PULLS*</u>	081=gun
039=by hand/arm	082=knife
040=by clothing	083=other weapon
041=by hair	<u>*USES*</u>
042=by leg	084=baton

FIGURE 1.--Behaviors defined as control-seeking--Continued

085=gun as baton  
086=gun  
087=knife  
088=other weapon

090=fingerprints  
091=places in cells  
092=strip-searches

---

twist an arm, or reach for a weapon. And because control-seeking behaviors and their outcomes can be described with precision, officer actions that lead to violence can be identified.

Control-seeking behaviors sometimes occurred in sequences, forming patterns of controlling. For example, an officer at a crowd scene was observed to order citizens and then to push them and later to arrest them. Under similar circumstances, other officers issued many orders, but did not progress to the more extreme means of control. The observation data from this system provide a good picture of such individual variation. An example of control-seeking during an incident observed in this study is described in figure 2.

Observers were asked to specify the objective for which they judged each control attempt to have been made, and then to note whether the objective was met. Thus, as figure 2 shows, the control-seeking model allowed evaluation as well as description of performance of individual officers or groups of officers. As in the case of physical activities, a separate observation form was used for the recording of control efforts.

Other Modes of Process Measurement--Citizen Interviews and Departmental Arrest Reports. Information about the quality of the policing process came from two additional sources--interviewing with citizens who had been victims of crimes or complainants in a previous patrol contact with an officer under study, and examination of arrest records.

Citizen interviewing had been shown feasible in previous studies of patrol. In the present case, citizens were asked to describe an officer's performance and this was coded with the same system used by first-hand observers. This approach provided a check on the validity of the observation process. However, client accounts from memory of an event between 1 and 3 weeks in the past were often sketchy compared with the wealth of information generated from first-hand observation. Ratings of the officer from the point of view of the client proved to be a more useful aspect of the citizen interviews because citizens were able to respond more fully to this questioning. Citizens' opinions about the officer's respectfulness, pleasantness of manner, quality of listening and explaining, and emotional sensitivity were each coded on a five-point scale. Citizens rated women significantly higher than men on these scales.

Reports of arrests made by officers under study were examined closely to explore the process of policing in still another way. Such analysis has typically been limited to disposition in the prosecutor's office and courts, but in the Vera-NYPD study, as in many situations where performance evaluation is needed quickly, the time lag until disposition was too great. Also, it was

FIGURE 2.--Sequence of control-seeking

Background

Type of incident : Family dispute--husband has left premises  
 Officers at scene: Female subject officer and male partner  
 Citizens at scene: Female complainant, physically intact,  
 heightened emotional state evidenced by weeping

Actions directed toward complainant	Type of control	Citizen response/ effectiveness of control
Subject officer says, "Please tell me how it happened."	Request	Citizen continues to sob/ control attempt not successful
Partner says, "Calm down, now!"	Order	After some seconds, citizen stops crying, starts to explain/control attempt fully successful
Subject officer says, "You ought to get him to a clinic."	Recommendation	Citizen continues explanation/success of control attempt cannot be rated
Partner says, "You shouldn't stay here tonight."	Recommendation	None; however, citizen later leaves with officers/control attempt fully successful
Noncontrolling behavior: Subject officer says, "Do you have a friend or relative where you can stay over?"		
Citizen responds in affirmative.		
Everyone departs.		

decided to stress those aspects of the arrest process which were thought to be within an officer's control rather than at the discretion of prosecutors. Therefore, in consultation with officials in the NYPD and the project's advisory committee, this list of points which were believed to be related to varying quality of arrest was developed:

A. Importance of charge(s):

- (1) More major charges were assumed to reflect a better quality arrest, at least in New York City;
- (2) It was agreed that a string of trivial charges often reflected over-reaction on the part of the officer.

B. Larger context:

- (1) Community demands being met was generally considered a positive indicator (e.g., making a prostitution arrest in an area where community meetings reflected great concern with keeping prostitutes off the streets);
- (2) Precinct supervisory expectations being met were also considered a positive indicator, implying that the officer had not just "taken the law into his own hands" (e.g., making a narcotics arrest where the precinct command emphasized drug abuse control).

C. Evidence of degree of officer self-control:

- (1) Locking up a prisoner when a "desk appearance ticket" (station-house summons) would usually have sufficed was considered a negative factor in officer self-control;
- (2) If a suspect was hurt without apparent cause during apprehension, this was also considered a negative factor;
- (3) The bringing of charges against someone not in the initial role of suspect was considered an additional negative factor, suggesting possible officer escalation of conflict.

To assess the degree to which subject officer arrests measured up against these standards, arrest reports were analyzed for this content by experienced police personnel. In addition, a separate general evaluation of each report on a five-point scale was made by high-ranking police officials to validate the set of criteria.

These criteria did not turn out to be useful. First, the negative factors and negative evaluations occurred too infrequently to make discriminations between officers. Perhaps this reflects the rarity of a "bad" arrest, but it seemed more likely to be a deficiency of the system and perhaps an insufficiently critical attitude by the rating personnel involved. Also, if the system is to be reused, the meaning of "supervisory expectation" and "community demand" must be more carefully specified with examples.

A Problematic Attempt to Use Self-Ratings and Peer-Ratings for Performance Evaluation. A questionnaire was developed to be self-administered by patrol officers. Extensive pretesting and use with male police officers resulted in four six-item subscales, forming indices which appeared to have adequate internal consistency,<sup>13</sup> acceptability by the officers, and face validity for measuring officers' patrol self-evaluations. These covered the areas of: self-confidence about controlling citizens, self-confidence about performing strenuous physical activities; level of fear, and level of self-control. Self-confidence in one's competence to control others in the course of patrol was tapped by agreement with items such as "I can handle just about any person I come up against in my job," and disagreement with items such as "I feel much better when another officer is there to back me up when I give orders." Confidence in one's physical abilities was assessed by items such as, "I feel I would never have to use my gun to handle an unarmed citizen." The degree to which an officer experienced active fear while on patrol was assessed by an item such as, "It would take the most extreme danger to make me feel afraid." Officers' self-control was established by items such as, "I find it easy to ignore verbal taunts"; the absence of such control by items like, "There is a certain type of person who brings out the worst in me on the street." The full scales are included in the appendix to this paper.

This kind of information, interesting in itself, can also help provide reasons for differences in performance discovered through observation. Unfortunately, it proved difficult to collect this self-report information from the New York City policewomen under study. Several of the women told project staff that they feared they could damage the cause of women in policing if they gave "wrong answers." In retrospect, it would have been preferable to utilize self-ratings in the diagnosis of training needs among all relatively inexperienced officers, including women, rather than to single out policewomen for mandatory self-revelations which they felt to be potentially self-incriminating.

Performance descriptions by police officer peers were collected but, as in the case of self-ratings, it was impossible to obtain complete and honest information. It was found that partners of the officers being studied were unwilling to provide ratings of the officer, even in an indirect way during an interview. The informal code, as well as overt union policy, discouraged any reporting on a peer's competence in this police department.

Further, in studies of policewomen the use of peer ratings is complicated by the problem of bias on the part of male peers.

#### REFLECTIONS AND SUGGESTIONS FOR OTHERS CONTEMPLATING SIMILAR PERFORMANCE ASSESSMENT

The most valuable indicators of patrol performance in the Vera-NYPD study derived from first-hand observation. For these research purposes the observers worked in police-civilian pairs and observed full 8-hour tours of duty. For nonresearch applications, specially trained police, riding singly for a few hours at a time, probably could perform the same kind of observation at less cost, providing they ride at times of high patrol activity. Based on N.Y.C. experience, officer controlling occurred in one patrol incident out of four. Therefore, to include controlling as a performance dimension to be eval-

evaluated would require substantial observation time. However, since differences in this kind of activity seem to be at the heart of policing success, the investment may be worthwhile.

The control model's usefulness could be extended by taking into account joint patterns of controlling by the officers and their partners, where teams patrol. Especially with experienced officers who have worked frequently together, an officer's controlling may be deliberately tailored to fit with that of his or her partner. Thus, one officer may cajole and humor a suspect because it has proven effective in concert with threats made by the partner.

It appeared that requiring highly specific observations, such as the behavior coding system used in the present study, led to highly reliable observation results. However, this should not be carried to the point that observation is addressed to behavioral minutiae which have only a tenuous relation to important job skills. Some of the data gathered in the present study were not useful for that reason, for example, who walked in front on the way to a patrol car.

In developing observation for police evaluation purposes, caution should be paid during selection and training to possible observer biases, especially when women or minorities are to be evaluated. The present study found some significant differences in perceptions by women as opposed to men observers and by civilians in contrast to police observers. Rotation of observers and sampling of several tours for a single officer seem indicated to maximize impartiality.

Likewise, it appeared that avoiding subjectivity in observation is generally desirable; subjective items were most susceptible to observer bias in rating policewomen. However, translating from objective behavior description to evaluation is problematic. Unless observers use ratings, their descriptions tend to be without an evaluative component. If women patrol officers are observed to "back up" other radio cars less than male officers, this may either be taken as a reflection of better or of worse police patrol. These problems are not insoluble; however, they are challenging. The author feels that careful specification by police experts and clients as to what will distinguish good and poor patrol performance should come either before or after the actual performance observations. For example, as in the New York State Trooper system, ability to make decisions might be an "element" considered by officers and supervisors to be essential to the job. These judges would then have to decide to what extent an officer making a decision to back up another radio car would be demonstrating decisive thinking--a positive performance factor. They could also decide that the same observed behavior was reflective of "over-aggressive policing"--a negative performance indicator. (In fact, with regard to this factor, present research found that as many as 98 percent of such "back-ups" appeared to be unnecessary.) It should be possible to develop a table of positive and negative points associated with specific observed behavior. Some aspects might be specific to particular jurisdictions, but could also have wide applications.

It is ironic that in an area where many questions have been raised about women on patrol, observation was the least economical. The Vera-NYPD study revealed that strenuous physical activity occurred in only 1 out of 14 patrol

encounters between police and citizens in N.Y.C., and that an attempt to assess these abilities for a single individual would be an unreasonably lengthy procedure using first-hand observation. The best alternative would seem to be deliberate testing in simulations of patrol incidents requiring physical skills, such as pursuing a subject and carrying an accident victim. Situational testing of related kinds has been used in various jurisdictions with police applicants and recruits.<sup>14</sup> Advanced forms of these tests could employ actors, other officers, stage sets, or real street or residential settings. The validity of the method could be assured by comparing the results of these exercises with observations of field performance by a small sample of the same officers. The method should probably also be applied to assess officers' physical controlling ability with violent citizens, instead of waiting for those rare events to be observed in the natural course of patrol. Simulation of both these sorts was planned as part of a second phase of Vera-NYPD research, but is presently not contemplated in New York City.

In summary, it is felt that actual first-hand patrol observation is the evaluation method of choice, to be supplemented by simulation of physically demanding patrol events since these occur so rarely in the natural course of patrol. Client ratings of police service are also recommended as supplemental information about officers.

While all extensions of available performance evaluation methods were originally developed out of the need for an unbiased evaluation of policewomen, the measures are equally applicable to policemen; it is hoped that they may enhance the ability of police administrators to evaluate the performance of all their patrol officers.

## NOTES

1. Adam Klimkowski. Women in Patrol: the Miami experience. Paper presented to I.A.C.P., Washington, D.C., 1976.
2. California Highway Patrol. Women traffic officer project: Final report. Sacramento, California, 1976.
3. Michael Rosenfeld and Richard Thornton. The development and validation of a multijurisdictional police test. International Personnel Management Association, Chicago, 1975.
4. Gary T. Marx. Alternative measures of police performance. Paper delivered to American Sociological Association, New York, 1973.
5. The New York City Police Department provides more extensive subratings than most jurisdictions, with the meaning of each subscale and rankings within it specified by examples.
6. Frank Landy. Police Performance Appraisal Project, June 1975.
7. George Tordy and Lorraine Eyde. Job analysis of the position of New York State Trooper: An application of the job element method. New York State Police, Albany, New York, 1976.
8. Cf. Albert J. Reiss, The police and the public, 1971; Block, Peter B. and Anderson, Deborah, Policewomen on Patrol: Final Report (Washington, D.C.: Police Foundation, 1974); Sherman, Lewis, "An evaluation of police-women on patrol in a suburban police department," Journal of Police Science and Administration, 1975, Volume 3, Number 4, pp. 434-438; Greenwald, Judith E. and Connolly, Harriet A., Policewomen on Patrol: New York City, unpublished manuscript, 1974.
9. Cf. Daniel Cruse and Jesse Rubin, Police Behavior (Part I), Journal of Psychiatry and Law, Summer 1973, 167-222.
10. Kansas City Preventive Patrol Experiment, Police Foundation, Kansas City, Missouri, 1975.
11. Cf. John Hafey. Issues in arrest productivity research. Unpublished manuscript, Vera Institute of Justice, 1975.
12. Both Cruse and Rubin (op. cit.) and William Brown ("Local Policing: A Three-Dimensional Task Analysis," Journal of Criminal Justice, 1975, 3, 1-16) have emphasized "controlling behavior" as a crucial dimension of police performance. In fact, Rubin found in his study of the quality of police-citizen interactions in Miami that "controlling" behaviors were more frequently exhibited by police than any other type of activity.
13. This was determined by calculating Cronbach's Alpha Coefficient which was required to be at least .60.

14. Robert B. Mills, Robert J. McDivitt, and Sandra Tonkin. Situational tests in metropolitan police recruit training. *J. Criminal Law, Criminology, and Police Science*, 1966, 57.

82920

PATROL BY HELICOPTER: AN EVALUATION

Dr. John F. Schnelle  
Middle Tennessee State University

Capt. Robert Kirchner  
Nashville Police Department

Experimental evaluations of police crime control strategies have recently been reported in several police departments including Kansas City (Kelling, Pate, Diekman, and Brown, 1975), San Diego (Boydston, 1976), and Rochester (Bloch and Bell, 1975).

These latter studies have reported mixed results concerning the efficiency of police patrol in reducing reported crime levels. The Nashville police department has also undertaken a program of systematic research over the last 3 years which, unlike research in the other departments mentioned, is based on funding completely internal to the police department.

This latter ongoing program of research has led to the development of a low-cost research format based on an initial retrospective evaluation of police department procedures. The retrospective evaluation is followed by a series of more controlled experimental evaluations of the same procedure if the retrospective evaluation produces data judged significantly important to warrant the cost of such future evaluations. The exact number and extent of such future experimental evaluations are determined by the significance of the data produced at each level of evaluation.

For example, an evaluation of a saturation car patrol directed toward home burglary prevention was conducted 3 months after the patrol was terminated (Schnelle, Kirchner, McNees, and Lawler, 1976). This retrospective evaluation indicated no change in reports of burglary even though patrol was allegedly increased by over 100 percent. These latter data were significant in the sense that a normally used police procedure with a significant cost factor led to no measurable benefit. Thus, a more intensive and experimental evaluation was completed which involved systematically changing levels of patrol in different target areas while extensively monitoring the actual performance of the increased patrol. This second experimental evaluation also failed to document the burglary deterrence efficiency of increased police patrol. Since the second evaluation was negative, and was also collaborated with similar results from Kansas City, it was decided to risk no further resources in evaluating police car saturation patrol. This paper describes the application of this "in house" research format to a police helicopter patrol procedure. The application of this format illustrates the logical sequence of experimental evaluations that can occur based on an initial retrospective evaluation. More importantly, the outcome measures which define the "significance" of the helicopter patrol procedure and which thus determine future levels of evaluation activity are explicated.

Project Description: The special operations section of the Nashville police department utilized a helicopter patrol procedure in an attempt to reduce home burglaries in one area of the city.

The experimental zone incorporated 5.66 square miles and a population of approximately 12,000. The target zone was selected by the patrol chief because of chronically high burglary levels. The time that most residential burglaries were reported to have occurred was between 8 a.m. and 4 p.m. in the experimental zone, as well as in all zones in the entire city.

The helicopter was scheduled to fly between 9 a.m. and 5 p.m. over the experimental zone and was to stay in the air as much as fuel and weather permitted. The decision about refueling and weather conditions was left to the pilots. The physical boundaries of the zone were shown to the pilots who recorded their air time at the end of each shift. The pilots could make radio contacts with ground patrol cars at any time and were told to fly low enough so that they could detect suspicious ground activity.

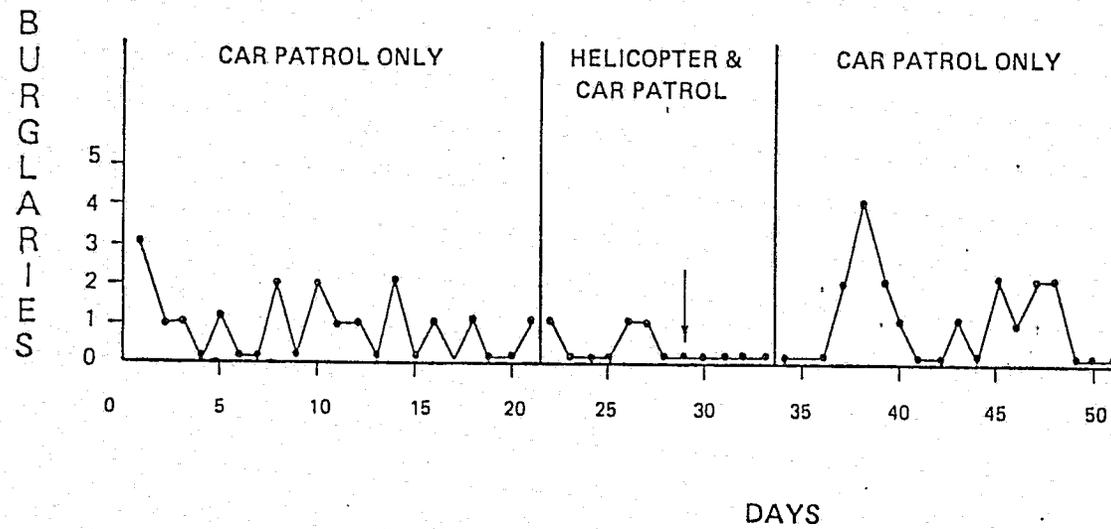
#### RETROSPECTIVE EVALUATION DESIGN AND MEASURES

The initial primary measure of the helicopter effects was report of burglaries. The frequency of burglary reports from all Nashville zones was monitored daily. Reported time of burglaries was taken daily from the reports of the complainants. In some cases a specific time range as to when the burglary occurred within one 24-hour period could not be identified. For example, if a homeowner was gone for several days and only discovered the burglary upon his return, then the estimated time range of the burglary occurrence would span several days. The latter types of burglary reports were analyzed separately from the reports in which a specific time span could be reported. Similarly, there were several instances in which the burglary could have occurred during a time span that crossed several shifts. In these cases, the burglary was assigned to the shift that accounted for most of the potential burglary hours. The addresses of all reported burglaries were crosschecked against a zone map to assure that burglaries would be assigned to the correct zones and the details of the burglary, i.e., method of entry, missing property, were crosschecked, when possible, by comparing the initial crime report prepared by a patrol officer against a supplementary report prepared independently by a detective.

The helicopter was initially flown on patrol for a 12-day period and then grounded because of cost limitations. This arrangement was conducive to an interrupted time series evaluation design with burglary measures taken before intervention, during intervention, and after intervention. This burglary occurrence data are illustrated in Slide One. As can be seen in table 1, there was a decrease in burglaries from baseline to the helicopter patrol period and an increase in home burglaries from the helicopter period to the postpatrol period.

These latter data were sufficiently dramatic and surprising to justify an immediate additional evaluation step designed to strengthen the initial evaluation design. Thus, the helicopter procedure was replicated for an additional 12-day period. This latter step permitted an analysis of the helicopter during five conditions: prior to initial patrol (baseline), during first patrol (first intervention), after the first patrol (second

TABLE 1.--Frequency of burglary during first experimental period



baseline), during the second helicopter patrol (second intervention), and after the second intervention (third baseline).

As can be seen in table 2, the burglary deterrence effects of the helicopter patrol were replicated in the second intervention, thus adding evidence that the helicopter patrol was the effective burglary deterrent.

The next step was to determine if the data were significant from a cost-benefit viewpoint. In other words, would cost-benefit data support investing more resources in helicopter patrol tactics? The costs most relevant to decisions concerning future patrol uses of the helicopter were the additional costs that were incurred when the helicopter patrol was experimentally validated. Costs which change because of a specific intervention are called marginal costs. These marginal costs are broken down on an hourly and total project basis in table 3.

Other cost factors that were not significantly affected by the patrol were potential costs that could have resulted from the helicopter being removed from other jobs because of the patrol activities (opportunity lost costs). In fact, the helicopter was not removed from its normal traffic and search activities; rather its overall use was increased by the patrol intervention. Before the intervention the helicopter spent more time on the ground. Furthermore, if an emergency arose in which a helicopter was called to assist in an emergency search, i.e., bank robbery, the response time of the helicopter would be reduced since it was already in the air. During the time period of this experiment, the helicopter was not called for such an emergency.

#### BENEFIT ESTIMATES

The benefits resulting from most crime prevention programs are generally measured as a reduction in costs which occur as a result of deviant behavior (Glaser, 1973). Since there is typically a margin of error in making cost savings estimates, savings should be estimated in several alternative ways. A reasonable upper and lower dollar savings boundary can be estimated based on alternative estimates of benefits and any cost savings estimate within these boundaries can be used to make decisions (Levin, 1975). If decisions about the cost significance of the data do not change within the upper or lower boundary levels, then clear conclusions can be drawn about the cost effectiveness of the particular set of data. Applied to the helicopter intervention, there seem to be two logical methods of calculating the cost savings that resulted from the lower burglary rates.

During the initial police investigation of the burglary, the victim is asked to describe the items taken and to estimate their value. A supplementary report filed by a detective at a later date allows the victim to add items to the list. Shoup and Mehay (1971) suggest that this latter cost figure should be corrected according to the amount of property that the police recover and return to the owner. Since a police record of recovered property is kept by the police property section, this correction factor is relatively easy to compute.

The basic problem with victim cost estimates is the possibility that the victim will intentionally or unintentionally inflate the actual dollar

TABLE 2.--Frequency of burglary during second experimental period

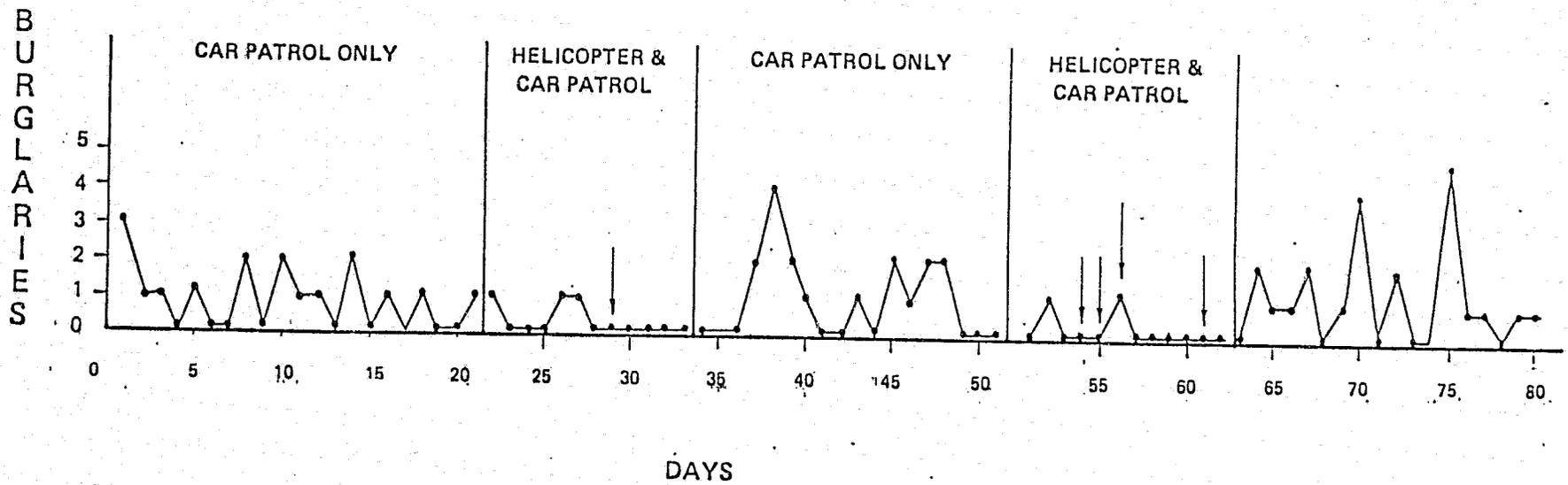


TABLE 3.--Marginal costs of helicopter procedure

	Hourly cost	Project cost
Fuel	\$24.00	\$2,445.60
Oil	.25	25.47
Motor overhaul costs	4.00	407.60
"Other" costs	1.50	<u>152.85</u>
Total project costs		<u>\$3,031.52</u>

loss of the burglary. The correction for recovered property does not control for this problem since the police recover property on a relatively infrequent basis. Nevertheless, even though the victim estimates might be distorted, it is still true that insurance companies base their payments directly on the victim and police reports. Thus, a second estimate of the cost of a burglary can be gained through insurance company records. These estimates represent an actual cash flow from the insurance company to the victim; hence, irrespective of possible inflations by burglary victims, the insurance company figures objectively reflect money lost due to burglaries.

The average dollar cost reported by the two largest companies with headquarters in the Nashville area which insure a total of 20,500 Nashville homes was \$486.50 (\$455.26 and \$517.74, respectively). These latter figures seem to be comparable to a national burglary loss figure reported by White, et al. (1975) of \$337 in 1973.

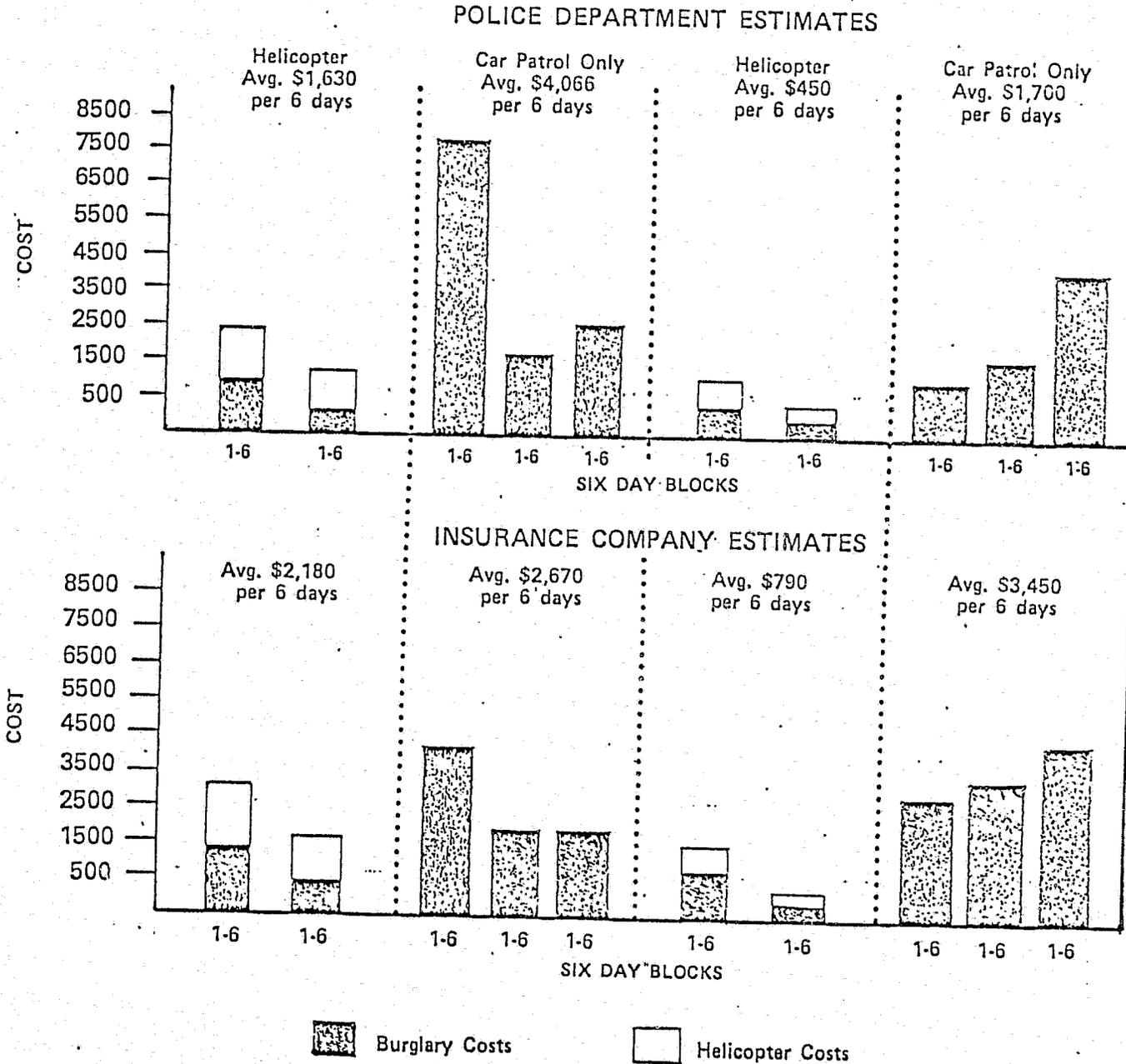
#### COST-BENEFIT COMPARISON

To estimate the cost-benefit comparisons between helicopter patrol and nonhelicopter patrol conditions, the total amount of property loss resulting from burglaries was added to the marginal costs of the helicopter patrol. When the helicopter was not used during the baseline conditions, there were no such marginal costs to add to the property loss estimates. The upper panel in table 4 illustrates the cost-benefit comparisons when the property loss estimates were based on police reports, while the lower panel in table 4 illustrates the cost-benefit comparisons when the property loss figure was based on average insurance company data. It is clear from both panels that the total costs incurred during the helicopter periods (helicopter plus burglary costs) are reliably less than the total costs incurred during the no-helicopter patrol periods (burglary costs only). Thus, the savings resulting from the helicopter intervention support administrative decisions to risk additional money to further investigate the usability of helicopter patrol procedures (table 3 indicates the total marginal cost to the police department for the current project as \$3,031.52). Although the cost-benefit data seem to be clear, there are three problem areas in this analysis.

First, long-term changes in both costs and benefits can drastically change the picture. The extreme reversability of the crime suppressant effect and the short intervention periods (forced short by cost factors) do not preclude the possibility that the helicopter might be a reactive intervention that would have a diminishing effect if continually employed.

A second area in which costs and benefits are difficult to assess concerns such potential benefits as the feelings of increased security that citizens may have when protected by helicopter patrol. This area also includes such costs as dissatisfaction with the noise or pollution produced by the helicopter. In cost-benefit jargon, the latter noncash factors would be labeled consumption benefits as opposed to investment benefits. Investment benefits are increased capitalized net economic worth attributable to a procedure (Neenan, 1974). Even though it is obvious that citizens in different areas might estimate these consumption benefits

TABLE 4.--Cost estimates of patrol types



differently (citizens in high crime, high ambient noise level areas would probably weigh security effects above noise effects), survey data collected by Simonson (1975) in Columbus, Ohio, indicate citizens strongly support helicopter patrol procedures to which they were exposed.

A final area concerns the potential distribution of benefits resulting from helicopter patrol as compared to the distribution of costs. The helicopter procedure would probably never be cost-effective and should not even be attempted in some areas of the city which experience low levels of home burglaries. The geographical areas of the city could be rated according to burglary cost levels and a picture could be developed of the different areas of the city that could potentially benefit from the helicopter procedure. If the number of such areas proves to be extremely low, given the optimum cost-benefit ratio of the helicopter intervention, then the significance of the helicopter as a general burglary reduction technique would be diminished. This latter point is true since it would be difficult to justify the marginal cost of the intervention which is borne by all taxpayers in Nashville against the benefit of the intervention, which might be received by only a few.

In sum, the crime and cost-benefit data from the initial evaluation justified the expenditure of resources on future patrol evaluations that could be directed toward answering two primary questions: (1) How many areas of Nashville could benefit from the patrol (distribution of benefits)? (2) How long-lasting are the patrol effects?

The first series of evaluations was directed toward answering the distribution of benefit question. The Metropolitan police area was divided into potential areas of helicopter coverage based on burglary reports and density of population. This analysis revealed the existence of three areas similar to the initial target zones. These three areas ranged in size from 9-15 square miles. There were also four other definable low-density areas which ranged between 50 and 140 square miles in size.

Two high-density and two low-density target areas were randomly chosen and a multiple baseline design was employed to evaluate the helicopter procedure.

Table 5 shows the effects of helicopter patrol on the two high-density target areas. Once again a decrease in burglary reports is visually obvious in both areas. Furthermore, there remains a positive cost-benefit ratio (table 6) in both areas.

Table 7 shows the noneffects of the helicopter in the two low population density areas. There is obviously no effect on crime and obviously, in the absence of benefits, an unfavorable cost-benefit ratio (table 8).

These latter data thus underline a direct limitation as to the distribution of benefits that could result from routine employment of the helicopter. This distribution of effect limitation in combination with the high marginal cost of the helicopter procedure reduces the overall significance of helicopter patrolling and brings into question the justification for future evaluations.

TABLE 5.--Frequency of burglary in high density areas

88

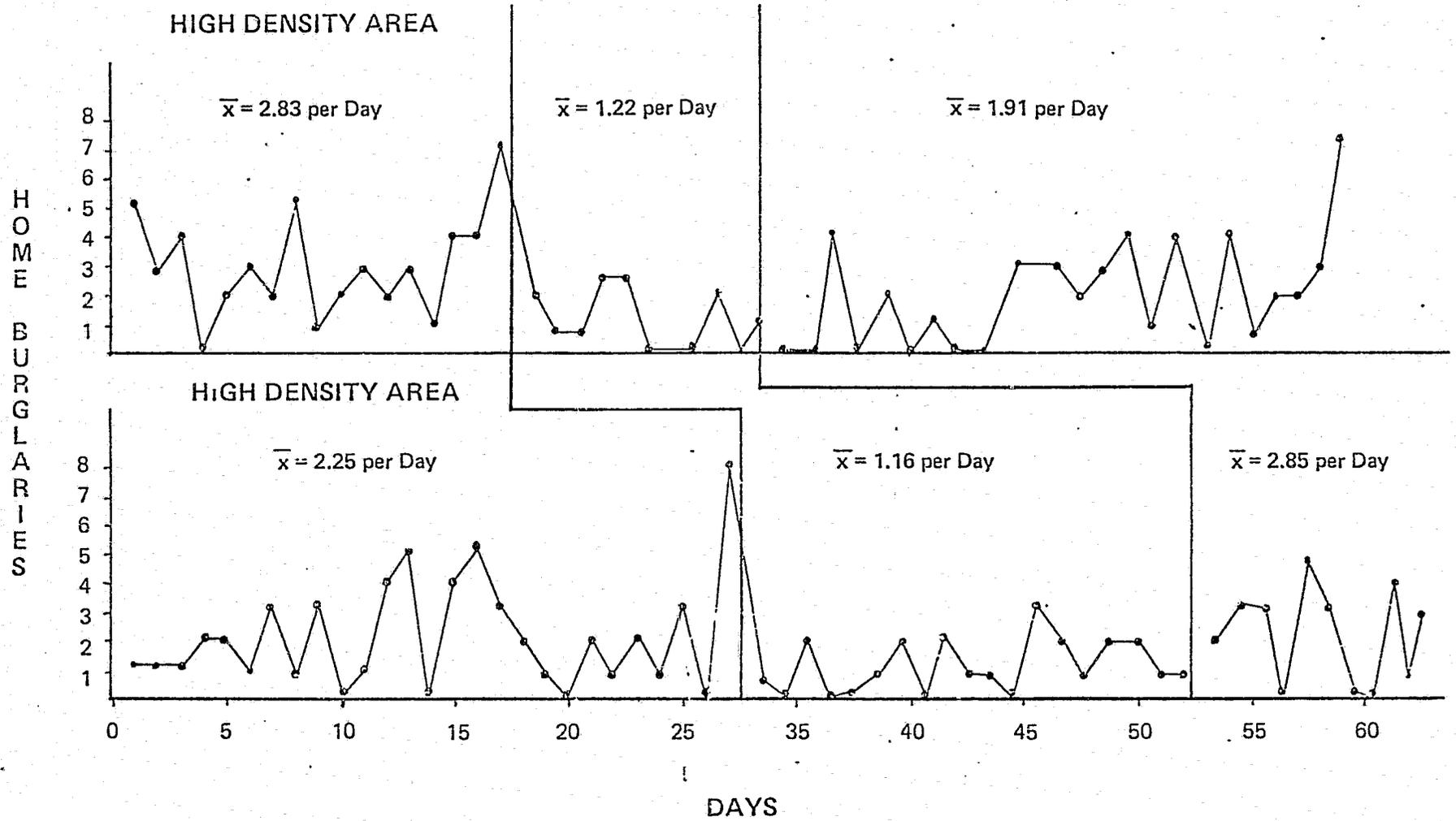


TABLE 6.--Average costs by patrol type in high density areas

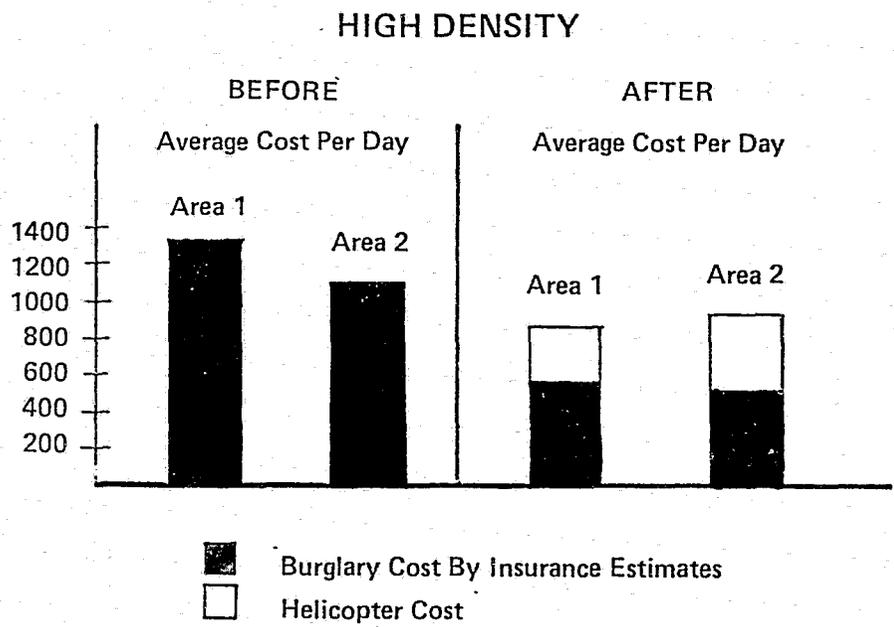


TABLE 7.--Frequency of burglary in low density areas

70

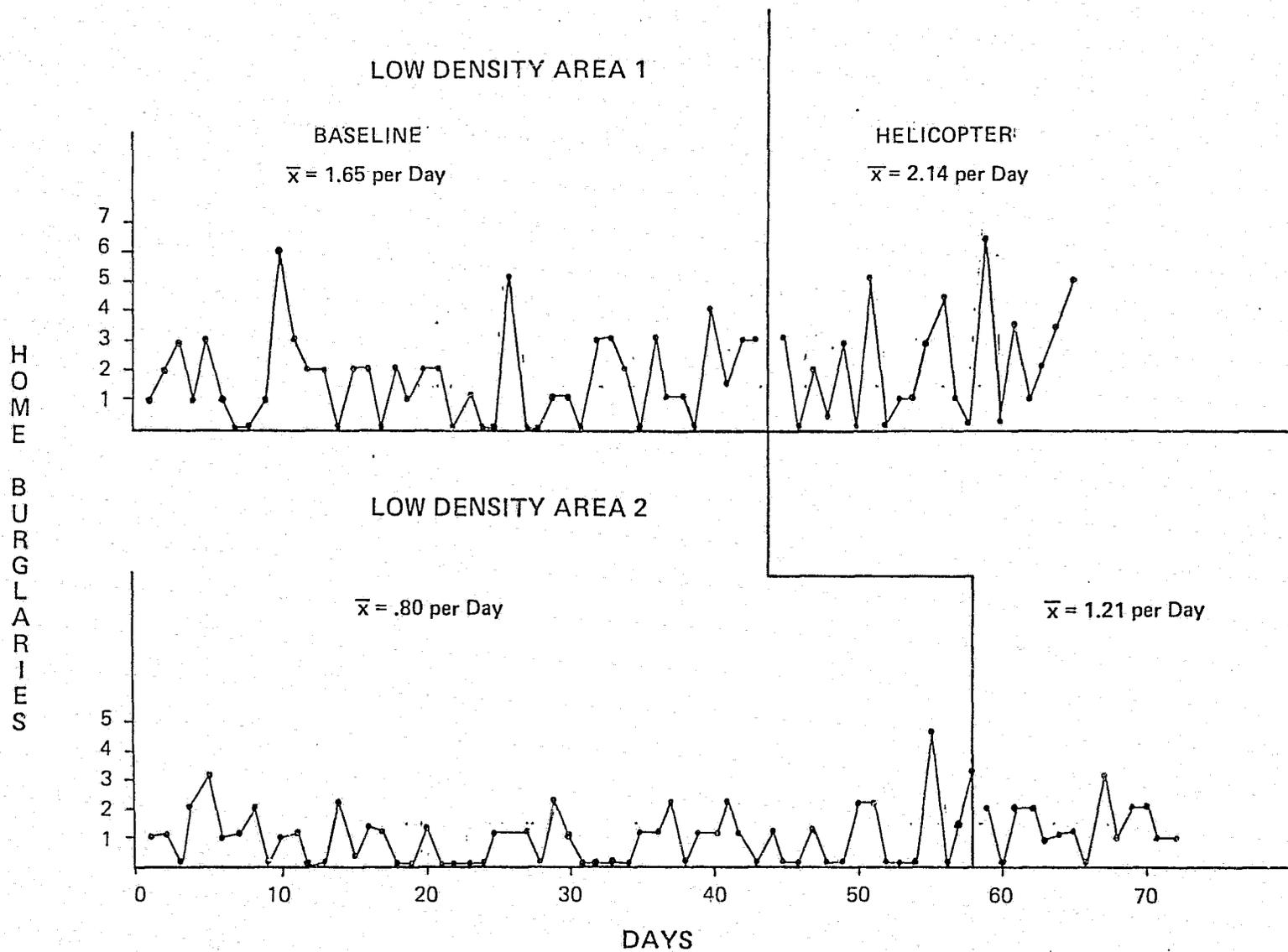
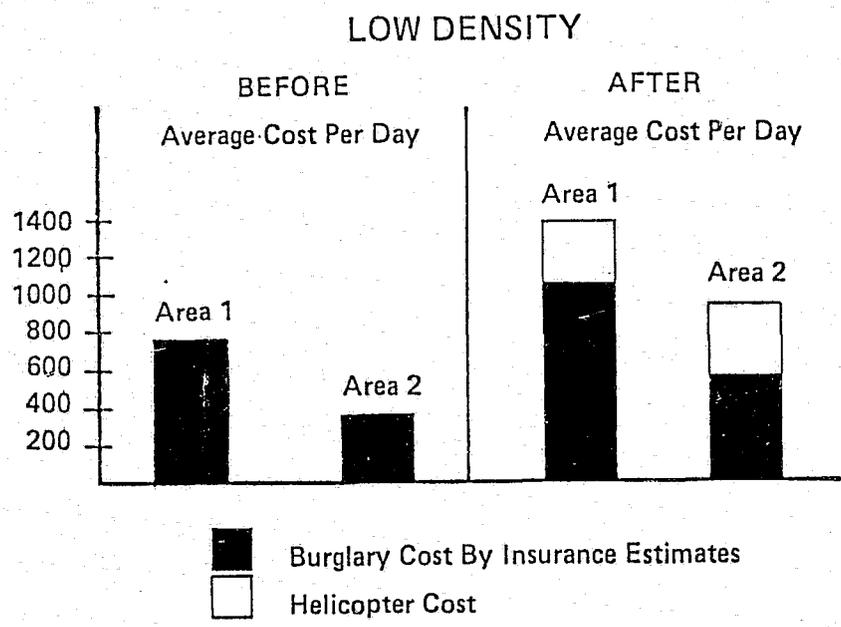


TABLE 8.--Average costs by patrol type in low density areas



If methods to improve the benefit distribution cannot be generated then administrative officials must decide one of three things: (1) find additional funds for a police procedure benefiting only a sample of a population; (2) generate a new funding formula based on the beneficiaries of the procedure paying for the procedure; (3) terminate the helicopter patrol as a routine police procedure.

In any case, the police department has done its job in clearly specifying the alternatives to policymaking personnel.

## BIBLIOGRAPHY

- Bloch, P. B., and Bell, J. Managing investigations: The Rochester System. Police Foundation, 1909 K Street, N.W., Washington, D.C., 1975.
- Boydston, J. E. San Diego Field Interrogation: Final report. Police Foundation, 1909 K Street, N.W., Washington, D.C., 1975.
- Glaser, D. Routinizing evaluation: Getting feedback on effectiveness of crime and delinquency programs. Washington, D.C., U.S. Government Printing Office, 1973.
- Kelling, G. L., Pate, T., Dieckman, D., and Brown, C. E. The Kansas City preventive patrol experiment: A summary report. Police Foundation, 1909 K Street, N.W., Washington, D.C., 1975.
- Levin, H. R. Cost effectiveness analysis in evaluation research. In Guttentage, M. and Struening, E. L. (Ed.), Handbook of Evaluation Research. Beverly Hills-London: Sage, 1975, pp. 89-121.
- Neenan, W. B. Benefit-cost analysis and the evaluation of mental retardation programs. In Davidson, P. O., Clark, F. W., and Hamerlynck, L. A. (Eds.). Evaluation of behavioral programs. Champaign, Illinois: Research Press, 1974, pp. 175-189.
- Schnelle, J. F., Kirchner, R. E., McNees, M. P., and Lawler, J. M. Social evaluation research: The evaluation of two police patrolling strategies. Journal of Applied Behavior Analysis, 1975, 8, 353-365.
- Shoup, D. C., and Mehay, S. L. Program budgeting for urban police resources. Los Angeles, California: University of California Institute of Government and Public Affairs.
- Simonson, C. E. Helicopter patrol. Police Chief, October 1975, pp. 30-33.

SECTION II  
EVALUATION IN THE COURTS

Preceding page blank

MEASURING THE EFFECTIVENESS OF THE BAIL BOND SYSTEM  
AS AN ASSURANCE OF TRIAL APPEARANCE

Helen Reynolds\*  
Department of Economics  
Southern Methodist University

It is widely held that there are two basic reasons for imposing financial bail upon accused offenders as a condition for release from jail before trial:

- (1) to assume court appearance by confiscating the bond in the event of nonappearance and
- (2) to have a method of retaining certain dangerous defendants by setting a bond price higher than those defendants can pay.

It is the first of these two functions of bail that we shall discuss. For an economist, the imposition of a financial bond to assure court appearance is the more interesting function because it involves decisions by the court about releasing risky defendants on bail and decisions by the defendants on bail about whether or not to show up for trial.

I. THE PROGRAM TO BE EVALUATED

The specific bail system evaluated here is the system in Dallas County for handling persons accused of committing both felonies and misdemeanors. The county jail is controlled and staffed by the Dallas County Sheriff's Department and therefore, the pretrial disposition of prisoners is handled through the Sheriff's Office. The bond information and the personal data used here were collected from the files on 600 persons released on bond in 1973. The limitations on the type of information available were imposed by the specific questions asked by the staff of the Sheriff's Office. The information in the sheriff's files gave the following information:

race  
age  
occupation and employment status  
marital status  
children  
place of residence  
place of birth

---

\* I would like to thank Rob Roby for data collection and sorting and Gerald W. Scully for initial conversations about bail.

offense charged with  
offense(s) previously charged with  
number of previous arrests  
date first arrested  
bond price  
appeared in court or not

From this information we shall identify some general characteristics of those who have posted bond but do not appear for trial.

## II. OBJECTIVES OF THE EVALUATION

The basic objectives of this study are three:

- (1) to select variables which significantly affect the probability that an accused will forfeit bail (not appear in court);
- (2) to determine the relation between bond price and the significant variables found in (1);
- (3) to compare the significant variables found in (1) to the criteria used for denying release in the Dallas County Pretrial Release Program.

It should be noted here that the decisions about granting bail and about granting pretrial release are made by different county departments, and those departments do not necessarily coordinate their activities. Part of the ensuing discussion will deal with the different methods used to estimate the probability that an accused offender will fail to appear in court and what effect those differences have on who is released by different agencies.

## III. HYPOTHESES USED

In economics, the analysis of decisionmaking is generally undertaken by considering the amount of utility (satisfaction) that an individual will gain from each possible choice. The choice that is selected is shown to contain more utility than the other choices by the mere fact that it was chosen. This choice is made by a decisionmaker weighing the costs and benefits of each choice and selecting the option that will yield maximum utility. So it is with the accused offender when, once released on bail, he decides whether to appear for trial or forfeit the bail bond. The utility derived from fleeing prosecution would be constrained by the risk of detection and apprehension for forfeiting bond as well as the possible outcome of the trial for the original charge. The utility from facing trial, on the other hand, is greatly determined by the peace of mind of not being a fugitive. There will be some risk involved in forfeiting bond, so the decisionmaker must choose while weighing the risks and possible outcomes, tempered by his attitudes about risk.

Tastes and attitudes about risk are characteristics of the utility function of an individual, but many people can have similar utility functions. If there develops a pattern of a type of person whose tastes and attitudes lead him to forfeit bail and flee, then we might be able to predict bail forfeiture from such a pattern. The personal characteristics of groups of

people often allow us to predict their behavior; for instance, housing purchase patterns can be predicted by finding age, income, schooling, and professions of the buyers. Here, similar personal data may be helpful in predicting court appearance or bail forfeiture. Information on age, education, marital status, and other socioeconomic variables categorizes groups of individuals into economic behavior groups.

Also, something about the activity in which a subject is engaged will boost the predictability of his behavior. For example, occupational groupings are helpful in organizing some data. Often people involved in the same occupation are in that occupation because of similar tastes or they have had their tastes molded by the occupation. Similarly, the offense with which one has been accused may tell us something about his behavior or at least may allow us to group him with others with similar tastes for criminal activity.

We shall use personal characteristics and criminal data for each subject, along with whether each subject appeared for trial or forfeited bail. From statistical analysis of these data, we shall be able to speak in general terms about the relationship of each of the variables to the likelihood of forfeiting bail or appearing in court.

Additionally, we shall examine the role played by the size of the bail bond price in influencing the accused offender to appear in court. Simple demand theory tells us that when the cost of forfeiture goes up (that is, when higher bond prices are set) then fewer people will forfeit. We shall test the data for the effect of bond price on court appearance.

In order to choose the variables that are significantly related to trial appearance, we shall use some fairly simple statistical techniques. The statistically significant variables will be selected from the data bank collected from 1973. Using previous cases of this sort to predict current or future behavior is an acceptable statistical technique as long as the socioeconomic structure of the community does not change appreciably, which we shall assume is the case.

#### IV. EVALUATION MEASURES

A. Court Appearance of Those Granted Bail. The data from the Sheriff's Office were evaluated by multiple regression analysis, which used the decision to appear for trial as the dependent variable and used the information about the accused as the independent variable. With ordinary least squares as the regression technique, the general form of the equation of regression is

$$A = C + \alpha V_1 + \beta V_2 + \gamma V_3 + \dots + u$$

A = decision to appear in court (A=1 for appearance, A=0 for failure to appear, and therefore forfeiture of bond)

C = constant term

$V_1, V_2, V_3, \dots$  = variables pertaining to accused, such as age, marital status, previous record, et alia

$\alpha, \beta, \gamma, \dots$  = coefficients of regression

u = disturbance term

For the sake of simplicity, the independent variables were regressed against court appearance one at a time. Otherwise, we would have tried simultaneously to regress 45 independent variables which would have resulted in vast computational problems, particularly multicollinearity. Therefore, in order to screen variables for statistical significance, each independent variable was regressed on court appearance with the coefficient and t-statistic noted. Table 1 shows the results of those regressions. Note in those results that the dependent variable was "court appearance"; therefore a variable with a positive coefficient is directly correlated with court appearance, and a variable with a negative coefficient, inversely related.

The regressions listed in table 1 enabled us to sort out the significant variables, ones with t-statistics that were statistically significant at the 95 percent confidence level. According to these preliminary results, the significant variables in predicting appearance at trial are:

- Negro
- Residence out of state
- Residence out of county
- Charged with sex crime
- Charged with weapons offense
- Previously charged with burglary
- Previously charged with robbery
- Previously charged with narcotics
- Number of previous arrests

After this initial selection process, these variables were regressed all together against court appearance.<sup>1</sup> (Due to severe problems of multicollinearity, the variables for previous charges were dropped; the other variables remained as is. The influence of "previous charges" will also be included in the "number of previous arrests.") Table 2 lists the results of the regression involving the previously selected significant variables.

In reviewing the regression results, the reader is directed to note the magnitude of the t-values (as well as the signs of the coefficients). For a 95 percent confidence level the critical value is  $t = |1.645|$  for a sample as large as ours here. By finding t-values greater than  $t = |1.645|$ , we can select the statistically most significant variables. Table 2 identifies the most notable characteristics of an accused offender as:

- Residence out of county
- Charged with theft
- Charged with weapons offense
- Charged with narcotics offense
- Number of previous arrests

It is interesting to note here the inclusion of "sex crimes" in the list of significant variables in table 2 and the strong t-value (within a 90 percent confidence interval). The positive sign assigned to the coefficient means that court appearance and the charge of sex crime are positively related; that is

TABLE 1.--Results of regressions of each variable  
against court appearance

Variable	Coefficient	t-Statistic
Negro	-.105	-2.01*
Spanish surname	-.043	-.381
Age	-.001	-.157
Residence out of state	-.381	-3.01*
Residence out of county	-.304	-4.13*
Nonskilled worker	-.067	-1.26
Place of birth not Texas	-.051	-1.07
Charged with:		
Burglary	-.007	-.110
Robbery	-.037	-.387
Assault major	-.099	-.463
Theft	-.111	-1.79
Sex crime	.530	2.71*
Rape	-.076	-.443
Murder	.116	.656
Manslaughter	.535	1.05
Auto theft	.175	1.52
Arson	.041	.182
DWI felony	.022	.332
Narcotics	-.078	-1.25
Weapons	-.226	-2.00*
Other	-.242	-1.74
Single	-.049	-.839
Divorced	-.032	-.436
Have children	.036	.656
Unemployed	-.066	-1.25
Previous charges:		
Burglary	-.056	-2.98*
Robbery	-.726	-2.98*
Assault minor	-.199	-1.53
Assault major	-.005	-.161
Theft	-.039	-2.31
Sex crime	.136	1.98
Rape	-.066	.575
Murder	-.035	.340
Manslaughter	-.496	-1.36
Auto theft	-.068	-1.85
DWI felony	.013	1.26
Narcotics	.063	-2.51*
Weapons	-.045	-1.17
Other	-.091	-4.83
Number of previous arrests	.020	3.74*
Bond price	.486	.486

\*Coefficients that are significant at the 95 percent level.

TABLE 2.--Regression results of previous selected variables regressed on court appearance

Variable	Coefficient	t-Statistic
Constant	.8319	1.050
Negro	-.3030	- .8052
Residence out of state	-.1367	-1.126
Residence out of county	-.3457	-4.867*
Charged with theft	-.1428	-3.005*
Charged with sex crime	.2271	1.589
Charged with weapons offense	-.1323	-1.667*
Charged with other offense	-.3316	-2.665*
Number of previous arrests	-.1780	-4.123*

\*Statistically significant at a 95 percent confidence level

R<sup>2</sup> = .1194.

to say that those charged with violating a sex crime tend to show up for trial. The term "sex crime" here is not to be interpreted as "rape"--a separate category altogether--but rather crimes against public decency such as prostitution, incest, fondling, homosexuality, etc. One might conjecture that those charged with sexual conduct outside the norm feel a strong enough tie to the community that they tend to appear for trial.

B. Bond Price. The price of the bond as well as the decision to grant bail is determined by a county judicial officer. The size of the bond is left up to the discretion of the judge, after consultations with the prosecutor or arresting officer or defense attorney. How are these bond prices determined? Does the size of the bond reflect the riskiness of freeing the defendant? If so, then defendants with a higher probability of fleeing would be assessed larger bonds. If not, then there are some other criteria for choosing the size of a bond imposed on a defendant.

To investigate the relationship between bond price and other variables, we first looked at the regression of bond price alone on court appearance to see if a larger bond size increased the likelihood of court appearance. The regression equation was:

$$\text{Court Appearance} = .6011 + .2708 \text{ Bond Price} + u.$$

(t = 23.21) (t = 3.736)

With an R<sup>2</sup> = .0228, one could say that bond price is significant (at the 95 percent confidence level) but only explains about 2 percent of the variance in court appearance. In this equation the constant term (representing all other factors) was much more significant, so we say that higher bond prices will not assure a proportional increase in court appearance.

Next we looked at the relationship between bond price and those variables found to be significant in determining bail bond forfeiture. The results of that regression are presented in table 3. According to those results, the variables that significantly affect court appearance have little effect on bond price. Of the variables entered into the regression, only "Residence out of county" and "Charged with weapons offense" are significant factors influencing bond price.

One may begin to ascertain that the county officials have little communication with each other concerning the bail system and forfeiture. At this juncture several problems in the use of data seem clear. First, the Sheriff's Department, which collects the data about defendants, may not be using the data to its fullest advantage. Examination of the data even in a simple formal manner would have revealed the relationships found in tables 1 and 2. Because of a lack of general knowledge about the determinants of bail forfeiture within the county offices, one can discern that the information about these determinants has not been gleaned from the data.

Second, the data collected by the Sheriff's Department may not be asking the correct questions to get a good picture of the determinants of bail forfeiture. The regression results presented in table 2 show that the most significant variables when used together only explain 11.94 percent of the variance in court appearance (the  $R^2 = .1194$ ). This indicates that there are some excluded variables that could help explain court appearance. Further data about wages or job stability and about home ownership or residential transiency might enhance the information about ties within the community. Even a department store inquiring about credit references for a charge account would delve into a person's stability and record of trustworthiness with more thoroughness than shown by the data used here.

Third, the bond prices set for accused offenders do not always correspond to their probabilities of appearing in court. We can assume that the bond price is imposed as a financial assurance that the accused will show up for trial, particularly in the smaller magnitudes of bond prices, such as in the data sample. (Larger bond prices tend to act as a financial barrier to being freed from jail.) Using conventional wisdom that the more serious crimes and "hardened criminals" should be dealt with more harshly at all levels, we took the variables which include serious crimes (FBI Index crimes, sex crimes, and narcotics) and "hardened criminals" (number of previous arrests) to be regressed against bond price. Also, to test for any racial biases, the variables for minority race (Negro and Spanish surname) were included. The results of this regression are found in table 3.

Note in table 3 that the variables which significantly affect bond price are the accused's being charged with burglary, robbery, assault-major, and murder. This means that a charge of one of these offenses influences a magnitude to increase the bond price. If the higher bond price is supposed to assure court appearance, the judge may in fact be imposing larger than necessary bonds on some defendants and smaller than sufficient bonds on others. Table 1 shows that of the serious offenses, only those charged with theft are significantly likely to forfeit bond, while the number of previous arrests shows up as a significant indication of a risky defendant. Table 3 shows that the number of previous arrests has little influence on bond price, however.

TABLE 3.--The results of regressing selected variables  
against bond price

Variable	Coefficient	t-Statistic
Constant	1594.0	.2303
Negro	- 73.94	-.3644
Spanish surname	-181.7	-.3312
Charged with burglary	912.8	3.186*
Charged with robbery	1562.0	4.615*
Charged with assault major	770.5	1.849*
Charged with theft	696.6	2.513*
Charged with sex crime	7.744	.9991
Charged with rape	228.8	.4443
Charged with murder	4756.0	9.371*
Charged with narcotics	396.6	1.441
Number of previous arrests	1.639	.7038

\* Statistically significant at a 95 percent confidence interval

$R^2 = .1526.$

The use of bond price as a deterrent to forfeiting bond would only be effective if in fact higher bond prices were imposed on riskier defendants.

C. Pretrial Release. The Dallas County Pretrial Release Program was instigated to release low-risk defendants without imposing a bond on them. The fee for pretrial release is nominal (\$15.00) and tends to remove the wealth bias from the bail system. The selection of low-risk defendants is made by a screening process that involves the judgment of the pretrial caseworker and statutory limitations imposed by the County Commissioner's Court. The caseworker uses his own evaluation of the reliability and strength of the community ties of the defendant, but he may not release a defendant if the following criteria apply:

- murder
- rape
- robbery
- assault with deadly weapon
- assault of a police officer
- aggravated assault
- felony drugs
- federal cases
- child abuse
- cruelty to animals
- city traffic violation
- sex related
- child molesting, fondling
- holds for other jurisdictions

sale of drugs  
forged prescription  
probation violators (felony probation)  
burglary of a habitation  
burglary of a building  
aggravated promotion of prostitution  
bond forfeitures  
prostitution  
no ties in community  
extensive prior records  
unstable residential record  
recidivistic tendencies  
retaliation  
arson  
AWOL  
impersonating a police officer  
false report to a police officer  
female impersonator  
carrying a prohibited weapon  
tampering with a witness  
bribery of an official  
permitting or facilitating an escape

By screening defendants through the above exclusions, the Pretrial Release Program collects many of the data that the Sheriff's Department does not. These data reflect an improved method of evaluating riskiness using information on such factors as residential stability and prior evidence of escape or unreliability (AWOL, bond forfeiture).

Since the decision about granting pretrial release precedes bail in the course a defendant must travel through the adjudication process, it is reasonable that pretrial release screens out the risky defendants and releases only those who have low risks. Then the bail system takes higher risk individuals and screens through them, sets bond prices, and releases those for whom the bond is posted. The last step in the bail system for a defendant, after being denied pretrial release, is to present his case to a bail bondsman. All three agencies involved in release, the pretrial release office, the magistrate granting bail, and the bail bondsman, have criteria for evaluating riskiness and predicting probabilities of bail forfeiture. Of these three agencies the most formalized system of evaluation is carried out in the Pretrial Release Program, whose standardized criteria seem to be more realistic than the ad hoc methods of the other agencies, with a few exceptions, to be noted later.

In looking at the exemptions from pretrial release, one can find some of the same general categories as found in the data from the Sheriff's Department. Many of the categories on the exclusions list are not included in the data from the Sheriff's Department. However, of the data from the sheriff, the variables that contain the pretrial release exclusions are:

Residence out of state  
Residence out of county  
Place of birth not Texas  
Charged with burglary  
Charged with robbery

Charged with assault major  
Charged with sex crime  
Charged with rape  
Charged with murder  
Charged with arson  
Charged with narcotics offense  
Charged with weapons offense  
Number of previous arrests

Together these variables were regressed against court appearance to see how significantly effective each of these was in explaining court appearance. If each of these variables is significant (has a t-value with an absolute value greater than 1.645) and has a negative coefficient, then there is a good argument for using it as an exclusion from pretrial release. The results of that regression are found in table 4.

The statistically significant variables in table 4 which support their use as exclusions for pretrial release are:

Residence out of county  
Place of birth not Texas  
Charged with weapons offense  
Number of previous arrests

This does not seem surprising if a defendant is considered a bad risk if he is transient ("no community ties"), has an extensive past record, and deals in prohibited weapons. The inclusion of these as significant indicators of riskiness would be anticipated even by conventional wisdom. (Also note that the variable "Charged with narcotics offense" is significant at the 90 percent confidence level, which lends support to its inclusion on the list of pretrial release exclusions.)

The interesting significant variables are those in table 4 that give contrary evidence to the exclusions for pretrial release. The variables

Charged with robbery  
Charged with sex crime  
Charged with murder

are statistically significant, but each has a positive coefficient. This means not only that those defendants tend to appear in court, but that they appear a significant percentage of times. One reason for putting these charges on the list of exclusions may lie in limiting the power of the Pretrial Release Program. If it is thought that pretrial release for felonies should be handled by magistrates, then the county government may not want to usurp that power and entrust it to a nonjudicial authority. Felonies are offenses considered by society to be more severe than others, and so releasing an accused felon before trial is considered a larger risk to society than releasing some charged with a misdemeanor, particularly in the event that another felony offense is committed during the release.

TABLE 4.--Results of regressing pretrial release variables on court appearance

Variable	Coefficient	t-Statistic
Constant	.8042	.7974
Residence out of state	-.1262	-1.053
Residence out of county	-.3406	-4.845*
Place of birth not Texas	-.1110	-2.931*
Charged with burglary	.7207	1.406
Charged with robbery	.1576	2.437*
Charged with assault major	.3834	.4961
Charged with sex crime	.2455	1.711*
Charged with rape	.1384	1.465
Charged with murder	.2135	2.309*
Charged with arson	.9358	.5495
Charged with narcotics offense	-.7924	-1.621
Charged with weapons offense	-.2079	-2.444*
Number of previous arrests	-.1839	-4.258*

\* Statistically significant at a 95 percent confidence level

$R^2 = .1426.$

#### V. BAIL VERSUS PRETRIAL RELEASE

The Pretrial Release Program was established in Dallas County in 1971 mainly as an economy measure. Its proponents argued that the program would save Dallas County tax revenues by diverting prisoners from the county jail and by keeping those with jobs from the welfare rolls. But more importantly the program has been annually releasing about 5,500 defendants, half of those who seek pretrial release.<sup>2</sup> Those defendants, who have an excellent record for appearing in court, may not have been able to pay the down payment on even a small bond. (The President's Commission Report, *The Courts*, cited the 1958 case in New York where 25 percent of all cases could not meet a modest bond of \$500<sup>3</sup>--equivalent to about \$1,000 today.) Certainly more misdemeanor defendants, particularly poor ones, have been released than would have been otherwise.

The bail system, on the other hand, falls heir to those defendants who have been charged with a felony or who otherwise have been rejected by the Pretrial Release Program. Bail serves as a means for a defendant to be released to prepare for his defense case for trial, but not all defendants can afford or qualify for bail. Even those who manage to gather the fee for the bail bondsman often have trouble paying the whole fee and often must further engage in illegal activities in order to pay the fee. As we saw above, the size of the bond does not necessarily insure court appearance. Larger bond prices only tend to insure that those who can post that bond are richer than those who cannot. If the courts cannot rely on the size of the bond to insure court appearance, then is bail a totally worthless institution? Certainly not

all defendants can be left in jail until trial. How can the courts insure trial appearance without wholesale retention of defendants?

Perhaps the bail system for felons and risky misdemeanants can learn something from the Pretrial Release Program. The bail system suffers from two problems that have been substantially solved by the Pretrial Release Program, namely:

- (1) the payment of a burdensome financial bond prevents some poorer defendants from being released, and
- (2) the selection process for deciding who shall be released on bail allows some high risk defendants to be set free.

The Pretrial Release Program solved the first problem by setting the price of the bond at a nominal \$15.00. The bond can be revoked for nonpayment (or other infractions), but most defendants have the ability to meet the bond, unlike larger bail bonds for felonies. Secondly, since the size of the bond price is not a good determinant of court appearance, the criteria used to grant bail should be revised to be more realistic, as they are in the Pretrial Release Program. The record of the two systems indicates the relative efficacy of the programs: the rate of bail bond forfeitures averages about 15 percent a year, whereas the rate of "no-shows" from the Pretrial Release Program is less than 2 percent. A better, more formalized selection process plus a reduction of the financial burden of the bail bond would make the bail system not only more effective (in assuring court appearance) but also more available to all income levels.

## NOTES

1. Using ordinary least squares for regressing both continuous and discrete independent variables on a dummy (0 or 1) dependent variable leads to a statistical problem of heteroskedasticity (the variance of the disturbances is not independently distributed). The estimates obtained for the coefficients are not as reliable as they would be without heteroskedasticity. A better but more complex form would be regressions using logistics analysis, which would yield better coefficients. However, since we are interested more in relative degrees of statistical significance, we shall assume that the variance in the disturbances affects all the coefficients uniformly, so that the ordering of the t-values remains unchanged.
2. According to a conversation with the staff of the Dallas County Pretrial Release Program.
3. President's Commission on Law Enforcement and the Administration of Justice, Task Force Report: The Courts (Washington, D.C.: Government Printing Office, 1967), p. 37.

82921

AN ALTERNATIVE TO COURT: AN EVALUATION OF THE  
ORANGE COUNTY (FLA.) BAR ASSOCIATION'S CITIZEN  
DISPUTE SETTLEMENT PROJECT

Ross F. Conner  
Assistant Professor  
Program in Social Ecology  
University of California, Irvine

Ray Surette  
Research Assistant  
Department of Criminology  
Florida State University

In early 1974, the American Bar Association (ABA) began a new program in criminal justice reform. The program was developed by the ABA's Commission on Correctional Facilities and Services and was named BASICS, an acronym which stands for Bar Association Support to Improve Correctional Services. As the name implies, the BASICS Program was intended to improve correctional services using a new method: bar association involvement and work on local criminal justice problems. This novel method for effecting correctional reform was developed by the director of the ABA's Commission on Correctional Facilities and Services and the vice-president of the Edna McConnell Clark Foundation, the funder for the new program. These men believed that lawyers and the bar associations of which they were members were an untapped resource in the correctional reform area. As the first BASICS informational brochure stated:

"There are several reasons why bar associations are eminently suited to accomplishing both the short and long-term goals. The bar is a self-perpetuating body. Its own leaders and members are frequently community leaders as well. Such influence creates access to the many local financial and human resources vital to a reform effort. In undertaking correctional programs, each association becomes part of a nationwide network with established lines of communication and the technical assistance of the ABA Corrections Commission. With financial support, therefore, bar associations have the potential to plan and execute specific, well-defined improvement programs."

The BASICS' plan, then, was to activate local bar association members first to plan and later to implement some type of correctional reform effort. The project solicited applications from bar associations across the country for small planning grants of approximately \$3,000. They received 106 applications and funded 80 planning projects in 40 states (Huff, Conner, and Geis, 1975). After approximately 3 months of planning, 62 bar associations applied for larger grants up to \$35,000 to implement their correctional reform efforts. Twenty "action" grants were awarded by the BASICS Program, including an award to the Orange County, Florida, Bar Association to implement a Citizen Dispute Settlement Project. It is this particular "action" project that we will be discussing here.

## PROGRAM DESCRIPTION

The Citizen Dispute Settlement (CDS) Project was designed to provide impartial hearings for residents of Orange County, Florida, who have complaints involving ordinance violations, misdemeanors, and minor felonies (for example, simple assault). This concept is not new and can be traced to the early function of the Justice of the Peace who served as an informal hearing officer for citizens involved in minor disputes. Currently, police officers often serve in this role when they are called, for example, to quiet a neighbor's barking dog or to end a family dispute. Police officers, however, often do not have the time to arbitrate a thorough hearing between parties. Minor disputes which are left unresolved can develop into major disputes, adding additional cases to the overburdened criminal justice system.

The idea of establishing a dispute settlement project outside the criminal justice system was developed in the Night Prosecutor Program in Columbus, Ohio (U.S. Department of Justice, 1974). There, law students served as hearing officers who aided the complainant and respondent in reaching a mutually satisfactory settlement. The purpose of the hearings was not to determine right or wrong but instead to reconcile differences between the parties and end the dispute. The Night Prosecutor's Program reported a high degree of success in settling disputes and was selected by the Law Enforcement Assistance Administration as a project worthy of replication.

The members of the Orange County Bar Association who were involved in planning their bar's BASICS project were aware of the Ohio Night Prosecutors Program. The bar association members' discussions with criminal justice personnel convinced them that a citizen dispute settlement project would be beneficial to Orange County. In addition, the project planners realized that they had a ready and willing group of hearing officers: individual members of the local bar association. The action grant proposal submitted by the Orange County Bar at the end of their planning period met the requirements which the BASICS Program had adopted; that is, the plan was to improve the local criminal justice system using bar association resources. The BASICS Program staff were quite impressed by the Orange County Bar Association's project because it intended to activate bar members in a very direct manner by recruiting volunteer attorneys to serve at all hearings.

Orange County was awarded a grant, and in late 1975 the program began. Hearing officers were easily recruited and trained, and the project was publicized to attract clients either through direct contact with the program or through referrals from the police, the sheriff, or others active in the Orlando area criminal justice system. At the same time, the evaluation plan for the project was developed and measures were pilot tested.

## EVALUATION PLAN

Several objectives underlay the evaluation plan that was developed for the CDS Project. The primary objectives were (1) to monitor both the types of clients and the types of complaints, and (2) to measure the effectiveness of the hearings. There were other objectives, but they will not be discussed here.

To monitor the types of clients and the types of complaints, a client intake form was developed from a standardized form in general use in the Orange County criminal justice system. The client intake form was administered to all complainants by CDS personnel. Although we would have wanted to collect the same descriptive data on respondents as well as complainants, this was not possible.

To measure the effectiveness of the CDS hearings, a two-part plan was developed. The first part of the plan involved ratings made at the conclusion of all hearings. Complainants, respondents, and hearing officers made two ratings: (1) their degree of satisfaction with the settlement just reached and (2) their judgment of the likelihood that the problem which underlay the conflict had been solved. The parties made these ratings on seven-point scales, with three degrees of positive judgment to one side of a neutral point and three degrees of negative judgment to the other. A research assistant, who had not participated in the hearing, administered the scales; all judgments were made independently and confidentially by the three parties. To assure their validity and hence their reliability, the scales were extensively pilot tested; revisions were made both in the scales and in the instructions given to clients until we were confident that the scales' validity was high.

The second part of the research plan to assess the effectiveness of the hearings involved drawing a random sample of complainants for a follow-up approximately 2 weeks after the hearing. Complainants and respondents who had participated in a hearing were contacted, as well as complainants and respondents who were scheduled for a hearing but no hearing was held. Clients were contacted either by phone or in person by a research interviewer. Those who had participated in a hearing were asked to rate their satisfaction with the settlement at this later time and to make a forced-choice judgment of whether or not the problem which underlay the complaint was now solved. Clients who had not participated in a hearing were only asked to make the latter judgment. We were especially interested in comparing complainants who had no hearing with complainants who had a hearing to assess the longer term effects of the CDS hearing procedure.

#### DISCUSSION OF FINDINGS

The results presented here are based on data collected between January and October 1976. Although the CDS Project began in late 1975, the program was not fully operational and the evaluation measures were not completed until January 1976.

Type and Number of Complaints Filed. A total of 306 complaints were presented for settlement during this period. The largest categories of complaints were harassment (28.5 percent) and simple assault (19.7 percent). Other kinds of complaints were presented much less frequently (see table 1). If clients who presented multiple complaints are reclassified by their primary complaint, harassment and simple assault constitute a majority of the cases (31.4 percent and 22.2 percent, respectively).

Harassment was defined as verbal assaults on the complainant's character, charges of rumor spreading, or charges of purposely creating a disturbance for the complainant. Although these kinds of charges would not automatically involve the complainant and respondent in the regular criminal justice system,

TABLE 1.--Types of complaints filed: percentages

Category	Percentage
Harassment	28.5
Simple assault	19.7
Family dispute	6.2
Neighbor dispute	6.2
Petty theft	5.9
Property damage	5.9
Menacing threat	4.6
Breach of peace	3.9
Animal control	2.3
Trespassing	2.0
Bad checks	1.3
Littering	.7
Disorderly conduct	.3
Multiple complaints	<u>12.1</u>
	99.6*

\* Less than 100 percent due to rounding.

n = 305; one missing case.

these problems were often precursors of the second largest category of complaint, simple assault. Defined as actual physical assault to the complainant, simple assault usually does result in formal charges and entry into the criminal justice system. Although filed infrequently, types of complaints presented (for example, family dispute or property damage) can also lead to involvement in the criminal justice system. The CDS Program, then, has attracted a majority of cases involving serious problems--problems which could lead to adjudication. If these problems are being solved out of court, the CDS hearings would be helping to reduce the burden on the courts. We will analyze this issue of the effectiveness of the hearings below.

Complaints Brought to a Hearing. All complaints did not result in a hearing between the complainant and respondent. Of the 306 complaints presented, 194 (63.4 percent) resulted in hearings. In the other 112 cases, no hearing occurred for a variety of reasons, the most frequent of which was the absence of the respondent at the hearing (33.9 percent). Other important reasons were that the complainant was referred elsewhere (18.8 percent) for settlement of his problem (for example, to a community agency) or that the complainant agreed to drop his complaint (17 percent). (See table 2 for other reasons.)

Comparisons between the type of complaint and the occurrence of a hearing showed that all types of complaints did or did not result in hearings with about equal frequency. The one exception was neighbor dispute, which was more

TABLE 2.--Reasons hearings were not conducted  
(n = 112)

Reason	Percentage
No show--respondent	33.9
Complainant referred	18.8
Complaint dropped	17.0
Complainant and respondent agree to no hearing	14.3
No show--both complainant and respondent	12.5
No show--complainant	3.6
Warrant filed	0
	100.0

likely to result in a hearing ( $\chi^2 = 5.65$ ,  $df = 1$ ,  $p. < .02$ ; contingency coefficient = .48).

Selected Characteristics of Complainants. When clients came to the CDS office to present their complaint, CDS personnel completed intake forms on each person. These forms included questions on a number of client characteristics, several of which are presented here.

The median age of complainants was 36 years ( $\bar{X} = 38.2$ ,  $SD = 14.2$ ). More females than males filed complaints: 62 percent of the complainants were females and 38 percent were males. Information was collected on clients' race: 70 percent were white, 28 percent were black, and 2 percent were Spanish or other ethnic groups.

Data on marital status of complainants show that 49 percent of the clients were married, 20 percent single, 13 percent separated, 12 percent divorced, and 6 percent widowed. Data on employment status indicate that 48 percent of the complainants were employed full time, 6 percent part time, and 36 percent were unemployed.

The occupational status of the majority of complainants was sales (54 percent), defined here as clerical, service, and craftsmen. Other large occupational categories represented were labor (18 percent) and professional (16 percent). The monthly income of 36 percent of the complainants was under \$300; 45 percent earned between \$300 and \$800 per month.

The average complainant was an angry client, by his or her own admission. We asked complainants whether they would have pursued a warrant if the CDS Project was not available, and 80 percent reported that they would. We would expect this figure to be inflated somewhat by clients' need to convince the project of the seriousness of their problem. Nonetheless, the figure is quite high and indicates the seriousness of the complaints from the clients' viewpoint.

Source of Clients. Through general announcements and special presentations, the CDS Project was publicized throughout the Orange County area. In this way, the project hoped to attract clients from a variety of sources. The majority of clients who came to the project between January and October were referred either by the Orange County Sheriff's Department (33 percent) or the Orlando Police Department (28 percent). The State Attorney's Office referred 14 percent of the clients, and 10 percent came on their own. Private attorneys and police departments of communities surrounding Orlando were other less frequent sources of clients.

Effectiveness of the Hearings. One of the major objectives of the evaluation was to determine the effectiveness of the hearings. In several other CDS-type projects, effectiveness has been judged simply by noting the number of people who file complaints. Complainants whose cases do not result in hearings are treated as successes, as are all complainants who do have hearings. No attempt is made either to determine client satisfaction with the hearings or to determine whether problems have been solved after the hearings.

The CDS program director was not willing to accept client intake measures as success measures. Instead, he agreed to look closely at the success of cases that had hearings, as well as cases that did not result in hearings. This kind of evaluation plan is unique, and we salute the CDS project director for his willingness to undertake an objective evaluation of the outcomes of his program.

Following each hearing, the complainant, respondent, and hearing officer independently made two ratings: their satisfaction with the settlement they had just reached and their judgment of whether the problem which caused the complaint was now solved. These judgments were made on seven-point scales, with the most positive rating equal to 1 and the most negative rating equal to 7.

Overall, complainants, respondents, and hearing officers were generally satisfied with the settlements reached at the hearings. (See table 3.) Among complainants, 37.2 percent were "very satisfied" with the settlements; 68.6 percent gave positive ratings for the settlement. Only 8.4 percent of the complainants were "very unsatisfied;" 17.3 percent of the clients gave negative ratings. Among respondents, 48.4 percent were "very satisfied" with the settlements; 78.6 percent gave positive ratings. Only 6.8 percent were "very unsatisfied;" 9.4 percent gave negative ratings. In general, then, complainants and respondents gave quite favorable opinions. It is understandable that respondents would be even more satisfied than complainants: a respondent has just avoided a potential court battle with the complainant. Satisfaction ratings for complainant-respondent pairs were correlated (Spearman  $r = .44$ ;  $p. < .001$ ).

Hearing officer satisfaction ratings were also favorable, although slightly less so than complainants or respondents (cf. means in table 3). The correlations between hearing officer ratings and complainant ratings and between hearing officer ratings and respondent ratings were high ( $r = .40$ ;  $p. < .001$  and  $r = .41$ ;  $p. < .001$ , respectively).

The three parties involved in each hearing also made independent ratings of the likelihood that the problem which underlay the complaint had now been

TABLE 3.--Complainant, respondent, and hearing officer  
satisfaction ratings: percentages

Scale	Complainant (n=191)	Respondent (n=192)	Hearing officer (n=194)
1 - Very satisfied	37.2	48.4	27.3
2 - Somewhat satisfied	21.5	20.8	21.1
3 - Just a little satisfied	9.9	9.4	15.5
4 - Neutral	14.1	12.0	12.9
5 - Just a little unsatisfied	3.7	1.6	4.6
6 - Somewhat unsatisfied	5.2	1.0	4.6
7 - Very unsatisfied	8.4	6.8	13.9
	<u>100.0</u>	<u>100.0</u>	<u>100.0</u>
	$\bar{X} = 2.7$	$\bar{X} = 2.3$	$\bar{X} = 3.2$
	SD = 1.9	SD = 1.7	SD = 2.1
	Med = 2.1	Med = 1.6	Med = 2.6

solved (see table 4). Opinions were optimistic among complainants and respondents. Nearly 31 percent of the complainants and 41 percent of the respondents thought that it was "very likely" that the problem was now solved. Optimistic complainant ratings tended to be associated with optimistic respondent ratings ( $r = .33$ ;  $p < .001$ ). However, overall judgments were cautious: the average complainant rating was 3 (that is, "just a little likely") and the average respondent rating was 2.6, only slightly more positive.

Hearing officers were less positive than either complainants or respondents in their assessment of the likelihood of problem solution: their average rating was 3.8, only slightly optimistic. The correlations between hearing officer and complainant ratings and between hearing officer and respondent ratings were high ( $r = .49$ ;  $p < .001$  and  $r = .44$ ;  $p < .001$ , respectively).

In sum, these posthearing ratings by complainants, respondents, and hearing officers indicate that all parties were generally satisfied with the solutions reached and cautiously optimistic about the likelihood that the problem which underlay the complaint was solved. A random sample of hearing cases was selected to determine whether the long-term effects of the hearings matched the positive short-term effects. Approximately 2 weeks following their hearing, complainant and respondent pairs were contacted to obtain satisfaction ratings concerning the settlements that had been reached earlier. In addition, complainants and respondents were asked whether the problem that caused the complaint had been solved.

Comparative analyses of the follow-up sample and the entire hearing population verified the representativeness of the follow-up sample. Consequently, we are able to generalize from this sample to all hearing cases. As a group, complainants or respondents continue to report high satisfaction ratings

TABLE 4.--Complainant, respondent, and hearing officer ratings of the likelihood of problem solution: percentages

Scale	Complainant (n=191)	Respondent (n=192)	Hearing officer (n=195)
1 - Very likely	30.9	41.1	15.9
2 - Somewhat likely	20.4	20.8	19.0
3 - Just a little likely	6.3	8.3	11.3
4 - Neutral	23.6	15.6	16.9
5 - Just a little unlikely	4.2	2.1	9.2
6 - Somewhat unlikely	4.7	2.1	10.8
7 - Very unlikely	9.9	9.9	16.9
	<u>100.0</u>	<u>100.0</u>	<u>100.0</u>
	$\bar{X} = 3.0$	$\bar{X} = 2.6$	$\bar{X} = 3.8$
	SD = 2.0	SD = 1.9	SD = 2.1
	Med = 2.4	Med = 1.9	Med = 3.7

2 weeks after their hearing (see table 5). For complainants, the average rating at follow-up ( $\bar{X} = 2.6$ ) is quite similar to the average rating at the time of the hearing ( $\bar{X} = 2.7$ ). For respondents, the average rating at follow-up decreases from the average rating at the time of the hearing (2.8 and 2.3, respectively).

TABLE 5.--Follow-up sample--complainant and respondent satisfaction ratings: percentages

Scale	Complainant (n=41)	Respondent (n=42)
1 - Very satisfied	51.2	38.1
2 - Somewhat satisfied	22.0	26.2
3 - Just a little satisfied	2.4	4.8
4 - Neutral	4.9	11.9
5 - Just a little unsatisfied	0.0	0.0
6 - Somewhat unsatisfied	0.0	7.1
7 - Very unsatisfied	19.5	11.9
	<u>100.0</u>	<u>100.0</u>
	$\bar{X} = 2.6$	$\bar{X} = 2.8$
	SD = 2.3	SD = 2.1
	Med = 1.5	Med = 2.0

The average ratings for complainants and respondents tell only part of the story. It is also important to determine whether individual clients who were satisfied following a hearing remain satisfied 2 weeks later. Among the complainants, there was a moderate association between the two judgments ( $r = .30$ ;  $p. < .058$ ); among the respondents, there was a lower association between the two judgments ( $r = .14$ ).

When we analyzed these judgments more closely, we discovered that 75 percent of the complainants either remained satisfied with the settlement or changed their judgment to satisfied. In only seven of the 40 cases analyzed (17.5 percent) did complainants either remain unsatisfied or change their judgment to unsatisfied. Among the 41 respondents analyzed, 65.8 percent either remained satisfied with the settlement or changed their judgment to satisfied. In 29.3 percent of the cases, respondents either remained unsatisfied with the settlement or changed toward unsatisfied.<sup>2</sup>

The complainants and respondents who were followed up were also asked to judge whether the problem which underlay the complaint had been solved. These ratings were forced-choice judgments of either "yes" or "no." A majority of both complainants and respondents reported that the problem was solved. The CDS Program then is solving problems in a majority of cases, but a large minority of problems (35.7-41.5 percent) remain unsolved. (See table 6.)

Although a large number of problems are not solved, many clients were as satisfied or more satisfied with the settlements 2 weeks following the hearing. This finding reflects the fact that many of the problems which underlay CDS complaints are complex. A short half-hour to hour hearing may result in a satisfactory short-term settlement to a specific complaint, while the long-term problem remains unresolved. An actual CDS case will illustrate this point.

TABLE 6.--Follow-up sample--complainant and respondent forced-choice judgments, "Is the problem now solved?": percentages

Judgment	Complainant (n=41)	Respondent (n=42)
Yes	58.5	64.3
No	41.5	35.7

A feud between neighbors had reached the point of open hostility between the children. At the hearing, a settlement was reached such that the parents would closely monitor their own children to prevent fights. The hearing had given both parties a chance to explain their side of the story, and each party was quite satisfied with the settlement. Neither party, however, believed that much progress had been made in resolving basic differences that had existed for years. Consequently, both parties made high satisfaction ratings but judged the problem as not solved.

In addition to follow-up contact with CDS clients, we contacted a random sample of potential clients who had filed complaints but who had not participated in a hearing. The objective was to analyze comparable groups of complainants who did or did not have a hearing and thereby determine whether a hearing really helped to solve underlying problems. We obtained forced-choice ratings from 27 complainants who did not attend a hearing and from 41 complainants who did attend a hearing. For the no-hearing group, 51.9 percent reported that their problem was now solved. For the hearing group, 58.5 percent reported that their problem was solved. These data support the idea presented earlier that it is difficult for the CDS program to solve long-term problems.

Our current analyses of the CDS evaluation data focus on interrelations among factors such as type of complaint and judgments of problem solution. With these analyses we hope to provide suggestions and recommendations to the project director about ways of improving his program. Although many complainants and respondents are satisfied with the short-term settlements, some are not. We hope to specify some possible reasons for this so that the program can be made even more effective.

## NOTES

1. The research reported here was supported by a grant from the American Bar Association and the Edna McConnell Clark Foundation; their support does not, however, necessarily indicate their concurrence with our conclusions. We wish to thank David Linden of the ABA, Rod Petrey of the Clark Foundation, and Wil Miller of the Orange County Citizen Dispute Settlement Project for their support and help to us in our work.
2. These analyses involved reclassifying both ratings for easier inspection. Client ratings of 1 and 2 were reclassified as "satisfied," ratings of 3, 4, and 5 were reclassified "neutral," and ratings of 6 and 7 were reclassified "unsatisfied."

## BIBLIOGRAPHY

- Huff, C. R., Conner, R. F., and Geis, G. Planning Correctional Reform: An Assessment of the American Bar Association's BASICS Program. Washington, D.C.: American Bar Association. Research monograph, 1975.
- U.S. Department of Justice. Citizen Dispute Settlement: The Night Prosecutor Program of Columbus, Ohio. Washington, D.C.: Law Enforcement Assistance Administration, National Institute of Law Enforcement and Criminal Justice. Replication Manual, 1974.

## RECIDIVISM RATES OF DIVERTED JUVENILE OFFENDERS\*

Suzanne B. Lincoln, Kathie S. Teilman,  
Malcolm W. Klein, and Susan Labin  
University of Southern California  
Social Science Research Institute  
Los Angeles, California 90007

This paper reviews a project that was designed to evaluate a California law enforcement program to divert juveniles away from the juvenile justice system. The evaluated juvenile diversion program was run by the Juvenile Bureau of a large law enforcement department in California, using funds granted by the California Youth Authority, the state's office of juvenile corrections. The program emphasized diversion with referral to local, community-based counseling agencies, as opposed to simple diversion with no further action taken.

Probably the most distinctive component of the evaluated diversion program was the allocation of a portion of the grant monies for purchase of the services provided to referred offenders. The purchase of service contracts provided that for each diversion client served, the agency would receive a \$50 fee for providing a minimum of six sessions of service, and another \$150 if the client was not rearrested within a 6-month period following the referral arrest. The intention was to motivate the agencies to orient their efforts toward keeping diversion clients out of trouble with the law. Purchase of service contracts did in fact lead to the provision of increased service to referred clients. Records from agency files showed that 100 percent of the subjects referred with purchase of service were actually contacted personally at least once. By contrast, only about 80 percent of the subjects referred without purchase of service were contacted personally.

The primary objective of the evaluation was to determine which of four possible arrest dispositions had the most beneficial impact on subsequent offense behavior. The four dispositions were

- (1) outright release--that is, simple diversion without referral for additional services;
- (2) referral without purchase of service;

---

\* Work leading to this paper was supported in part by Grant No. MH-26147 from the National Institute of Mental Health (Center for Studies of Crime and Delinquency).

- (3) referral with purchase of service; and
- (4) normal petitioning through juvenile court, without detention.

There were two evaluation measures used to assess the impact of the four experimental conditions on the subsequent offense behavior of the project juveniles. These measures included (1) official rearrest data and (2) a self-reported index of involvement in delinquent activities. Rearrest data were retrieved using a central, countywide repository of juvenile arrest reports. The self-reported delinquency indices were obtained in the course of two waves of personal interviews with the juvenile offenders within the 18-month period following the referral arrest. The indices consisted of 18 behavior items for which a juvenile could be arrested. These included a broad range of status offenses, "victimless" offenses, and both minor and serious offenses against persons and property. Respondents were asked, in a modified card-sort technique, to indicate which, if any, of these things they had done subsequent to the arrest for which they entered the subject pool.

The project presented an opportunity to speak to two theoretical orientations much debated by social scientists and justice systems practitioners alike. These are

- (1) labeling or societal reaction theory, which emphasizes the role of official institutions in inadvertently encouraging illicit behavior, and
- (2) deterrence theory, an approach that examines the efficacy of official penalties in deterring antisocial behavior.

These two orientations appear in many ways antithetical to each other. Labeling theory would imply that the more involvement delinquents have with agents of social control, the more likely they are to assume deviant identities, and to get into more trouble as a result. Deterrence theory would imply that the actions of social control agents penalizing deviant behavior are likely to deter individuals from further wrongdoing. Actually, these two orientations agree on a fundamental point, that social sanctions may influence the subsequent behavior of sanctioned individuals. Proponents would tend to differ only on whether this influence tends to encourage or discourage further misbehavior. The commonality of these theories allowed us to address them both.

Due to the enlightened cooperation of the police administrators in charge of the program, it was possible to assign project offenders to the four treatment conditions on a random basis. The procedures used to accomplish this were simple. Except for being instructed to try to include moderately serious offenders into the referable pool, juvenile officers were allowed to select offenders for the referable pool using their normal discretionary criteria. After selection to the pool, offenders were assigned to disposition categories by the juvenile bureau commanders, who used lists of randomly generated dispositions. Because of delays in processing the purchase of service contracts through county bureaucracy, assignment of subjects to the refer-with-purchase condition began late. As a result, fewer subjects were assigned to this condition. At the end of the subject assignment period, 306 offenders had been randomly assigned to the four treatment conditions.

Halfway through the subject assignment period, evidence concerning possible selection bias was compiled from both police records and from informal conversations with station officers. This evidence suggested that, in fact, officers were assigning some subjects to disposition on a nonrandom basis. The station personnel were reprimanded and forcefully instructed to adhere to the randomization procedures. After subject assignment was completed, subjects in the four conditions were compared to determine whether officers had succeeded in entering significant bias into the assignment process. Subjects were compared on age, sex, ethnicity, nature of referral offense, and presence or absence of prior record. Although the differences across disposition are not significant with respect to most of these variables, there is a statistically significant difference among groups on prior record. Officers overassigned subjects with prior records to the most serious disposition, court petition. Fortunately, it was possible to adjust rearrest scores for this bias. The adjustment for prior record will be discussed below.

### FINDINGS

The earliest rearrest data were collected and tabulated for a period of 6 months from the date of each subject's referral arrest. These 6-month rearrest data were examined for differences across disposition with respect to both simple recidivism, that is, the number of recidivists within the 6-month time period regardless of number of rearrests, and with respect to multiple recidivism, which is the number of recidivists who were rearrested more than once. These data are presented in table 1. At 6 months after the referral arrest, 36 percent of the cohort of 306 subjects had been rearrested at least once. The simple rates for each disposition were: counsel and release, 28 percent; refer without purchase, 32 percent; refer with purchase, 35 percent; and court petition, 48 percent. Although there was a tendency for more subjects to recidivate in the more serious dispositions, the trend was not statistically significant at the .05 level.

TABLE 1.--Official recidivism frequencies (6 months)

	Release	Refer w/o purchase	Refer w. purchase	Court petition
Simple recidivism (1+ rearrests) $\chi^2=7.52, <.10$	23 (28%)	28 (32%)	19 (35%)	39 (48%)
Multiple recidivism (2+ rearrests) $\chi^2=15.35, <.005$	5 (6%)	14 (16%)	11 (20%)	24 (29%)
n=	81	88	55	82

When multiple recidivism was calculated for the four disposition conditions, ignoring prior record, results were obtained which picked up and magnified the pattern suggested in the simple recidivism data. The proportions of recidivists rearrested more than once across dispositions were: counsel and release, 6 percent; refer without purchase, 16 percent; refer with purchase, 20 percent; and court petition, 29 percent. These differences in proportions of rearrested recidivists are statistically significant at the .05 level. The impact of police disposition appears more pronounced for multiple recidivists than for simple recidivists, although in both cases the release condition shows the lowest rearrest rates, the court petition condition shows the highest rates, and the two referral conditions show rates intermediate between those two.

Later rearrest and self-reported delinquency data are now available for a period of about 15 months from the date of the referral arrest. At this later point, indicators of both simple and multiple recidivism show the earlier pattern that appeared most clearly in the figures for multiple recidivism. These data are summarized in table 2. There are now statistically significant differences in simple as well as multiple recidivism. The rates are: for counsel and release, 37 percent; refer without purchase, 45 percent; refer with purchase, 58 percent; and court petition, 62 percent. Again, multiple recidivism reflects the same pattern of increasing rearrests for the referral and court conditions. The proportions are: for counsel and release, 16 percent; refer without purchase, 27 percent; refer with purchase, 40 percent; and court petition, 41 percent.

TABLE 2.--Official recidivism frequencies (15 months)

	Release	Refer w/o purchase	Refer w. purchase	Court petition
Simple recidivism (1+ rearrests) $\chi^2=12.52, <.01$	30 (37%)	40 (45%)	32 (58%)	51 (62%)
Multiple recidivism (2+ rearrests) $\chi^2=41.73, <.001$	13 (16%)	24 (27%)	22 (40%)	34 (41%)
n=	81	88	55	82

The comparison of simple and multiple recidivism for both short-term and long-term follow-up suggests that the impact of differences in police handling is manifest earlier in multiple recidivism than it is in simple recidivism figures. Evaluators may be well advised to select indicators of impact carefully, according to whether program effects are to be measured in the short or long term. In particular, we have found that multiple recidivism appears

to be a more sensitive short-term indicator of the impact of police disposition decisions on juvenile rearrest rates.

None of the figures presented so far have taken into account the likelihood that offenders with prior records of arrest were overassigned to the most serious disposition, due to the deliberate nonadherence of some police referral officers to the random assignment procedures. The dimensions of this problem are described in table 3, in which rearrest means and frequencies are presented cross-classified by disposition and the presence or absence of a prior record of arrest. A chi-square value significant at the .01 level indicates that subject assignment was biased so that prior offenders were over-represented in the court petition condition. The row means, describing rearrest levels regardless of disposition, reveal that subjects with priors are, indeed, re-arrested twice as often.

TABLE 3.--Simple official recidivism at 15 months,  
by disposition and prior record

		Release	Refer w/o purchase	Refer w. purchase	Court petition	Overall	
PRIOR RECORD	NO	Mean rearrests	.62	.98	1.25	.92	.93
		n =	40(28%)	45(32%)	32(23%)	24(17%)	141
	YES	Mean rearrests	1.00	1.64	2.80	2.71	2.05
		n =	28(22%)	36(28%)	20(15%)	45(35%)	129
		$\chi^2=11.76, <.01$	.75	1.22	1.82	1.95	1.40
		n =	68	81	52	69	270

When rearrest patterns are expressed in terms of means rather than proportions, the pattern across dispositions seen so far remains the same. The pattern within categories of priors is slightly altered, especially for subjects with no priors. In both rows released subjects remain easily the least re-arrested, and subjects referred with purchase of service are rearrested most often. To assess the importance of these differences, analysis of variance tests for main and interaction effects of disposition and prior record on rearrests were performed. The presence of significant interaction effects here would make it difficult to discuss the impact of dispositions independent of prior record. As table 4 shows, the F test for interaction is not significant at the .05 level. The F tests for disposition and prior record were both significant at the .01 level, on the other hand. Together, these factors explained about 11 percent of the variation in rearrests; disposition explained 4 percent of the variance; prior record, 6 percent.

TABLE 4.--Analysis of variance: simple official recidivism  
at 15 months with disposition and prior record

Source of variation	Sum of squares	Df	F	Significance of F
Main effects	138.002	4	8.488	.001
Prior record	68.765	1	16.918	.001
Disposition	52.713	3	4.323	.006
Interaction effects	22.262	3	1.826	.141
Explained variation	160.264	7	5.633	.001
Residual	1,064.913	262		
Total	1,225.177	269		

A discussion of analytic solutions to such problems of selection bias in evaluation research designs appears in a recent article by Alwin and Sullivan, in the journal Sociological Methods and Research.<sup>\*</sup> Because there is no significant interaction between these independent variables, it was possible to employ a covariance adjustment, which removes the effect of prior record from means on the criterion variable. The unadjusted means, which appear as column means in table 3, increase with severity of disposition. Reflecting the bias of referral officers in overassigning prior offenders to the court disposition, the adjusted means change in the expected directions: the mean for court petition falls relative to the means for the three other conditions. The adjusted mean for release is .85 rearrests; refer without purchase, 1.31; refer with purchase, 1.95; court petition, 1.91. It is of interest to note that subjects in the refer-with-purchase condition are rearrested most often of all subjects when means are adjusted for prior record. Also, there is a noteworthy disparity between the means for the court and refer-with-purchase conditions, on the one hand, and the release condition, on the other. Clearly, the most effective treatment for reducing rearrests was to release subjects outright.

To this point, although randomization and adjustment for selection bias have excluded the possibility that background characteristics might have accounted for differing rearrest rates, it would remain unclear whether juveniles not released by the police become more involved in delinquent activities or whether they only find themselves arrested more often for the same level of delinquent activity engaged in by their released co-subjects. To address this problem, reports were obtained from the subjects themselves regarding their delinquent activity, for which they may or may not have been arrested.

<sup>\*</sup>Duane F. Alwin and Michael J. Sullivan. "Issues of Design and Analysis in Evaluation Research," Sociological Methods & Research, Vol. 4, No. 1, (August 1975), pp. 77-100.

Because subjects were contacted for interview about 15 months from the date of the referral arrest, the completion rate was disappointing. One hundred fifteen subjects were interviewed, about 38 percent of the subject pool. The primary reason for this low rate was unavailability of subjects due to mobility, rather than refusal. When respondents were compared with nonrespondents, however, there were virtually no differences in age, sex, prior record, nature of referral offense, mean number of subsequent offenses, and treatment condition. None of the small differences between the two groups were statistically significant. It is reasonable to conclude from this that the respondents adequately represent the entire subject pool.

Surprisingly, respondents across the four disposition conditions did not differ appreciably with respect to the number of illegal activities they reported. Table 5 presents the mean number of offenses reported by respondents, the mean number of times respondents were rearrested, adjusted to control for prior record, and resulting proportion of offenses for which respondents were apprehended. Although the respondents' self-reports do tend to follow the trend across disposition categories for rearrest, none of the self-report differences are statistically significant. Thus, higher rearrest rates for the court and refer-with-purchase conditions cannot be attributed to increased delinquent activity on the part of offenders not previously released by the police. Speculation concerning differing rearrest rates across dispositions should center instead on the possibility of increased responsive efficiency of social control agencies. That is, subsequent to arrest, offenders who have been brought to the attention of either community agencies or the courts are more likely to be rearrested than their comrades who have been released outright, even though both groups of offenders subsequently engage in about the same level of offense behavior.

TABLE 5.--Self-reported delinquency and rearrest at 15 months

	Release	Refer w/o purchase	Refer w. purchase	Court petition	Overall
Mean self-report	49.75	46.88	54.58	53.97	51.17
Mean rearrests	.94	1.41	2.56	1.71	1.62
Offenses apprehended	1.9%	3.0%	4.7%	3.2%	3.2%
n =	28	33	24	30	115

#### IMPLICATIONS

The theoretical implications of these findings appear straightforward in that the more severe societal responses did not deter subsequent rearrests, but rather, tended to encourage them. Released offenders were rearrested least often of the four groups. This tends to support the societal reaction focus of labeling theory, rather than the identity change focus, and it tends not to

support deterrence theory. In addition, the use of both official and self-reported delinquency indices provides more specific indication of the mechanism by which societal reaction tends to encourage rearrest. Rearrests are higher for nonreleased offenders not because their behavior differs from that of released offenders, but because their activities more often become known to authorities, and they are apprehended more often for the same level of offense behavior. This supports the increasing awareness that official crime and deviance rates are considerably affected by the policies of official social control agencies, aside from the activity of individuals subjected to official social control.

The implications for criminal justice evaluation are somewhat similar. In evaluating social programs, it is important to consider more than one indicator of impact, and to examine unofficial indicators as well as those routinely produced by the very office or sector being evaluated. In the course of this project, not only did we find that multiple recidivism proved to be a more sensitive short-term indicator of the impact of disposition on rearrest, but that rearrest data taken alone provided a misleading picture of the impact of police handling. Although first multiple recidivism and then simple recidivism showed that disposition alternatives do affect subsequent rearrests, self-reported delinquency indices shed new light on the specific mechanism in this effect. Whereas the official figures might lead us to believe that more severe police handling aggravates delinquent behavior, unofficial interview figures show that this is not the case. Evaluators are well-advised to employ multiple indicators in the attempt to clarify some of the complex possibilities of program impact.

Finally, these data have rather surprising implications for police diversion programs and policy administrators. It may be that referral to community counseling agencies with purchase of service may actually result in higher official recidivism than outright release. However, none of the dispositions considered here had a significant impact on behavior, relative to the other dispositions. Administrators can expect difficulty explaining to the public that rearrest rates for diversion programs are a function of improved surveillance rather than of aggravated delinquent behavior, especially if public opinion favors eliminating expensive social programs. It is conceivable that increased surveillance could be adopted as the goal of diversion programs, although according to these data that would not reduce delinquent behavior compared to low surveillance techniques. However, if the goal of diversion programs is to reduce rearrests, the present data would suggest that it may prove impossible to justify them on that basis. The disposition alternative that resulted in the lowest official recidivism was also the cheapest disposition--outright release.

VOLUNTARY PRETRIAL DIVERSION AND THE QUESTION OF COMPLIANCE:  
A PRELIMINARY EVALUATION

Diane L. Gottheil, Ph.D.  
Director, Adult Diversion Program  
Champaign County, Illinois

Many of the more serious criticisms of pretrial diversion programs that appear in the literature are related to the maintenance of control over program participants by legal authorities. Customarily, criminal charges against participants are pending and dismissal occurs only after favorable termination by program personnel. Some programs accept clients precharge, but with the provision that unfavorable termination will result in the filing of charges and prosecution. The voluntary aspect of most programs, therefore, rests exclusively on the individual's choice of whether to participate in diversion rather than to allow the charge to be pursued through the traditional criminal justice process. Even for those individuals uncertain themselves about their culpability for the offense, the merits of the evidence on which the charge is based, or the likelihood of prosecution, the alternative of pretrial diversion is one of those offers that cannot be refused.

Gorelick (1975) has described several issues of concern to those observers who note that diversion may be a means of delaying the operation of criminal processing, rather than diverting people out of the criminal justice system. Diversion formalizes and broadens preexisting discretionary prosecutorial practices, he asserts, by delaying further processing pending the meeting of certain conditions. Social control is expanded to individuals who, in fact, have not been convicted of a criminal offense. Moreover, eligibility requirements may enable only participants likely to succeed to be diverted, and these tend to be cases with low priority for prosecution and least likely to be convicted. Thus, diversion may bring people into the criminal process who would not have been processed at all, or who would have been screened out at an early stage.

Other issues relate to the reinstatement of criminal charges against unfavorably terminated participants in diversion, the threat of which is the mechanism of control. The defense effort may be prejudiced in that unfavorable termination in a diversion program may reduce the probability of negotiating a favorable plea, or such prosecution may be given higher priority by prosecuting authorities. The defendant's right to a speedy trial is usually waived at time of entering a diversion program. If several months of participation is followed by unfavorable termination resulting in prosecution, such a waiver may reasonably be questioned. Furthermore, if the threat of prosecution is not carried out--due to priority given to more serious offenders, not eligible for diversion, and the customarily high case load of prosecutors--program credibility may be reduced. Effective prosecution of such cases may even be

hampered by the time lapse between the offense and the unfavorable termination of the diversion participant. Finally, the claimed costs savings of diversion are reduced if the prosecutor and courts must deal with unsuccessful diversion clients.

The exercise of control over participants is behind the criticism that diversion programs may just be mirrors of probation. Minimally, they require a period of supervision-rehabilitation; at the most extreme, they delay prosecution to a later time, after some participation in rehabilitative services and perhaps followed by conviction and probation or incarceration. Moreover, even for those participants favorably terminated, although not subject to traditional prosecution they were placed in another system that may be as intrusive as the criminal justice system given the element of official control. In sum, diversion programs, it is charged, do not truly divert people out of the criminal justice system.

This paper describes a pretrial diversion program that avoids many of these criticisms by enabling client participation that is genuinely voluntary. The Adult Diversion Program of Champaign County, Illinois, began operation on October 1, 1975. Individuals arrested for criminal offenses who are eligible for diversion are referred to the program by the Champaign County State's attorney. Referrals are made before formal charges are filed and if the individual is accepted into the program, no charge is filed. Furthermore, once accepted into the program, the participant's behavior in carrying out the agreed-upon terms of participation, and the specification of the terms themselves, is completely voluntary. That is, there are no legal consequences of unfavorable termination and thus no means of coercion or compulsion. The prosecutor has agreed that under no circumstances will charges be filed against individuals accepted by the program for the offense for which they were referred. The accused is genuinely diverted out of the criminal justice system.

Upon acceptance, participants meet with a Diversion Counselor to develop a program agreement. This agreement specifies the length of participation (from 3-12 months) and the terms. Terms are defined as goals in three dimensions: program goals, short-term goals, and long-term goals. Terms specifying regular meetings with the counselor to deal with personal or family problems, enrolling in a school or college, obtaining vocational training, or seeking employment may be placed in any of these categories. Such terms as restitution to a victim, volunteer service to a community group or organization, or follow-up on referral to an alcohol treatment program, however, are usually considered program goals. An attempt is made to limit the intrusiveness of the program while providing support to the client in areas of agreed-upon needs.

Evaluations of diversion or deferred prosecution programs generally seek to demonstrate that the program has succeeded in achieving one or more of the following goals:

- (1) a cost savings in comparison to traditional prosecution,
- (2) a reduction in the rate of recidivism,
- (3) effective delivery of services, e.g., a high number of job placements, or

- (4) a high percentage of favorable terminations, i.e., prosecution was not reinstated, or charges were dismissed, for most program clients.

The Adult Diversion Program is being evaluated with respect to the first three of these goals, but the focus of this paper is on a variation of the fourth goal. The voluntary nature of participation in the ADP is, as far as has been determined, a unique feature. The implications of this aspect of the program are potentially far ranging in light of the substance of criticisms of diversion reviewed above. Participants in a voluntary diversion program are no longer subject to possible prosecution for the offense of which they were accused. Although they are asked to take responsibility for the offense, this has no legal implications. Rather, it is considered a criterion of the appropriateness of diverting them out of the criminal justice system to a nonpunitive program of social services.

The obvious drawback to any voluntary diversion program is an expected higher probability of noncompliance. Inasmuch as diversion programs have the dual goal of reducing the burden on the criminal justice system, and providing social services aimed at rehabilitation to keep people from reentering the system, some level of compliance would appear necessary to goal attainment as well as the integrity of a program. Favorable termination in diversion customarily requires no subsequent law violation and the carrying out of other terms of a program contract or agreement. Lacking means of compelling compliance, can a program that in effect offers clients "a free ride" demonstrate an acceptable level of favorable terminations?

Program personnel were aware from the outset that due to its uniqueness the voluntary aspect would require periodic scrutiny. For the purpose of recordkeeping and for evaluating level of compliance, termination types were defined and as participants were terminated each case was placed in one of the categories by rigorous application of the definitions. The decision to terminate a participant is always made by the counselor. If the participant has carried out the terms of the program agreement satisfactorily, it is considered a favorable completion. Decisions on such terminations are made in consultation with the client and are dependent on mutual agreement that the client no longer needs program services. Although charges are not filed against any participant in the Adult Diversion Program, noncompliance with the program agreement is considered a termination in one of five other categories.

The first year of program operations was chosen as a cutoff point for inclusion in a sample of participants. Inasmuch as some sample members were still active participants on January 1, 1977 (15 months after the initial participant was accepted), the evaluation presented should be considered preliminary. During the first year of operation, from October 1, 1975 through September 30, 1976, of 288 individuals referred, 200 were accepted into the Adult Diversion Program. Data on current status for the 200 first-year participants and on termination type for those terminated are presented in table 1. Most of those 43 individuals who remained active program participants as of January 1, 1977, had been accepted into the program during the months of June through September 1976. The termination categories are based upon those developed by the Illinois Law Enforcement Commission for the purpose of that organization's ongoing evaluation of Illinois deferred prosecution and diversion programs.

TABLE 1.--Current status and termination category of  
first-year participants  
(N=200)

Current status: January 1, 1977	Number	Percent of all cases	Percent of terminations
Active:	43	22%	
Terminated:			
Favorable completion	87	43	55%
Did not complete	36	18	23
Voluntary withdrawal	16	8	10
Abscond	12	6	8
Conviction	5	2	3
Other	1	-	-
Total terminated:	157	100%	100%

Favorable Completion. The individual has carried out the terms of the program agreement; i.e., (s)he has carried out program goals and has made progress toward achieving other short-term or long-term goals. Of those cases terminated by January 1, 1977, over half (55 percent) carried out all terms of their program agreements. It should be pointed out that participants who are favorably terminated tend to remain active longer than those in other categories. We would expect, therefore, a somewhat higher proportion of the currently active cases to be favorable terminations, in comparison to those already terminated.

Did Not Complete. The individual has not carried out program agreement terms and either discontinued meeting with the counselor or has otherwise indicated that terms are unlikely to be carried out. The 36 cases in the Did Not Complete category represent 23 percent of those terminated. In general, cases so categorized are individuals who did not carry out one or more of the terms of their agreement with the ADP. This may be, for example, failure to make contact to keep an appointment with a referral agency or failure to do volunteer work. Rarely has a participant failed to pay restitution to a victim, even when this is the only program goal they succeed in carrying out.

Voluntary Withdrawal. The individual has explicitly stated that the program agreement terms will not be carried out and/or (s)he will not continue to meet with the counselor. Generally, unwillingness to cooperate in a counseling relationship, rather than unwillingness to carry out other types of program terms, differentiates this category from that of the previous. The 17 cases in the Voluntary Withdrawal category are 12 percent of those terminated. These are not completely dissimilar from the previous category, but indicate that the termination decision was made rather explicitly by the client rather than by the counselor who makes the termination decision for other categories. In other diversion or deferred prosecution programs, where prosecution is a consequence of unfavorable termination, this distinction is more significant.

Absconds. The individual fails to meet with the counselor after being accepted into the program, or has an initial meeting but fails to work out and sign a diversion agreement and to meet again with the counselor. To date, 13 cases accepted into diversion either failed to appear for their first counseling meeting or appeared only for this meeting without establishing a program agreement. This represents 9 percent of cases terminated.

Conviction. The individual is arrested and convicted for an offense committed subsequent to acceptance into the program and prior to termination. Five cases have been terminated for this reason, 3 percent of the total terminations.

Other. This category has only been used once for an individual charged with an offense committed prior to referral to ADP, who was arrested for this prior offense while a diversion participant. He was terminated after being convicted and sentenced to prison outside the community.

In the absence of any external standard of success, no target figure for favorable terminations was established. Program goals refer to successful diversion but this paper does not purport to evaluate overall program success. For some, diversion of 200 individuals out of the criminal justice system is a sufficient demonstration of success, provided the recidivism rate is acceptably low. The goal here, however, is to evaluate compliance behavior. With no predetermined figure for an acceptable "drop-out" rate, the discussion will focus upon those independent variables that are related to compliance rates and some additional factors to be considered in evaluating compliance behavior.

It is proposed that three independent variables (or groupings of variables) affect the level of compliance: (1) the staff service provided to clients; (2) the structure of the program; and (3) characteristics of the client population. The higher the quality of the services delivered by staff to participants, the higher the proportion of clients who favorably complete the program. In the spirit of its voluntary nature, Adult Diversion Program counselors have adopted a style of assertiveness, but not aggressiveness. Supervision of clients is defined as maintaining contact in order to aid in reinforcing commitment to program agreement terms--or in the case of counseling, to carry out these terms--and being accessible for supportive or advocacy services. Measuring the effectiveness of counseling service is a difficult enterprise. Data on client responses to counselors, for example, cannot be gathered from those clients that drop out. Analyses of service would require before/after measurements of participants using several indicators, and this is not within the scope of this paper. In any case, while counseling service is indeed a contributing factor, its effectiveness cannot be gauged by outcome alone, inasmuch as it is strongly related to the other factors identified above, i.e., structure of the program and characteristics of the client population.

The structure of the program is primarily characterized as voluntary, and as stated above, this characteristic is likely to have a strong effect on compliance behavior of program participants. Comparison of the figures in table 1 may be made with figures from other diversion programs. A cursory review of materials available suggests that other programs tend to report favorable terminations in the 80 to 95 percent range. Were the Adult Diversion Program to eliminate its voluntary structure, the percentage of favorable terminations would no doubt be within this range. The various criticisms of diversion

discussed earlier, however, suggest that it may not be desirable to sacrifice the voluntary structure of the program for a higher percentage of favorable terminations. In any case, just what is an acceptable level of favorable terminations has not yet been established.

With this in mind, it should be emphasized that in placing individuals in categories of termination, a voluntary program can very rigorously apply the category definitions without concern for consequences for the participants. Where prosecution is a consequence of all terminations other than favorable, it is improbable that many of the cases in the "Did not complete" category in table 1 would, indeed, be returned for prosecution. Many of these clients have carried out some terms of their program agreement and few show evidence that they are likely to be involved in a further law violation. In other words, in evaluating overall program performance, some portion of the drop-outs might be considered "successes." Moreover, a voluntary compliance rate of over 50 percent may be considered high by some standards. Rather than comparing this figure with diversion programs that have mechanisms for compelling compliance, a more appropriate comparison may be with data from other voluntary social service organizations, such as community mental health clinics. The proportion of clients who continue to followup in counseling or therapy is closer to the 55 percent favorable termination in the Adult Diversion Program.\*

Another question to raise concerning a voluntary structure and compliance is whether anything is to be gained in a nonvoluntary program given the dual goals of diversion and rehabilitation. It has become widely accepted that genuine rehabilitation cannot occur in a coercive setting. This applies to institutional settings in particular, but it may be accurately generalized to diversion programs. Would greater positive change be induced in people by requiring that they comply under threat of prosecution? Compliance would occur and there may be positive outcomes to that alone. It might also be argued that some people, compelled to meet with a Diversion Counselor and carry out certain program agreement terms, would experience positive attitudinal or behavioral changes as a result. The question is how great the extent of such change or the number of people involved, and whether this is worth the trade-off of a truly voluntary program.

The integrity and credibility of a diversion program is perhaps less threatened when compliance rates are high. This may be an important consideration for gaining community acceptance and for staff morale. Nevertheless, in a coercive setting, program personnel are likely to be perceived as authority figures and overt compliance is not likely to go beyond what is judged by the client as minimally acceptable. Overt compliance may often also be accompanied by covert resistance as a result of resentment or hostility to authority.

Mental health and counseling professionals have recognized that individuals must be "ready" for change; that is, willing to work on defining their

---

\* Although data from mental health centers are difficult to gather and do present some problems of comparability (e.g., counselor-client contracts are open-ended), figures for Illinois report that about two-thirds of mental health clinic outpatients leave the counseling relationship.

problems and seeking alternatives or solutions. Change cannot be imposed upon people with expectations of a great degree of success. The participants in a voluntary program who favorably complete their programs, and even those in other termination categories who carry out some terms of their program agreements, do so on their own initiative and demonstrate self-directed responsible behavior. A voluntary program does not seek to impose change but attempts to motivate compliance by providing an active support person as a counselor and a program agreement consisting of mutually developed and agreed-upon needs. This situation can help the client identify opportunities or resources previously unknown or thought to be unobtainable. Motivation to comply that exists or is developed tends to be based upon self-interest rather than response to authority. Program personnel, therefore, devote their resources to those individuals who choose to take advantage of the opportunity of services provided by the program. This may in itself be a strong argument in support of retaining a voluntary structure and should be considered in determining an acceptable level of noncompliance.

The third independent variable to be considered as it affects compliance is client characteristics. Eligibility requirements of any program can produce a client population representing those most likely to comply with a diversion agreement. These tend not to be those people, however, most in need of program services. It is among those most in need that the "risk" of noncompliance is highest. The screening and acceptance process in the Adult Diversion Program and the consequent characteristics of those accepted for diversion undoubtedly affect compliance. In a voluntary program, there may be an even greater temptation to screen out those clients whose current offense, prior record, and attitudes or life situation suggest a higher risk of noncompliance. The Adult Diversion Program eligibility criteria are nevertheless rather broad, although the initial consideration depends upon referral from the prosecutor's office. The decision to accept an individual into the program is based on a recommendation from a three-person citizen screening panel which interviews each applicant. These panels are drawn from approximately 35 Citizens' Advisory Committee members who have volunteered and are trained to serve on the panels on a rotating basis. Impressionistic evidence suggests that volunteers from the community appear more likely to accept higher risk participants into the program than do staff. This is considered a positive contribution of citizen participation in that it helps to assure that program services are available to those most in need.

Although a complete analysis of client characteristics is not within the scope of this paper, it may be useful to present data from some indicators of probability of client compliance in order to better interpret the results presented in table 1. Two different measures of the probability of compliance were used. The first combines the current offense and prior offense record of each client as an index of low or high risk of noncompliance (Risk Index No. 1). The second is based upon a subjective assessment of each individual referred to the program (Risk Index No. 2); following an initial assessment interview with individuals referred, the program's Intake Officer scores each as high or low risk of noncompliance. The scores on each of the indices for all cases terminated as of January 1, 1977, were cross-tabulated with termination type (combining all categories other than "Favorable completion"). The frequencies and percentages are reported in table 2.

TABLE 2.--Indices of risk of noncompliance by termination type

<u>Risk Index #1</u>	<u>Favorable completion</u>		<u>Total</u>
	<u>Yes</u>	<u>No</u>	
Low	74 (62%)	42 (38%)	120 (100%)
High	13 (35%)	24 (65%)	37 (100%)
<u>Risk Index #2</u>			
Low	68 (70%)	29 (30%)	97 (100%)
High	19 (32%)	41 (68%)	60 (100%)

It can be seen that both indices appear to have some predictive ability of the probability that an applicant will carry out his/her program agreement. The index based on current offense and prior record--Risk Index No. 1--suggests that applicants accused of misdemeanors or felonies with no prior criminal record are almost twice as likely to be favorable terminations than those accused of felonies who also have some minor or more serious prior offense or accused of misdemeanors and also have a prior conviction record. The index based on the subjective assessment of probability of noncompliance--Risk Index No. 2--appears to be an even better predictor.

The figures presented in table 2 are only suggestive and the indices are not proposed to be sophisticated measuring devices. The figures do suggest, however, that a change in screening and acceptance policy might be an effective method of reducing the rate of noncompliance in diversion without changing its voluntary structure. Indeed, many diversion programs carefully "screen out" these individuals who they assess to be "unready" for change. However, this assessment is difficult to make with complete accuracy and, furthermore, a trade-off is clearly involved here as well. By eliminating from diversion those least likely to succeed a large portion of those people most in need of program services are also eliminated. What may be most significant about the figures in table 2 is that approximately one-third of those in the high risk category of each index do in fact voluntarily carry out the terms of their program agreement.

One final consideration before concluding is the question of recidivism in a voluntary program. Some members of the Champaign County Adult Diversion Program's Citizen's Advisory Committee have expressed the point of view that compliance is a low priority concern for evaluating pretrial diversion. The argument is made that if those individuals who drop out are not involved in subsequent law violations, then diversion was an appropriate alternative. Such an argument is advanced in the context of greater concern for reducing the backlog in the criminal justice system and perhaps for the apparent lower costs of diversion. Others express this point of view in terms of providing humane, nonpunitive alternatives to the impersonal processing of the criminal justice system which is, in any case, a system designed to deal with criminals. This argument holds some weight only if the rate of recidivism among diversion

participants--and especially among those who drop out--is acceptably low. If not, the individuals are soon back into the system and little reduction in the backlog or costs savings have been realized. Furthermore, a high recidivism rate might suggest that a system designed to deal with criminals may be more appropriate.

The rate of recidivism is not likely to be affected by the voluntary nature of a program; the threat of prosecution for the original offense is not likely to deter subsequent offenses any more than does the Adult Diversion Program's present practice of informing participants that any new offenses will be prosecuted and they will not have another opportunity for diversion. It may be argued that programs with mechanisms of compelling compliance will experience more frequent contact of clients with staff, and consequently more opportunity to solve problems and prevent subsequent law violations. This stretches the point made earlier, however, that greater compliance might produce some change in some clients. It also points to the criticism that diversion programs mirror probation for it suggests a supervisory type relationship with program participants that tends to be quite intrusive. Acceptance of higher risk clients, rather than a voluntary structure, is more likely to affect the rate of recidivism for the very same reason that it affects the level of compliance. Those people less likely to carry out their program agreement are also more likely to get involved in new law violations.

The concept of an acceptably low rate of recidivism, however, raises the obvious question of what is acceptable. To those of us concerned with achieving valid and reliable evaluation techniques and methods, an experimental design alone can perhaps provide the data required for responding to this question. This applies as well to the problem of evaluating the overall success of a voluntary diversion program. In the particular setting of the program examined here, as in most other settings, implementation of a control group design is problematical. In fact, involved community people who were instrumental in initiating the program explicitly rejected a proposal for such a design. In the absence of an experimental design, community acceptance with respect to both the rate of recidivism and the level of compliance may be a legitimate criterion of success. This is particularly true for a voluntary program in which the prosecutor has decided to divert a significant number of criminal offenders out of the criminal justice system, and by mutual consent with diversion program personnel, has relinquished his control over these individuals. In addition, systematic analysis and display of data from the program is a necessary ingredient of providing accountability, especially to criminal justice and law enforcement agencies and the attentive public. If costs savings and reduction of the case load in the criminal justice system can be demonstrated, one question that remains to be answered is whether a particular community is willing to tolerate a certain proportion of noncompliance among diversion participants and what the limits of such tolerance may be.

In conclusion, this paper has attempted to review the early experience of a voluntary diversion program and to suggest some of the factors that should be considered in developing such programs or proposing changes for existing programs. Given the multiple and sometimes conflicting goals of diversion, it seems unlikely that any program design can avoid vulnerability to criticism. Program personnel need to be aware of what features of their program may be subject to criticism, and to be able to place any criticism in a perspective that specifies trade-offs involved in proposed structural or procedural changes.

## BIBLIOGRAPHY

Gorelick, Jamie S. "Pretrial Diversion: The Threat of Expanding Social Control," Harvard Civil Rights--Civil Liberties Law Review, vol. 10, 180-214.

## PROSECUTORIAL POLICY, IMPACT, AND IMPLEMENTATION\*

Joan E. Jacoby  
Research Associate  
Bureau of Social Science Research

### A. INTRODUCTION

In 1975, the Law Enforcement Assistance Administration (LEAA) supported through its National Evaluation Program a Phase I Evaluation of Pretrial Screening Projects<sup>1</sup> to determine whether a nationwide evaluation was feasible. Based on an examination of the issues and observation in 19 prosecutor's offices throughout the nation, the study concluded that such an evaluation was not desirable until further research was performed. While this finding was primary to the Phase I study, another important discovery was made; namely, the overwhelming effect of prosecutorial policy on an office's performance as seen by its disposition rates, its allocation of resources, and use of various prosecutive strategies. Additionally, it was found that these policies could be classified and presented in a model typology format.

To quote Lewis Katz's Preface to the study:<sup>2</sup>

"The material developed by Joan Jacoby and her associates. . . offers a major breakthrough in the development and understanding of case screening. . . . The typology. . . sets forth a methodology for evaluating screening mechanisms in operation in a given office to determine whether those devised are fulfilling the purposes for which the prosecutor adopted them. Nowhere does Jacoby or her cohorts seek to dictate a policy to prosecutors, instead they recognize that there are many possible policies from which a prosecutor may choose one or more. Their role was not to evaluate or criticize those policies but to develop measures which would enable a prosecutor to determine whether his or her policies are being serviced. The typology deals with four of the policies and provides tools for measurement at each stage of the criminal process."

The purpose of this paper is to

- (1) summarize the findings of this study,

---

\*This paper is excerpted from: The Prosecutor's Charging Decision: A Policy Perspective, a prescriptive package published by LEAA, January 1977.

- (2) show that the identification of policy is a primary requirement in an evaluation effort,
- (3) illustrate how differing strategies and allocation patterns are needed to support different policies, and finally
- (4) make the prosecutor and public aware of the policy choices available and the effect of these choices on their local budgets as well as their communities.

The latter purpose is important. Already certain trends are occurring that give dominance to questioning the type of prosecutorial services we can afford. More and more the prosecutor is being pressured to become increasingly selective in what he is going to prosecute and how that prosecution should proceed. As a result, policy criteria have to be stated so that sound decisions can be made. The workload created by the demand for improved delivery of legal services, the increased utilization of public defender agencies, the impact of the Argesinger decision,<sup>3</sup> and increased system efficiency could result in one response by an increase of staff at additional public expense. More likely, however, it will require the prosecutor to become more selective in accepting cases for prosecution.

As states examine the possibility of abolishing plea bargaining (as has occurred in Alaska) or as individual prosecutor's offices move to this stance, success can only be fostered if court capacity is increased to meet trial needs, alternatives to criminal justice processing are expanded, and screening of cases for proper charging is considered one of the most important decisions to be made in an office. What the public should fear is not the tough prosecutor but the sloppy one. Where more intensive scrutiny of cases occurs, the probability of prosecuting the innocent defendant is diminished.

Prosecutorial policy is not circumscribed by the limits of the office. Its impact can be measured in other criminal justice agencies, particularly corrections. Depending upon the prosecution policy, the future quantity and characteristics of the correctional population can be anticipated. Where treatment programs are used, prosecutorial policy may well indicate the needs and requirements of such activity. Needless to say, this predictive power can be turned into a highly effective planning and management tool as well as a testing mechanism for attempted solutions to some of the problems confronted by the criminal justice system.

The remainder of this paper will examine the four "ideal" policies that are likely to be found in one form or another in prosecutor's offices. The results of the policies will be discussed in terms of expected case disposition rates and the strategies and resource allocation patterns that support the policies will be briefly examined.

## B. PROSECUTORIAL POLICIES

No matter what the external environment or a prosecutor's perception of his discretionary authority, the prosecutor operates with a policy (usually either the one for which he was elected or the one inherited) and implements the policy by various strategies. At first thought, one might expect the policy of the District Attorney to vary as much as the characteristics of the

approximately 3,400 prosecutors vary. However, our experience and observation have shown that generalized classifications can be delineated. Of the four discussed here, all have been observed operating in almost "pure" form in offices throughout the United States. The differences due to policy were so startling that the abstraction of these operations into policy models was not a difficult task. The examples presented below are therefore discussed as ideals or models. It should be emphasized that the four policy types presented here are neither exhaustive nor mutually exclusive. We are quite sure that other policies exist which result in different treatment modes and disposition patterns. We have also observed that in some offices a mix of these policies exists. For the purpose of this presentation, however, the policy types have been abstracted and presented as pure types.

The policies have been given the abbreviated descriptive titles of Legal Sufficiency, System Efficiency, Defendant Rehabilitation, and Trial Sufficiency.

Legal Sufficiency Policy. Some prosecutors believe that if any case is legally sufficient (if the elements of the case are present), then it is their responsibility to charge and prosecute. For example, in a breaking and entering case, if there was evidence of forcible entry, that is, if the entry was without the permission of the owner and if the person arrested was found to have in his possession items belong to the victim, the case would be prosecuted because it was legally sufficient. The elements of the case are present. However, what may on the surface seem to be a prosecutable crime, may indeed be lost because of constitutional questions, for example, an illegal search and seizure. Implementing this policy at the charging level requires only an examination for legal defects. If the basis for a charge is not legally sufficient, either additional investigation could be ordered or the case would be rejected. The legally sufficient policy is most prevalent in the lower, misdemeanor courts. Here cases are routinely but quickly examined for obvious defects prior to court appearance. This is usually the extent of screening that a case receives. As a result, overloads occur, plea bargaining is encouraged to reduce the volume and, with scant case preparation time, dismissals and acquittals abound. This is a policy quite common in lower, misdemeanor courts handling large volume workload. (See figure 1.)

System Efficiency Policy. Another prevalent policy can be labeled "system efficiency." It aims for the speedy and early disposition of cases by any means possible. Time to disposition and the place in the court process where disposition occurs are measures of success in addition to favorable dispositions. Under this policy, the breaking and entering case would be rejected because emphasis is placed on screening as a way to minimize workload and the search and seizure issues would have been spotted. If there were no search and seizure issue, the case would have been accepted, charged as a felony, and the defendant would be allowed to plead at the committing magistrate hearing to a reduced charge of unlawful trespassing or larceny (both misdemeanors). This policy usually emerges when the court is overloaded and heavily backlogged, and the resources of the prosecutor extremely limited.

Under these conditions, in addition to the emphasis placed on pretrial screening, any other method of case disposal that can be found will be used. The prosecutor himself may be an active searcher for additional avenues of case disposition. Cases will be examined for their ability to be plea bargained (to achieve this overcharging may occur). Extensive use will be made

FIGURE 1.--Legal sufficiency policy  
expected frequency of dispositions

POLICY: If the elements of the case are present, accept for prosecution.

Disposition universe (numeric base for rates)	Disposition	Frequency
Cases presented	Reject for prosecution	Low
	Accept for prosecution	High
	Divert--non-CJS	Not predictable
	Refer--other CJS	High
Cases accepted	Dismiss at preliminary hearing	High
	Bound over	Minimize
	Plea to reduced charge	Maximize
	Plea as charged	Low
Cases bound over	No true bill (grand jury only)	High
Trials	Guilty--trial	Low
	Acquittal--trial	Low
	Dismissed--trial (insufficient evidence)	High

of community resources, other agency resources, and diversion programs so that cases may be kept out of the criminal justice system. Charges will be broken down for handling in the lower courts, if possible, or modified and referred to another court with a different jurisdiction (e.g., a county court case referred to municipal court). The full utilization of the court's resources and the charging authority will be made to dispose of the case as soon as possible. Particular emphasis will be placed on the disposal of the case prior to a bind-over to the higher court or grand jury. (See figure 2.)

Defendant Rehabilitation Policy. A third approach, based on a policy of rehabilitating the defendant, utilizes some of the elements of the early and speedy disposition policy but should not be confused with it. Under this policy, the prosecutor believes that the most effective treatment for the majority of defendants who pass through his office is any alternative treatment other than processing through the criminal justice system and more particularly, through the correctional system. He believes that any treatment other than this is better for the vast majority of defendants. To cite our breaking and entering case again, if the defendant were a first offender or had a drug problem and restitution was made to the victim, he might very well be placed in a pretrial diversion program or if none available, and with the court's concurrence, he could receive a sentence of probation without verdict. The charging and prosecution decision depends primarily on the circumstances of the defendant and secondarily on the offense that he was alleged to have committed.

FIGURE 2.--System efficiency policy  
expected frequency of dispositions

POLICY: Dispose of cases as quickly as possible, by any means possible.

Disposition universe (numeric base for rates)	Disposition	Frequency
Cases presented	Reject for prosecution	Not predictable
	Accept for prosecution	Not predictable
	Divert--non-CJS	Maximize
	Refer--other CJS	Maximize
Cases accepted	Dismiss at preliminary hearing	Low
	Bound over	Minimize
	Plea to reduced charge	Maximize
	Plea as charged	Low
Cases bound over	No true bill (grand jury only)	Not predictable
Trials	Guilty--trial	High
	Acquittal--trial	Low
	Dismissed--trial (insufficient evidence)	Low

Thus the goal is the early diversion of many defendants from the criminal justice system coupled with serious prosecution of cases allowed into the system. It is logical to expect vigorous prosecution of this latter category especially if the defendant's history includes prior convictions with no evidence of rehabilitation. Offices using this policy tend to rely heavily upon the resources in the community as well as in the criminal justice systems to move eligible defendants out of the judicial and correctional systems. A close cooperation with the court often ensues, particularly in using the sentence recommendation power of the prosecutor to insure consistency in the recommended treatment plan for the defendant. (See figure 3.)

Trial Sufficiency Policy. The fourth policy in less common use is that of trial sufficiency. This policy states that a case will be accepted only if the prosecutor is willing to have it judged on its merits and expects a conviction. Under these circumstances, the prosecutor views his prosecutorial responsibility very stringently but not without leniency. If a decision was made to charge the defendant of our hypothetical breaking and entering case, and again if the constitutional question of the search was overcome, the defendant would be charged with the felony and a conviction expected at this level. Under this policy once the charge is set, it is difficult to change. To implement this policy, good police reporting is required since the initial charging stage closes out most options. It also requires alternatives to prosecution since not all cases will be prosecuted. Most importantly, it requires court

FIGURE 3.--Defendant rehabilitation policy  
expected frequency of dispositions

POLICY: Divert, since the vast majority of defendants cannot benefit from criminal justice processing.

Disposition universe (numeric base for rates)	Disposition	Frequency
Cases presented	Reject for prosecution	Not predictable
	Accept for prosecution	Minimize
	Divert--non-CJS	Maximize
	Refer--other CJS	High
Cases accepted	Dismiss at preliminary hearing	Low
	Bound over	High
	Plea to reduced charge	Not predictable
	Plea as charged	Not predictable
Cases bound over	No true bill (grand jury only)	Low
Trials	Guilty--trial	High
	Acquittal--trial	Low
	Dismissed--trial (insufficient evidence)	Low

capacity since each case accepted is expected to go to trial. Finally, this policy, as compared to the others, mandates the tightest management control in the office to insure that the initial charge is both proper and, once made, not modified or changed. (See figure 4.)

#### C. A CHARGING TYPOLOGY

Using the policies just described, it is possible to develop models that

- (1) demonstrate the various goals that are consistent with each policy and
- (2) predict the expected outcomes for each policy and goal.

A comparison of the models illustrates the power of policy by showing that unless policy is taken into account, it is impossible to determine prosecutorial effectiveness from dispositional data.

Figure 5 shows goals and predicted outcomes for each of the four policies previously discussed. The goals are shown in terms of outcomes which should be maximized (Mx in the figure) or minimized (Mn) for each policy. Because of the interrelatedness of the prosecutorial process, once these goals are

FIGURE 4.--Trial sufficiency policy  
expected frequency of dispositions

POLICY: If a case is accepted for prosecution, it will be charged at a level capable of sustaining a conviction or a plea to charge.

Disposition universe (numeric base for rates)	Disposition	Frequency
Cases presented	Reject for prosecution	High
	Accept for prosecution	Low
	Divert--non-CJS	Not predictable
	Refer--other CJS	Not predictable
Cases accepted	Dismiss at preliminary hearing	Minimize
	Bound over	High
	Plea to reduced charge	Minimize
	Plea as charged	High
Cases bound over	No true bill (grand jury only)	Low
Trials	Guilty--trial	Maximize
	Acquittal--trial	Low
	Dismissed--trial (insufficient evidence)	Minimize

established other outcomes may be expected to occur with predictable regularity. Some dispositions are expected to occur with high frequency (H in the figure), others with low frequency (L), while the likelihood of other outcomes appears to be independent of the policy and goals in some instances (shown as N in the figure).

The reader is cautioned to recognize that the frequencies listed as high or low do not have numerical values at this time. They are instead high or low relative to the universe specified for each disposition. It is expected that the designation of what is high or low relative to the universe will be defined by each individual prosecutor's office. Whether numbers can be generated that have nationwide applicability is yet to be determined.

If one reads across the row for any particular disposition in figure 5, it is obvious that the expected dispositional values may change drastically depending on the policy being used. For example, the number of cases dismissed<sup>4</sup> at preliminary hearing or a probable cause hearing is expected to be high under the Legal Sufficiency policy. Because cases receive only routine screening for obvious defects, other more serious defects may not be noticed until this point in processing is reached. Use of this policy also suggests that the courts are relied upon to function as the determinant of legal sufficiency rather than the prosecutor. On the other hand, the low dismissal

FIGURE 5.--Expected frequency of selected dispositions as a function of policy

Disposition universe (numeric base for rates)	Dispositions	Policies			
		Legal sufficiency	System efficiency	Defendant rehabilitation	Trial sufficiency
Cases presented	1. Reject for prosecution	L	N	N	H
	2. Accept for prosecution	H	N	Mn	L
	3. Divert--non-CJS	N	Mx	Mx	N
	4. Refer--other CJS	H	Mx	H	N
Cases accepted	5. Dismiss at preliminary hearing	H	L	L	Mn
	6. Bound over	Mn	Mn	H	H
	7. Plea to reduced charge	Mx	Mx	N	Mn
	8. Plea as charged	L	L	N	H
Cases bound over	9. No true bill (GJ only)	H	N	L	L
Trials	10. Guilty--trial	L	H	H	Mx
	11. Acquittal--trial	L	L	L	L
	12. Dismissed--trial (Insuff. evid.)	H	L	L	Mn

128

KEY

GOALS

Mx - Maximize this disposition  
Mn - Minimize this disposition

EXPECTED OUTCOMES

H - High frequency  
L - Low frequency  
N - Not predictable

rate expected for the System Efficiency policy and the Defendant Rehabilitation policy may be traced to the fact that relatively few cases are being processed through a preliminary hearing under the System Efficiency policy since the tendency is to first screen, then "break it down and plead it," thereby producing fewer cases at this level. Those cases that do survive are probably better prepared since they are likely to be nonpleadable. The same pattern occurs for the Defendant Rehabilitation policy but for different reasons. Namely, all the lesser defendants (cases) have been handled by other means, with the remaining cases being the most serious defendants who are vigorously prosecuted. Finally, the Trial Sufficiency policy, which anticipates trial and conviction, mandates that dismissals be minimized since, if one occurs, it is a direct reflection on the quality of the intake division's decision and may point up errors on their part.

A special note should be made of dispositions by dismissal, not all of which may be adverse measures of prosecutorial performance. As already cited, a dismissal of other pending cases may be sought after a conviction has been obtained on another case. In other instances, the case may be dismissed because the complaining witness refused to prosecute, the police officer failed to show, or the defendant was placed in a medical or health treatment facility. The dismissals that should be used to evaluate the performance of the prosecutor are those which reflect an insufficient case or lack of adequate preparation. Thus, generally, they can be classified as "dismissed--insufficient evidence." One would expect this to be a relatively high outcome under the Legal Sufficiency policy, since only cursory examination is given to a case, and relatively low under the System Efficiency and Defendant Rehabilitation policies since both seek other forms of dispositions. Probably of all dispositions recorded, a purified dismissal rate (that which attributes responsibility to the proper participant in the system) is the most sensitive in evaluating prosecutor performance and the most accurate in measuring the effect of the charging policy.

Not only does a comparison among policies affirm that prosecutorial performance varies with regard to the policy of the office, but also that the policy must be determined before performance can be evaluated within an office. If one reads down any column in figure 5, one sees that the expected distribution of outcomes can be made consistent with the policy. For example, the Trial Sufficiency policy, that of insuring that the charge is correct and the case convictable, logically should result in a high rejection rate at intake, an indeterminate number of referrals to other criminal justice systems (for example, a municipal court), a minimizing of dismissals both at the probable cause hearing and at the trial level, a high frequency of bindovers since the goal is to try the case, a minimizing of plea bargains, high rates of pleas to the original charge, and correspondingly a maximizing of convictions.

Under the System Efficiency policy (the earliest and speediest disposition of cases) an evaluator would measure success or failure in terms of the number of persons diverted from the criminal justice system, the number of cases referred to other court systems, the number of cases disposed of by a plea bargain, and the number of cases bound over (the latter should be minimal). Collection of time-in-process statistics and the court phase at which disposition occurs is also essential to the evaluation of this policy.

The typology thus permits the examination of prosecutorial performance within a rational and logical system. Since the relative frequency and pattern of dispositions are expected to vary according to the policies being pursued by prosecutors, any evaluation should take this into account. While the pattern of dispositions is expected to vary across policies, once policy is taken into account the pattern of dispositions is expected to be reasonably regular and interpretable as prosecutors strive to maximize desirable outcomes or dispositions and minimize undesirable dispositions of their cases.

#### D. STRATEGIES TO IMPLEMENT POLICY

Just as we have seen that certain policies force certain outcomes, so too can we examine the use of certain strategies to implement policy. If strategies can be viewed as choices among options available to accomplish certain tasks, at least three are immediately identifiable. They are plea negotiation, discovery, and diversion. This section will explore how a prosecutor chooses and uses them to attain his policy objectives.

Plea Negotiation. One of the most important strategies used by prosecutors in disposing of cases is that of plea negotiation or plea bargaining. Its use or prohibition is so controversial and has generated such volatile discussion, that its role as a strategy to implement policy has been ignored. The abolition of plea bargaining by 1978 was incorporated into the National Advisory Commission on Criminal Justice Standards and Goals.<sup>5</sup> It generated so much discussion, controversy, and argument that this issue dominated all other criminal justice issues at the national conference called to promulgate these standards. Whether a plea to a reduced charge as a result of plea bargaining is an acceptable form of case disposition should not be argued in the abstract. Plea negotiation should be examined in light of its ability to implement the policy of the office. While it is recognized that not all plea negotiations result in a disposition called "plea to reduced charge," for simplicity, we have defined it as such here.

The use of plea negotiation is consistent with both the Legal Sufficiency and System Efficiency policies. With little preparation and review time, the assistants working under a Legal Sufficiency system will tend to accept pleas to reduced charges as a means of either correcting a charging mistake or minimizing the time required for more substantive case preparation. Under the early and speedy disposition policy of System Efficiency, it is essential that this be the primary means of disposing of cases because it is the fastest and least costly conclusion. If the Defendant Rehabilitation policy is in effect, it is difficult to predict whether plea bargaining will be used because it is not an expected outcome of the policy. Whether the more serious cases are allowed to plead to a reduced charge is both a function of court capacity as well as prosecutorial policy. Finally, under the Trial Sufficiency policy, it is entirely consistent that plea bargaining be minimized since the initial premise for accepting a case for prosecution is that it be properly charged, capable of being sustained in a trial, and expected to produce a conviction. Hence to permit plea bargaining would be to contradict the policy.

Discovery. The implementation of discovery is a procedure whereby the prosecutor opens his case file to the defense counsel, showing him the evidence and the strength of the case. Where discovery does not exist, the defense counsel is usually limited to that information which has been filed in the

court (usually the accusatory instrument) and that which he may glean from his client or from witnesses suggested by the client. Sometimes the defense counsel may not even see a copy of the arrest report until it is entered as evidence, nor may he know in advance the witnesses for the State.<sup>6</sup>

The most commonly expressed opposition to the use of discovery is based on the prosecutor's fear that by exposing his case to defense scrutiny, he may jeopardize his chances of winning. Indeed, this fear may be well justified if the review and charging process is nonexistent or weak. Whether this is a function of the resources and/or experience of the police department (which may produce less well-made cases), or a result of prosecutorial policy, would have to be determined before an evaluation of its use as a strategy could be ascertained.

Ideally, we would expect use of discovery to vary according to the policy being followed by the prosecutor. Under what we have termed the Legal Sufficiency model, discovery is not likely to be used, precisely because that policy tends to result in processing less well-made cases. On the other hand, it has been observed that use of discovery results in a high rate of disposition by pleas--either to the original or a reduced charge. Thus we would expect that the use of discovery as a strategy would frequent both the System Efficiency model and the Trial Sufficiency model. At this point, use of discovery under a policy of Defendant Rehabilitation is not predictable since the outcomes produced by discovery, namely pleas, are not necessarily relevant to the outcomes sought under this policy. It may be used, however, to assure the diversion of a defendant to a proper treatment program.

Diversion. Diversion is the third strategy available to the prosecutor in implementing his policy. Like plea negotiation, diversion has been previously identified as a disposition, but it is also a strategy that results in a disposition of the same name. Diversion, as a strategy, may be characterized by its referrals. A case may be diverted from the criminal justice system to alternative treatment programs, for example, the drug abuser to TASC, or the first offender to an employment program. Or a case may be diverted from one part of the criminal justice system to another. This latter type has been called "referred to" to distinguish it from the treatment function cited above.

To divert a case from the criminal justice system to treatment programs such as educational training or medical treatment programs, is a strategy that is consistent with all policies. Yet the reasons vary. Under a Legal Sufficiency model that deals mostly with misdemeanors, the universe of defendants diverted will more likely be first offenders and thus most eligible for non-criminal justice diversion. The System Efficiency model would make extensive use of all available diversion programs or facilities as a means of disposing of cases and reducing workload. The Defendant Rehabilitation policy views diversion as a treatment option. The Trial Sufficiency model does not necessarily need a diversionary exit; since its decisions are essentially binary in nature (either go or no go), the use of diversion is more a matter of individual preference.

When cases are referred to another criminal justice system (notably another lower court or court with concurrent jurisdiction), the reasons for this decision may be due to one or more of the following factors:

- (1) the police charges may not be accurate reflections of the prosecution charges (this is particularly true if police tend to overcharge);
- (2) referral to another court may be a technique to reduce workload;
- (3) because the lower court usually has jurisdiction over minor offenses, it may be used as part of a plea bargain; and
- (4) because lesser charges mean lesser sentences, it may also serve as an alternative form of diversion.

Figure 6 summarizes the strategies likely to be employed by an office to implement policy. Since the ultimate goal of the prosecutive function is case disposition, how the case is disposed of by using these strategies is reflexive of the policy of the office and the choices that are available and consistent with the policy.

FIGURE 6.--Expected use of strategies to implement policy

Policy	Strategies			
	Discovery	Plea negotiation	Diversion	
			Refer other CJS	Divert non-CJS
Legal sufficiency	Not predictable	Yes	Yes	Yes
System efficiency	Yes	Yes	Yes	Yes
Defendant rehabilitation	Yes; to expedite treatment	Not predictable	Yes	Yes
Trial sufficiency	Yes; to insure adjudication	No	Yes	Not predictable

#### E. RESOURCE ALLOCATION CONCEPTS

No matter what policy is being implemented, work has to be distributed in a rational manner if the desired outcomes are to be attained. Many resource allocation options which theoretically could be available to the prosecutor in actuality may be precluded by the external environment. For example, it would be difficult to organize an office around a trial team concept (wherein one or two assistants handle a case all the way from charging, through trial, to disposition), without a court processing system geared to support it. Successful trial teams flourish when cases are assigned by the clerk of the court to a specific judge or a specific courtroom or when the prosecutor controls the docket.

Most prosecutive resource allocation plans are primarily responses to the external environment. From an evaluation focus, one must account for resource allocation responses due to the characteristics of police, defense, and courts before a critique of any plan can be initiated. But critiques are possible. After the exogenous factors have been identified and their constraints determined, one should evaluate the resource allocation patterns with respect to their consistency with the policy and priorities of the prosecutor. Just as different policies require a rational implementation of strategies, so, too, do the goals of the office establish rational resource allocation patterns. This section briefly examines some of the ways resources can be distributed to insure consistency with policy. It focuses only on those areas under the prosecutor's control--charging, case assignment and trial preparation, and sentence recommendation--even though the importance of external factors is recognized. Figure 7 summarizes these distributions with regard to the above three areas and the policy model.

While the timing and completeness of police reporting is essential to the charging process, equally important are the qualifications of the person making the charging decision. Figure 7 shows that the experience level of the charging assistant may vary according to the policy of the office. For example, if the policy of the office is to examine cases only for legal sufficiency, as is the common practice in misdemeanor courts, then it is not necessary to use the most experienced assistant. Third-year law students are capable of examining a case for the elements, with minimal review of their decisions by junior assistants.

On the other hand, the System Efficiency policy requires that the charging decision be made with respect to a speedy and early disposition. Thus the charging assistant should have enough trial experience to know what is negotiable, enough system experience to know what can be diverted elsewhere (either to another court or other noncriminal justice programs), and what should be tried. There is little need for internal review of his charging decisions since the case is either sent elsewhere or the charge is expected to be changed. Satisfaction is guaranteed as long as speedy dispositions are occurring. Final evaluative review should focus on the disposition of those cases that were processed through various steps in the system.

Similarly, the Defendant Rehabilitation policy requires minimal review of the charging decision. Since the goal is to divert the treatable defendant from the system and to prosecute the recidivist who would not be eligible for diversion, the charging assistant must not only be trial experienced but trained in some type of social work. The delicate decisions of who to prosecute and who to divert offer potentially dangerous situations to an elected prosecutor. A defendant released to a community treatment program always represents a certain level of risk. The prosecutor must feel confident that his decisionmaker is competent, experienced, and ideologically attuned to his philosophy. Since the operators of the diversion programs can accept or reject the referral, the need for a review function in the prosecutor's office is minimized.

Finally, the Trial Sufficiency policy requires the utilization of the most experienced trial lawyers to make the charging decision. With this policy, once the decision is made to prosecute, the strategy is set; the case will go to trial, and a conviction is expected. Under minimal conditions, the charging

*FIGURE 7.--Expected patterns of resource allocations by type of policy*

Policy	Resource allocation needs			
	Charging		Case preparation for trial	Sentence recommendation
	Minimum qualifications for charging	Personnel needed to review charges	Trial experience necessary	Personnel needed for sentence recommendations
Legal sufficiency	Paralegal; 3rd-year law students; new assistants	Yes	Minimal	None
System efficiency	Trial and criminal justice system experience	Not necessary	Minimal	None, unless basis for plea bargain
Defendant rehabilitation	Trial and social work background	Not necessary	Moderate	Yes, to insure consistency with treatment
Trial sufficiency	Extensive trial experience	Yes	Extensive	Yes, to insure consistency with charge

decision will be made by an experienced assistant; under optimal conditions, it will also be subjected to another experienced review, thereby minimizing the chances of something being overlooked at the initial step.

Not only will the allocation of personnel to the charging and review process vary according to policy, but so too will the assignment of personnel for preparing and trying the case. The Legal Sufficiency and System Efficiency policies both move the cases after charging to assistants who first attempt to strike a bargain, and failing this, prepare the case or transfer this task to other assistants. Since the goal is to minimize trials, it is not necessary that the assistants have extensive trial experience. Under these conditions, it is interesting to question whether the character of the resources in the office (namely, young, inexperienced assistants with high employment turnover rates) creates a policy to accommodate to this environment or whether the policy creates and supports the existence of this environment.

The Defendant Rehabilitation policy can sustain a mixture in personnel. Inexperienced assistants may handle misdemeanor court and perhaps monitor the diversion programs, if programs are under prosecutor control. Since the cases accepted for prosecution should tend to deal with more serious, repeat offenders, assignment for case preparation and trial would be given to the more experienced assistants. A similar strategy would apply to the Trial Sufficiency policy.

With regard to the prosecutor's authority to make recommendations at sentencing, office resources would, at the most, be only minimally employed under the Legal Sufficiency and System Efficiency concepts since so few cases are expected to be disposed of by trial and since the majority will be disposed of by plea negotiations.<sup>7</sup> For the Defendant Rehabilitation and Trial Sufficiency policies, it would be expected that sentence recommendation would be used extensively since, for the former, it would insure the consistency of treatment with the needs of the defendant and, for the latter, insure the consistency of the charge with its expected punishment.

Although this discussion merely summarizes, in the briefest form, various patterns of work distribution, it does suggest the validity of evaluating resource allocation patterns in light of the policy and priorities of the office. Consistency with goals is obviously the critical factor. Just as it makes little sense to assign third-year law students at intake to determine whether a case can be bargained, so, too, it is just as unreasonable to use experienced lawyers to determine that the elements are present.

#### F. CONCLUSION

The critical impact of policy on dispositions, prosecutorial strategies, and resource allocation patterns cannot be discounted. If one is to evaluate the performance of the prosecutor and, even more broadly, the direction of our criminal justice activity, it is fairly obvious that the first task is to determine what the prosecutor is attempting to do; the second, to assess how well he is performing; the third, to see whether the community agrees with him. Ultimately, the public should be given a role in making this choice. Too often the questions are not so simply presented, if at all. At least one benefit of this study has been to support the thesis that the system, within limits, is rational or at least explainable.

## NOTES

1. The major findings of this project are available in published form. See Pre-Trial Screening in Perspective, Joan E. Jacoby, National Evaluation Program, Phase I Report. Series A, Number 2, U.S. Department of Justice, Law Enforcement Assistance Administration, Washington, D.C., January 1976. Other unpublished materials are available on a loan basis or microfiche.
2. Pre-Trial Screening in Perspective, op. cit., p. vii.
3. Argersinger v. Hamlin, 407 U.S. 25 (1972).
4. In some jurisdictions a nolle prosequi may be used in lieu of or in conjunction with disposition. For purposes of this discussion, this type of disposition will be called a dismissal.
5. National Advisory Commission on Criminal Justice Standards and Goals: Courts Standard 3.1, Abolition of Plea Negotiation, p. 46.
6. See Brian Grossman, The Prosecutor: An Inquiry into the Exercise of Discretion, (Toronto Press, 1969) for an excellent discussion on abolishing this practice and the merits of implementing discovery.
7. A major exception to this statement occurs when the prosecutor bargains for a sentence, not charge. Under these circumstances, he would make extensive use of this power.

THE POTENTIAL IMPACT OF INFORMATION SYSTEMS  
IN EVALUATING PROSECUTION AND COURT POLICY

Sidney H. Brounstein  
Institute for Law and Social Research

ABSTRACT

Among the problems inhibiting any improvement in criminal justice system performance has been the lack of reliable statistical information that would reveal patterns of criminal and related citizen behavior, describe how the present interagency system is operating in processing arrestees and their cases, and help detect and explain problems such as why so many arrests do not result in conviction. An information system that could generate such statistics periodically, including the ability to track changes over time, could help court and prosecution policymakers evaluate the impact of their policies and programs.

One example of such an information system, PROMIS (Prosecutor's Management Information System), designed to support daily operations management of the prosecutor and to provide necessary research and evaluation data, has been operating in the local prosecutor's office of the District of Columbia since 1971. The PROMIS Research Project is a demonstration of how such automated case management information systems can be utilized to provide timely information by which judicial administrators may evaluate the impact of their decisions.

The PROMIS Research Project has analyzed police operations from the court perspective, prosecution operations from the standpoint of their potential impact on crime control, and criminal justice system effectiveness from the victim's vantage point. These and several other studies, including recidivism and bail, will be described to illustrate their relevance in court and prosecution policy evaluation.

INTRODUCTION

The public has been increasingly expressing dissatisfaction with the effectiveness of the criminal justice system in controlling crime. Various court and prosecution practices have been singled out as contributing to this lack of control. These practices include what the public perceives as lenient bail and sentencing, excessive plea bargaining, and delay. Research findings suggest that the criminal justice system is a revolving door with respect to processing repeat offenders. In an era of rising costs and ever tightening governmental budgets, the option of adding resources no longer seems to be a viable alternative in attempting to improve effectiveness; for example, in conducting special investigations to enhance evidence, and in reducing court

congestion and delay. Like other criminal justice agencies, the prosecutor and court are being forced to find better ways of managing their available resources to achieve any goals that they may set.

The PROMIS Research Project is a demonstration of how automated case management information systems serving the daily operational needs of the prosecutor and court can be utilized to provide timely information by which criminal justice policymakers may evaluate the impact of their decisions. Presenting highlights of selected findings of the project, now in its third year, this paper will attempt to describe and illustrate how such information can be used for evaluative purposes.

#### DATA REQUIREMENTS FOR PROSECUTION AND COURTS EVALUATION

There are a wide variety of policies and programs that prosecutors and courts may be interested in evaluating, and it is unrealistic to attempt to anticipate all the data needed to support such evaluations. Nevertheless, a sufficient body of knowledge exists, from prior research and experience, to suggest the structure of a data base that would satisfy basic needs.<sup>1</sup> Before discussing the characteristics of that data base, it should be helpful to consider several examples of the types of policies and programs that prosecutors and courts may want to evaluate.

In implementing a program for targeting on career criminals, prosecutors may wish to measure differences in conviction, incarceration, and recidivism rates between those defendants processed by the career criminal unit and those not so processed. Prosecution management may wish to ascertain if office policies regarding priorities of prosecution are actually being pursued, and if so, whether the policy objectives are being met; for example, is vigorous prosecution of weapons possession cases having the desired effect of getting criminals off the streets?

Judicial administrators may wish to know whether criteria used in making bail and sentencing decisions are being applied fairly and uniformly by all judges. They may also wish to know whether certain criteria, for example, those designed to predict the defendant's likelihood of failure to appear in court, appear to be valid based on the accumulated statistical evidence.

Some typical problems which prosecutors and courts may encounter are court congestion and delay, victim and witness noncooperation, and excessive crime attributed to defendants on conditional release.

In attempting to alleviate these problems, administrators might introduce and attempt to evaluate a revised court scheduling method, a victim/witness identification and processing unit, or a revised pretrial hearing procedure designed to consider probation revocation as a means of detaining recidivists.

A flexible data base is needed to support such evaluations. Because of ethical and legal considerations, controlled experiments have little or no appropriateness in the court system. One type of evaluation design would provide for quasi-experiments, which though lacking the classical elements of experimental controls, do provide for statistical comparison groups. For example, the information system should have the capacity of generating

performance measures before and after introduction of a change in policy or procedure, while providing for the structure of comparison groups.

We can also gain insights by applying multivariate analysis techniques to nonexperimental data (i.e., data that accumulate in the normal course of operations).<sup>2</sup> Nonexperimental data are becoming more abundant in the criminal justice system, particularly due to the growth of automated management information systems.

The type of data base that appears to be needed in conducting prosecution and court evaluations is one that records the decisions and actions taken in each case, as it progresses through the court system over time. The data base should provide the date and outcome of each court event so that case-oriented data can be related to the dates of changes in organization, policies, and programs.

Each court case should be described in sufficient detail to support crime specific analyses, controlling for defendant characteristics. It should uniquely identify the police officers, attorneys, judges, types of victims, and witnesses involved at each stage of case processing. The data base should be flexible enough to link data from police, prosecutor, court, and correctional systems.

#### OPERATIONALLY BASED SYSTEMS VERSUS SPECIAL DATA COLLECTION

The usual course of action for evaluative researchers has been to devise a special data collection effort to develop the information they need for studying a particular problem. Such efforts are always limited by the time and cost constraints of the study. These data collection efforts rarely result in data that are as accurate as that collected as part of an ongoing operation. They tend to be limited with respect to the time frame that can be covered. Later, in the analysis, if any gaps occur or changes in the data collection design appear desirable, it is often impractical to modify the data base, and the research objectives need to be compromised.

Operationally based systems are those that support day-to-day operational applications. Any evaluation data provided by such systems are a byproduct of the operational functions. If an evaluative research program is conducted in close coordination with an operationally based system, it can obtain the following advantages over special or one-time data collection efforts:

- improved accuracy, since data being used in day-to-day operations are subjected to validity checking by automated systems, manual data control personnel, and the users themselves;
- familiarity with the data base on the part of the information systems personnel and the operational "users" so that nonobvious problems in the data and true meanings can be determined;
- ability to structure and maintain longitudinal data following individual persons or cases over time; and
- feasibility of modifying the information system to meet data needs newly determined by evaluators.

The data needs of the evaluation may not be totally met by any existing operationally based information system because those needs cannot be totally predefined in building the system. The nature of research and evaluation is such that new problems are always emerging for analysis, and little may be known about the factors that might explain the variation under scrutiny. Until the variables are properly defined, it may be necessary to collect a large amount of data about each individual "observation." Generally, a sample of "observations" is all that is necessary for evaluation purposes, rather than processing of all observations. For example, information on all criminal court cases is essential in an operational system for the criminal court, but only a representative sample of cases is needed to develop statistics for most evaluation purposes. These are some of the arguments often given to justify special data collection efforts by evaluators, but the aforementioned advantages of linking research/evaluation with ongoing operationally based systems are too significant to ignore. The remainder of this paper will attempt to demonstrate, by example, why this is so.

#### PROMIS DATA BASE

The major prosecution agency for Washington, D.C. is the Superior Court Division of the United States Attorney's Office for the District of Columbia. Although part of a Federal agency, the Superior Court Division closely resembles a local district attorney's or state attorney's office in that it has trial jurisdiction for common law, misdemeanor, and felony crimes. The District of Columbia Superior Court, in which this local prosecutive element functions, is the equivalent of a state court of general jurisdiction with trial responsibility for local civil and criminal litigation in Washington, D.C.

The data base that provides the unique source of information about the criminal court process has been and continues to be generated as a byproduct of a computer-based information system known as PROMIS (Prosecutor's Management Information System). PROMIS is used to support the daily operations of the Superior Court Division.

The PROMIS data base contains approximately 100,000 closed criminal cases including 170 different types of information on each case including:

- Information about the defendant. This includes name, alias, sex, race, date of birth, address, facts about prior arrests and convictions, employment status, and alcohol or drug abuse;
- Information about the crime. The date, crime and place of the crime, the number of persons involved in the crime, and information about the gravity of the crime in terms of amount and degree of personal injury, property damage or loss, and intimidation using the seriousness scale developed by the criminologists, Marvin Wolfgang and Thorsten Sellin;
- Information about the arrest. The date, time, and place of the arrest, the type of arrest, and the identity of the arresting officers;

- Information about criminal charges. The charges originally placed by the police against the arrestee, the charges actually filed in court against the defendant and the reasons for changes in the charges by the prosecutor, the penal statute for the charge, the FBI Uniform Crime Report Code for the charge, and the Project SEARCH Code for the charge;
- Information about court events. The dates of every court event in a case from arraignment through motion hearing, continuance hearing, and final disposition to sentencing, the names of the principals involved in each event including the defense and prosecution attorneys and judges, the outcomes of the events, and the reasons for the outcomes; and
- Information about victims and other witnesses. The names and addresses of all witnesses, the prosecutor's assessment of whether the witness is essential to the case or not, any indications of reluctance to testify on the part of the witness, and other witness characteristics, such as whether related to the victim or defendant.

In order to make longitudinal analysis work, PROMIS employs a unique offender identification number assigned by the police based on fingerprints and photographic checks. The unique ID is carried on every case record, following the offender through court processing in PROMIS, and even later in the corrections system. Whenever the offender is rearrested, that same ID is used. This enables tracking the offender over time, identifying and measuring recidivism, and analyzing relationships with recidivism. The unique offender identification numbers are used in merging historical files of closed cases. Similarly unique numbers are assigned to each court case and criminal event. The criminal event number is used to tie together codefendants in a particular case.

#### THE PROMIS RESEARCH PROJECT

Given the type of information system described in the previous section, what kinds of evaluative research studies can be conducted? The PROMIS Research Project provides several examples of such studies.

Like most administrators, criminal justice practitioners are forced to make decisions with incomplete information. There are always uncertainties; for example, anticipated witness cooperation in the prosecutor's decision to carry forward a case to trial, and the defendant's likelihood of flight in the bail decision. The challenge of evaluation is to provide the decision-makers with additional information that will improve the quality of their decisions with respect to the office's goals or performance standards.

Five examples will be drawn from the PROMIS Research Project to illustrate various types of evaluative studies. The first example concerns the evaluation of how well the criminal justice system is controlling specific crimes, and the development of diagnostic information useful in strategic planning for more effective crime control. The second example describes a longitudinal analysis of recidivism over a 5-year period that attempts to evaluate criminal justice actions with respect to recidivism. The next

example is an evaluation of prosecution policy before and after implementation of a career criminal program. The fourth example describes an evaluation of police effectiveness from the court perspective; that is, in producing arrests that lead to conviction. Finally, the fifth example describes an evaluation of the criteria used in making the bail decision.

#### EVALUATING EFFECTIVENESS IN CRIME CONTROL

Why do statistics that are valuable indicators of the performance of individual agencies often tend to obfuscate the combined, systemwide effectiveness of those same agencies? How might the collection of crime data be improved to enhance their utility to policymakers? Addressing these questions, in the PROMIS Research Project, various statistical adjustments were made so that court, prosecutory, police, and victimization data could be compared to obtain systemwide performance measures for various crimes and to analyze at what points--from victimization to conviction--criminal incidents dropped out of the criminal justice process.

Among the problems inhibiting any evaluation of criminal justice performance has been the lack of statistical performance measures that provide a systemwide perspective. It would be useful to be able to trace the incidence of crime from the original victimization through reporting to police, arrest, prosecution, and court disposition. LEAA's victimization surveys were designed to provide insights regarding the substantial number of crimes not reported to the police. Although it would be very desirable to be able to link victimizations with data from the Uniform Crime Reports (UCR), these two statistical systems are largely incompatible. In addition, while the statistics gathered by police, prosecutors, and courts may be useful for their individual purposes, they have tended to obscure attempts to view criminal justice performance from a systemwide perspective.

Research based on use of the Prosecutor's Management Information System (PROMIS) in Washington, D.C.<sup>3</sup> has attempted to explain the incompatibilities of the present criminal justice statistical systems, and to illustrate through certain manipulations made possible by the flexibility of PROMIS, the usefulness of statistical measures designed to provide a systemwide perspective. As indicated in exhibit 1, for example, the indictment-based conviction rate of 81 percent (police recorded a 72 percent clearance rate) for aggravated assault in Washington, D. C. during 1973 suggests a well functioning felony justice system from a prosecutor's perspective. However, when one bases the conviction rate on aggravated assault cases accepted for prosecution, the figure plunges to 37 percent. (Upon analysis, preindictment dismissals were a major reason for the drop; 65 percent of them were caused by witness problems. This area of performance--attrition between arrest and indictment--is one widely ignored by agency statistical systems, particularly by prosecutors in computing conviction rates.) As one proceeds down the chart, the criminal justice process appears to be increasingly ineffectual. From the victim's perspective, less than 7 percent of all aggravated assaults led to a conviction and less than 2 percent resulted in an incarceration. Clearly, the availability of such information on a crime-specific basis can yield valuable insights for policymakers and the public:

- Are citizens not reporting these crimes?

- Are the police not apprehending these criminals?
- Are prosecutors not obtaining convictions?
- Are judges too lenient in sentencing?

#### PATTERNS OF RECIDIVISM

PROMIS-assisted research has revealed that a small number of individuals were responsible for a significant portion of the prosecutor's and court's workload in the District of Columbia during the 56-month period ending September 1975: persons arrested at least four times accounted for only 7 percent of all arrestees but as many as 24 percent of all arrests; those prosecuted at least four times accounted for 6 percent of all persons prosecuted and 20 percent of all prosecutions; persons convicted at least three times constituted 5 percent of all those convicted and 15 percent of all convictions.

In the 56-month file, all of the defendants do not have an equal length of time in which to be rearrested. Specifically, persons first arrested on August 30, 1975 had only 1 day to be rearrested, while those first arrested on August 30, 1971 had 4 years. To overcome this problem, a sample panel of defendants was chosen to be studied in depth on a longitudinal basis. All panel defendants had been arrested at least once between November 1, 1972 and February 28, 1973. (A defendant's first arrest in this period is referred to as his or her "panel case.") The panel group selected for more intensive analysis constituted approximately 10 percent of the 45,575 defendants from the 56-month study.

The intent of the panel analysis is to determine ways of predicting the likelihood of recidivism based on data available during case processing and contained in PROMIS. If the defendant had five or more PROMIS arrests prior to the panel case, the probability of subsequent arrest approached certainty. The same was true for prosecuted cases. The probability of another conviction after a conviction in the panel case increased with the number of prior convictions. (However, not enough time was available for a defendant to have had many convictions.) Thus, the extensiveness of criminal history (regardless of whether expressed in terms of arrests, prosecutions, or convictions) seems to be a good predictor of future criminality.

Whether the seriousness of crimes committed increases or decreases over time was addressed by looking at arrests for felonies and misdemeanors. Defendants arrested for felonies in the panel case were more likely to be rearrested for felonies, and defendants arrested for misdemeanors in the panel case were more likely to be rearrested for misdemeanors. However, a considerable proportion of the defendants switched between felonies and misdemeanors.

Analyses were also performed by type of crime. The crimes were first classified into violent, property, and victimless offenses, and then into specific crime categories, such as burglary.

Defendants arrested for robbery, burglary, larceny, consensual sex offenses (mainly prostitution), and bail violations were found to involve a large percentage of recidivists.

EXHIBIT 1.--Aggravated assault: Perspectives on criminal justice performance  
(Washington, D.C.: 1973)

	Measures of performance	N	Rate	Comments
Prosecutor's perspective	A. <u>Conviction rates (closed cases)</u>			
	1. <u>Guilty pleas and findings (F)</u> <u>Indictments less dismissals (F)</u>	$\frac{232}{263}$	88%	39 of postindictment dismissals: witness problems.
	2. <u>Guilty pleas and findings (F)</u> <u>Indictments (F)</u>	$\frac{232}{286}$	81%	
Police perspective	3. <u>Guilty pleas and findings (MF)</u> <u>Cases accepted at screening (MF)</u>	$\frac{480}{1,284}$	37%	65% of preindictment dismissals: witness problems (no show, no prosecute).
	4. <u>Guilty pleas and findings (MF)</u> <u>Arrests (F)</u>	$\frac{480}{1,879}$	26%	30% of arrests not accepted for prosecution.
Victim's perspective	B. <u>Conviction rates (criminal incidents)</u>			
	5. <u>At least one adult guilty (MF)</u> <u>Reported offenses (F)</u>	$\frac{477}{3,591}$	13%	Arrest likely, if crime reported.
	6. <u>At least one adult guilty (MF)</u> <u>Victimizations (F)</u>	$\frac{477}{>6,906}$	47%	Victim reporting behavior poor.

(F) = Felonies      (M) = Misdemeanors

Data sources: All data from PROMIS (Prosecutor's Management Information System), except for (1) the denominator of measure 6, which reflects survey-based victimization data, and (2) the denominator of measure 5, which reflects Uniform Crime Reports data. The denominator of 6 is expressed as "greater than 6,906" because the victimization survey seems to underestimate aggravated assault incidents.

Defendants whose panel case was a violent crime (i.e., homicide, assault, sexual assault, or robbery) had the highest proportion of rearrests for violent crimes. However, the rearrests of these persons for violent crime were less than 50 percent of all their rearrests. There also appeared to be some specialization in property offenses. Defendants arrested for property crimes in their panel cases were more likely to be rearrested for property crimes than for any other offense. Overall, it was quite striking to observe the large amount of crime switching by defendants.

Also explored was the relationship between the final disposition in a case and the later rearrest rates for defendants. Those convicted in their panel case were more likely to be rearrested than those who were not convicted. (This may be an understated finding since periods of incarceration were not yet available for defendants in the panel.)

This longitudinal study will attempt to evaluate the impact of various court and prosecution actions, such as plea bargaining, bail, and sentencing, on patterns of recidivism.

#### PROSECUTION POLICY REGARDING RECIDIVISTS

Given the disproportionately large share of crime attributable to repeat offenders, prosecutors would seem justified in structuring their discretion so that an appropriate percentage of time and staff is focused on recidivists.

Such a decision is analogous to an investment whereby a person reduces current purchases of consumer items in favor of acquiring stock whose appreciation potential promises greater benefits over the long term. Similarly, the prosecutor may have to give up some convictions in the current period by diverting some office resources from readily convictable cases to relatively difficult ones involving repeat offenders in order to secure a greater reduction in future crime rates and future workloads. This reduction would result from the incarceration of those whose criminal histories reflect their relatively high potential for future criminality.

INSLAW analyzed about 6,000 felony cases in the Washington, D.C. PROMIS data base. The objective was to estimate the degree to which crime seriousness, defendant seriousness, and convictability actually influenced prosecution decisions for these calendar year 1973 cases. A carefully structured multivariate analysis revealed that convictability was the dominant influence, followed by crime seriousness. No evidence was found that defendant's criminal history independently influenced prosecution priorities.

This study was replicated using 1974 data with similar results. However, there are strong indications that a recently established "career criminal" program will change this situation. PROMIS is being used to monitor the effects of this program.

Launched in August 1976 by the Metropolitan Police Department and the U.S. Attorney's Office for the District of Columbia, a special unit (Operation Doorstop) is staffed by four experienced prosecutors and six police investigators, who focus exclusively on the serious, habitual criminal. According to the prosecutor's office, the objective of Operation Doorstop is to "stop the

revolving-door justice that permits repeat offenders to escape the punishment they deserve."<sup>4</sup>

Press accounts,<sup>5</sup> among other sources, have highlighted some of the policies governing Operation Doorstop:

- Repeat offenders are investigated, as appropriate, after arrest by the unit's prosecutors and police officers to build as strong a case as possible in order to minimize the chances of subsequent dismissals and to maximize the probability of conviction. (Similarly, the police department itself seems to be even more aware than in the past that arrests must be of a quality that they not only are accepted for prosecution but also can withstand closer scrutiny at later prosecutive stages.)
- A case involving a habitual criminal is not passed from one prosecuting attorney to another, assembly-line fashion; rather, it receives detailed attention from one prosecutor.
- Career criminals arrested while on probation or parole can expect the unit to try to hold them in jail while seeking to expedite the revocation of parole or probation. (The Superior Court has ruled that a probationer or parolee charged with a serious offense may be held without bond for 5 days, to permit sufficient time for a decision regarding revocation.) During its first 2 months, the unit identified 60 repeaters: 52 were jailed because of failure to make high bail or because of revocation of parole or probation.
- Case processing time is said to have been significantly reduced; for example, indictments have been returned within 8 days of arrest. And the court is reported to have agreed to attempt to schedule preliminary hearings in a manner that will conserve the unit members' time in court.
- Criteria for pretrial release are being tightened. To keep repeat offenders off the streets, preventive detention procedures are expected to be used more frequently than in the past. In a related matter, legislation is pending that would reduce the likelihood of pretrial release for those arrested while already on conditional release (bail, probation, parole).

A more thorough evaluation of the effectiveness of Operation Doorstop will be possible after sufficient time has elapsed to provide an adequate sample of closed court cases. Comparison groups can be structured using PROMIS data, contrasting the handling and outcomes of cases assigned to the Operation Doorstop unit versus those not assigned, and contrasting results before and after implementation of the program.

#### PRODUCING ARRESTS THAT LEAD TO CONVICTION

In adopting a broader systemwide perspective on crime control, it appeared to be desirable to explore why so many arrests do not end in conviction; that is, to evaluate police performance from the court's perspective. Since each PROMIS case record contains the badge number(s) of the arresting police

officer(s), it was feasible to link prosecution and court data with police personnel data records.

The data base was organized in two ways to analyze performance. First, a case-oriented analysis showed how various categories of arrests fared in court; secondly, an officer-based analysis, which grouped all cases of a particular officer together, revealed the performance patterns of various types of officers.

The case-oriented analysis revealed that conviction rates were quite sensitive to whether tangible evidence was recovered, the number of witnesses identified, whether the victim/defendant relationship was stranger-to-stranger, and delay between offense and arrest.

When tangible evidence, such as stolen property and weapons, is recovered by the police, the number of convictions per 100 arrests is 60 percent higher in nonviolent property offenses. When the police bring to the prosecutor arrests with more witnesses, the probability of conviction is also substantially higher, both for the violent and property crimes.

Related to the role of witnesses is the finding that a conviction was much more likely in an arrest in which the victim and arrestee did not know one another prior to the occurrence of the offense. This holds for robberies, other violent crimes, and nonviolent property offenses.

Another feature of the arrest influenced the likelihood that the arrests would result in conviction--the length of the delay between the time of the offense and the time of the arrest. This delay was found to be longest in robberies, with 55 percent of the arrests made more than 30 minutes after the offense. The conviction rate for robbery arrests, especially the stranger-to-stranger variety, declines steadily as the delay grows longer. In stranger-to-stranger robbery episodes, 40 percent of all persons arrested within 30 minutes of the offense were convicted; for the suspects apprehended between 30 minutes and 24 hours after the occurrence of the offense, the conviction rate was 32 percent; for arrests that followed the occurrence of the crime by at least 24 hours, the conviction rate was only 23 percent. This pattern was not apparent, however, in arrests for other offense categories.

To the extent that arrest promptness does increase the conviction rate, it appears to do so largely out of the enhanced ability of the police to recover tangible evidence when the delay is short. In stranger-to-stranger robbery episodes, recovery of evidence is more than twice as likely when the arrest is made within 30 minutes of the occurrence of the offense than when it is made at least 24 hours afterward. This pattern was similar for violent offenses other than robbery, and somewhat less extreme in the case of non-violent property offenses.

The ability of the police to recover tangible evidence, obtain witnesses, and arrest suspects promptly after the offenses occur is limited. Victims and other witnesses who notify the police of an offense--and not all witnesses do--often learn of the offense after some delay (especially in burglary and homicide cases); witnesses do not always notify the police promptly after becoming aware of the crime; tangible evidence and witnesses may often be unobtainable.

At the same time, the police who respond to the calls of victims and other witnesses may not be fully aware of the crucial importance to the success of the arrest in court of recovering physical evidence about the crime and the person who committed it--evidence such as stolen property, weapons, articles of clothing, samples of hair, and items marked with fingerprints. Further potential for reducing the enormous volume of arrests that fail to end in conviction is likely to lie in informing police officers of the importance of obtaining more than one good witness in serious crimes.

One way to induce arresting officers to obtain better evidence is to expand their perspective of their own performance beyond the number of arrests they make. Arresting officers are likely to bring better evidence to court when their incentive to increase the number of convictions they produce, particularly in cases involving serious offenders, exceeds their incentive to increase the number of arrests that they make.

The officer-oriented analysis was aimed at examining the characteristics of police officers that are associated with varying levels of performance; that is, arrests and arrests ending in conviction.

Especially striking is the fact that over half of the 4,347 MPD arrests made in 1974 that ended in conviction were made by 368 officers--a mere 15 percent of all the officers who made arrests, and 8 percent of the entire force. Eighty-four percent of all the convictions were produced by less than 1,000 officers (41 percent of all arresting officers and 22 percent of the force). And this phenomenon was not the result of a few officers making large numbers of arrests leading to convictions for victimless offenses. Over half of the 2,047 MPD arrests for felony offenses that led to conviction were made by a handful of 249 officers.

Nor do these prolific officers appear to have produced a large quantity of arrests at the expense of quality. The conviction rate for all the arrests made by the 368 officers who produced over half of all the MPD convictions was 36 percent--substantially higher than the conviction rate for the arrests made by all the other MPD officers who made arrests in 1974 (24 percent).<sup>6</sup>

The research showed empirically that officers with more experience were more likely to make arrests that ended in conviction than the more junior officers, which suggested that improved training may increase the quality of arrests. Another finding pertained to the residence of police officers.

In recent years, the combination of increased costs and a dwindling tax base have encouraged officials in a number of American cities, including Washington, D.C., and Detroit, to advocate that city employees be required to live within the city proper. Besides arguing that such a requirement would result in increased tax revenues, officials in both cities have argued that there would be an additional benefit in enforcing such a requirement for police officers: their residence within the city would result in better protection for citizens.

Our research relating officer characteristics to performance indicates that an officer who resides outside of the District of Columbia tends to make more arrests and is more likely to have arrests result in convictions than the

officers who reside within the District, even after accounting for differences in experience and other factors among these officers.

This finding that police officers residing inside of the District were no more productive in terms of both the quantity and quality of arrests suggests that local government officials should carefully weigh the trade-offs involved when considering residency requirements for police officers.

## THE BAIL DECISION

The judge's decision to grant a defendant release pending trial is usually based on two criteria: the likelihood of appearing in court and the likelihood of committing new crimes. Statistical evidence may help the judge make better predictions.

This paper reports the preliminary findings of an analysis of the determinants of the pretrial release decisions, ability to post money bond, and failure to appear.

Under the terms of the D.C. bail law, release on bail is to be based solely on the defendant's likelihood of appearance in court. A separate provision in the bail statute, called "preventive detention," permits denial of pretrial release to certain classes of defendants, based on their dangerousness to the community and strength of evidence in their cases. Designed to protect the community from new crimes rather than to assure court appearances, this provision rarely has been invoked. In requesting high money bond in cases that might otherwise qualify for preventive detention, prosecutors have argued that such defendants are likely to flee rather than face a severe sentence. This research has been examining whether judges appear to have been accepting the argument and reports on models being developed to test the validity of that argument.

The purpose of the research has been to provide policy- and decisionmakers with empirical information on the functioning of the bail system. Judges tend to be reluctant to use statistics, since each case must be decided ultimately on its individual merits. The thesis of this research is that there is no conflict with that philosophy. The judge at the bail hearing is faced with making a prediction of the defendant's future behavior (failure to appear in court or rearrest while on conditional release). He has to consider the conditions of the individual case before him. But if in addition he is provided with statistical information to help him make a better prediction, the interests of justice and the community will be better served.

Based on the preliminary analysis of the determinants of the decision to set financial conditions of release rather than release on personal recognizance, it appears that the judge is most influenced by the prior criminal record of the defendant, and to a somewhat lesser extent, by the seriousness of the current offense and the strength of the evidence. Older defendants and those with other cases pending against them were also more likely to have financial conditions of bail set. When money conditions were set, the amount was influenced most by the seriousness of the offense and next by the criminal history of the defendant. Therefore, there is some indication that judges have been accepting the prosecutor's rationale that defendants who have serious criminal records facing serious charges with strong evidence are more

likely to flee. On the other hand, it may indicate that high money bond is being used in lieu of preventive detention to protect the community from dangerous releasees.

Preliminary findings of an analysis of the determinants of failure to appear suggest that cases involving more delay, stronger evidence, and a stranger-to-stranger relationship between the defendant and the victim were more likely to result in failure to appear in court. No support was found for the hypothesis that defendants with more serious prior criminal records and charged with more serious crimes are more likely to fail to appear.

Using the PROMIS data base as a sample frame, those cases in which financial conditions of release were set were sampled to undertake a special data collection effort aimed at ascertaining which defendants were able to post bond and secure actual release. Such information is not collected routinely, but is necessary in estimating failure-to-appear rates. An analysis was made of the likelihood of securing release by posting money bond. Defendants released on cash bond rather than surety, those who had their bond conditions changed most often (suggesting the importance of a good defense counsel), those with lower amounts of bond (not surprisingly), and those with less serious criminal records were more likely to be released. A simultaneous equations model is being developed to sort out the effects of the determinants of the pretrial release decision on failure to appear.

Although the results reported in this paper are preliminary, they suggest that statistical information can be used to evaluate the criteria used for the pretrial release decision, and can help judges better understand how the system is working so that they can make more informed bail decisions.

#### COMPARATIVE EVALUATIONS AMONG JURISDICTIONS

About 50 state and local jurisdictions are in various stages of implementing PROMIS. These include some of the largest prosecution entities in the United States, such as Los Angeles County District Attorney's Office, the New York County (Manhattan) District Attorney's Office, and the Wayne County (Detroit) Prosecuting Attorney's Office.

More recently, PROMIS has begun to be accepted at an increasingly rapid rate by the courts themselves. For example, Milwaukee County, Wisconsin, has made some modifications to PROMIS, renamed it "Justis," and implemented it as its trial court information system. The unified court systems of the States of Florida and Rhode Island have decided to adopt PROMIS as the nucleus of their state judicial information systems, serving both trial court needs and state-level judicial planning and management requirements regarding criminal cases.

Since a substantial proportion of the data elements collected in each PROMIS jurisdiction will be common to all PROMIS jurisdictions, comparative evaluations will be feasible. In addition, a set of standardized report generation programs are supplied to each jurisdiction to facilitate such comparisons.<sup>7</sup>

Various aspects of the court's problems and activities are reported in the media. Judicial administrators tend to be apprehensive about looking bad

in the newspapers, and may resist rather than cooperate with outside evaluators. Certain performance measures, such as average case processing times, incarceration rates, and disposition rates may be viewed as too high or low, depending on the eyes of the beholder. Experience has shown that administrators tend to be less defensive when they can be shown that their court's performance is not so extreme when compared to that of other courts.

Eventually the wealth of data on varying criminal procedures in the PROMIS jurisdictions can provide a natural quasi-experimental setting for cross-jurisdictional evaluative studies. Each jurisdiction tends to do certain things differently. Hence, comparable data describing their case processing practices and performance may suggest more effective practices.

## CONCLUSION

This paper represents an attempt to show the power of soundly designed operationally based management information systems, at the local level, to support prosecution and court evaluations. Through the use of several examples drawn from the PROMIS Research Project, it illustrated the use of such an information system for evaluating:

- The effectiveness of the criminal justice system in controlling crime, utilizing flexible counting mechanisms to relate UCR and victimization data through PROMIS;
- The impact of court and prosecutory actions on patterns of recidivism, utilizing a 5-year longitudinal data file;
- The conduct of prosecutive policy with respect to cases involving recidivists, utilizing techniques feasible only with a large number of observations and variables;
- Police performance from the court's perspective, utilizing unique identifiers to link agency files; and
- The criteria for the bail decision, utilizing the data base to select a random sample of cases for further data collection to supplement the routinely collected data.

The evolution of a unique multijurisdictional data base for conducting comparative evaluations among many jurisdictions has also been discussed. Since these PROMIS jurisdictions will have many data elements in common and utilize standard reporting packages, this data base has the potential for providing a natural quasi-experimental environment for cross-jurisdictional and time series analyses of various court and prosecution policies and procedures.

## NOTES

1. See, for example, Technical Report No. 12 SJIS, State Judicial Information System Final Report, (Phase I) Search Group, Inc., Sacramento, California, June 1975.
2. For example, some of these techniques are described in "Statistical Techniques and Their Limitations in the Analysis of Criminal Justice Data," Brian Forst, Quantitative Tools for Criminal Justice Planning, Law Enforcement Assistance Administration, U.S. Department of Justice, Washington, D.C., 1975, pp. 113-121.
3. "Expanding the Perspectives of Crime Data: Performance Implications for Policymakers," Institute for Law and Social Research, Washington, D.C., 1977.
4. Earl J. Silbert, U.S. Attorney for the District of Columbia, as quoted in the Washington Post, August 12, 1976.
5. Washington Evening Star, August 11, 1976; Washington Post, August 12, 26, and October 19, 1976.
6. This is not to imply that officers who make large numbers of arrests generally have higher than average conviction rates. Indeed, we find a conviction rate of 33 percent in the arrests made by the officers who made only one arrest and a rate of only 27 percent for the 3,081 arrests made by the 828 MPD officers who made at least 20 arrests each.
7. See Special Report #1, the PROMIS Management Report package for PROMIS, and Special Report #2, the Generalized Inquiry Package for PROMIS, Institute for Law and Social Research, Washington, D.C., 1977.

SECTION III  
EVALUATION IN CORRECTIONAL PROGRAMS

AN EXAMINATION OF INTENSIVE SUPERVISION  
AS A TREATMENT STRATEGY FOR PROBATIONERS\*

Joseph H. Sasfy  
The MITRE Corporation

INTRODUCTION

The interest in intensive supervision and the increase in the implementation of intensive supervision projects are responses to the general failure of traditional probation and parole projects in which caseload sizes of 75 to 125 were typical. Intensive supervision essentially means smaller worker/client ratios and, as such, represents a general treatment approach rather than a specific treatment modality or program. There is a widespread belief among practitioners that reductions in caseload size free the worker to provide treatments and services in a more intensive, responsive, and individualized manner. One of the recommendations of the President's Commission on Law Enforcement and the Administration of Justice explicitly reflected the assumption that reduced caseloads translate into more effective supervision and, thus, reductions in recidivism:

"All jurisdictions should examine their need for probation and parole officers on the basis of an average of 35 offenders per officer, and make an immediate start toward recruiting additional officers on the basis of that examination."<sup>1</sup>

The correctional caseload research which has followed the growing interest in intensive supervision has clearly failed to answer what seemed to be the essential question--do reduced caseloads reduce recidivism? If there has been an evolution in this research, it has been from a search for the proper (or most effective) caseload size to a concern for the more immediate variables affecting the quality of the supervision process. Thus, the general question of the effectiveness of reduced caseloads has been supplanted by more specific questions relating to those variables central to the treatment process.

It should be noted (somewhat ironically) that the growth of interest in intensive supervision and individualized treatment as specific alternatives to traditional probation and parole practices has been followed by a broad disillusionment in the rehabilitation ideal in general. The demise of the rehabilitation model is at least partly the result of evaluative reviews of the

---

\*This research was funded as part of the national-level evaluation of the High Impact Anti-Crime Program under Contract J-LEAA-028-75 for the Law Enforcement Assistance Administration.

literature on correctional treatment programs. For instance, Martinson,<sup>2</sup> in an exhaustive review of over 200 studies embracing educational, vocational, intensive supervision, counselling, and other treatment approaches, concluded that there were no techniques which clearly demonstrated the ability to reduce recidivism. There is little doubt that the correctional philosophy that has emerged in the last few years (with its emphasis on the concepts of deterrence, punishment, and offender rights) would present a strong contrast to the generally optimistic philosophy of rehabilitation in a community context which formed the basis for many of the correctional projects of the last 20 years.

This study was designed to address the general question of the effectiveness of intensive supervision by aggregating data across a number of projects featuring reduce caseloads. In addition, more specific questions related to offender characteristics and the prediction of recidivism were addressed.

#### METHODOLOGY

Five projects, serving mostly juvenile probationers and featuring caseloads of no more than 35 per worker, were the subject of the research. The projects were New Pride (Denver), Providence Educational Center (Providence), Probation Aftercare #6 (Los Angeles), Essex County Probation Department's Specialized Caseload Project (Newark), and Case Management Corrections Services (Portland). Data, collected on 388 clients in these projects, were used to describe the following categories of variables:

- Client-descriptive--includes age, ethnicity, grade level, living situation, educational lag, and three attitudinal variables.
- Preprogram criminal offense--includes length of preprogram criminal career, and number, frequency, and severity of preproject criminal offenses.
- Baseline criminal offense--includes frequency and severity of offenses for 1-year period prior to project entry.
- Project criminal offense--includes length of stay in project and number, frequency, and severity of project offenses.

These variables were employed to conduct a number of analyses including:

- Comparison of the frequency and severity of offenses from the baseline to program period; and
- Prediction of the frequency of offenses in the program period (recidivism) using client-descriptive variables, preprogram and baseline criminal offense variables.

The dependent variable selected for the assumptions research is a recidivism measure (frequency of offenses during project) based on arrest data for the client samples. There are a number of limitations related to arrest data. Undoubtedly, the decision to arrest a person and to charge him with a particular offense is often dependent on a number of arbitrary and/or chance factors. Nevertheless, arrests are probably the best available measures of criminal behavior other than self-reports of victimization. If other definitions of

criminal behavior that occur further on in legal proceedings (referrals, for instance) were used, there would be a serious underreporting of criminal activities. For example, in Denver, approximately two-thirds of all juvenile arrests are not referred to court. By using arrest data, however, we will be assuming a high correlation between arrests and the actual commission of offenses and also between the offense charged and the suspected illegal behavior. Thus, in this research, offenses are considered synonymous with arrests.

## RESULTS

Frequency Comparisons. The frequency comparisons (see table 1) represent the most direct assessment of project-level effectiveness in the assumptions research. The t-values for the baseline to project period comparisons of frequency of offenses indicate that all five projects achieved significant reductions. Expressed in terms of percentage change, the reductions in frequency ranged from 28.5 percent for New Pride to 62.0 percent for Providence Center. The effectiveness of these intensive supervision projects taken as a group is reflected in the reduction in frequency from a baseline average of about two offenses to a project average of one offense; in other words, a 50 percent reduction in the offenses was achieved by these projects. Additionally, almost half of all clients (46.1 percent) committed no offenses during the project period which averaged over 15 months ( $\bar{x} = 15.2$ ). The percentage not arrested in project ranged from 61.8 percent for Case Management to 19.6 percent for New Pride.

Although all five projects demonstrated significant reductions in recidivism, the design employed (baseline to project comparisons) has serious limitations. Without control groups, it is impossible to determine if the reductions achieved by clients under intensive supervision were greater than those achieved by clients receiving traditional probation services or no services at all. Some basis for comparison is provided in the study of juvenile recidivism in Denver conducted by Carr and Molof.<sup>3</sup> This study provides 1-year rearrest data for a cohort of 2,203 juveniles arrested during a 1-year baseline period. Because the Denver study showed prior number of arrests to be a significant predictor of recidivism and the present research findings are similar (see below), a comparison was made between project frequency for clients in this study and the 1-year followup frequency for a sample of Denver juveniles matched on prior number of arrests.

Table 2 presents a comparison of frequencies and percentages not arrested for juveniles from the Denver study and from this research matched on number of prior arrests. As can be seen, clients under intensive supervision had lower frequencies than Denver juveniles at every level of prior offenses. On the other hand, a higher percentage of clients under intensive supervision were arrested during the comparison periods. This finding is partially accounted for by the fact that the service period averaged 16.6 months for these clients while the Denver data are based on 12 months. The intensive supervision clients had a lower frequency of rearrest than the Denver cohort even though the Denver cohort had a proportionally greater number of individuals with only one prior offense. This is reflected in the mean number of prior offenses for the two groups; intensive supervision clients average around six prior offenses, while the Denver juveniles averaged three.

TABLE 1.--Baseline to project period comparisons of frequency of offenses

Variables Projects	Baseline				Project				Percent reduction in frequency	t	Percentage not arrested in project
	Number of clients	Number of offenses	Frequency	S.D.	Number of offenses	Time in project months	Frequency*	S.D.			
New Pride (Denver)	51	116	2.28	1.54	79	11.42	1.63	1.71	28.5%	2.49 <sup>1</sup>	19.6%
Providence Center (St. Louis)	70	105	1.50	.72	48	14.41	.57	1.00	62.0%	6.59 <sup>2</sup>	51.4%
Essex County (Newark)	69	103	1.49	1.07	92	19.46	.82	.98	45.0%	4.53 <sup>2</sup>	44.9%
Case Management (Portland)	144	282	1.96	1.35	130	12.76	.85	1.53	56.6%	7.87 <sup>2</sup>	61.8%
Aftercare #6 (L.A.)	54	187	3.46	1.85	137	20.87	1.46	1.21	57.8%	6.73 <sup>2</sup>	24.1%
Total	388	793	2.04	1.46	486	15.20	.99	1.39	51.5%	12.28 <sup>2</sup>	46.1%

158

1.  $p < .05$

2.  $p < .01$

\*Project frequencies are adjusted to a 12-month base.

TABLE 2.--Comparison of frequency of offenses for two groups matched on prior offenses

Variables Number of prior offenses <sup>1</sup>	Denver recidivism study			Intensive supervision research (total sample) <sup>2</sup>		
	n	1-year followup frequency	Percent not arrested	n	Project frequency	Percent not arrested <sup>3</sup>
1	921	.59	70.2%	39	.52	66.7%
2	348	1.17	48.3%	32	.88	31.2%
3	213	1.40	37.6%	27	.77	48.1%
4	130	2.05	23.1%	24	11.23	33.3%
5	119	2.55	17.6%	24	.74	45.8%
6	74	2.38	13.5%	12	1.63	25.0%
7+	398	2.53	18.6%	86	1.73	23.3%
Total	2,203	1.36	46.8%	244	1.05	37.3%

1. The average prior arrests for Denver juveniles was 2.98, while it was 5.98 for assumptions re-search clients.

2. No data on prior arrests were available for Case Management.

3. These arrests percentages are based on a 16.6-month average service period.

It must be noted, however, that this comparison data should be interpreted cautiously, since no attempt was made to match clients on any variable other than prior offenses. Additionally, the experience of the Denver cohort during the followup year is unknown; undoubtedly, many were under traditional probation supervision, some were receiving no formal supervision, and some may even have been under intensive supervision.

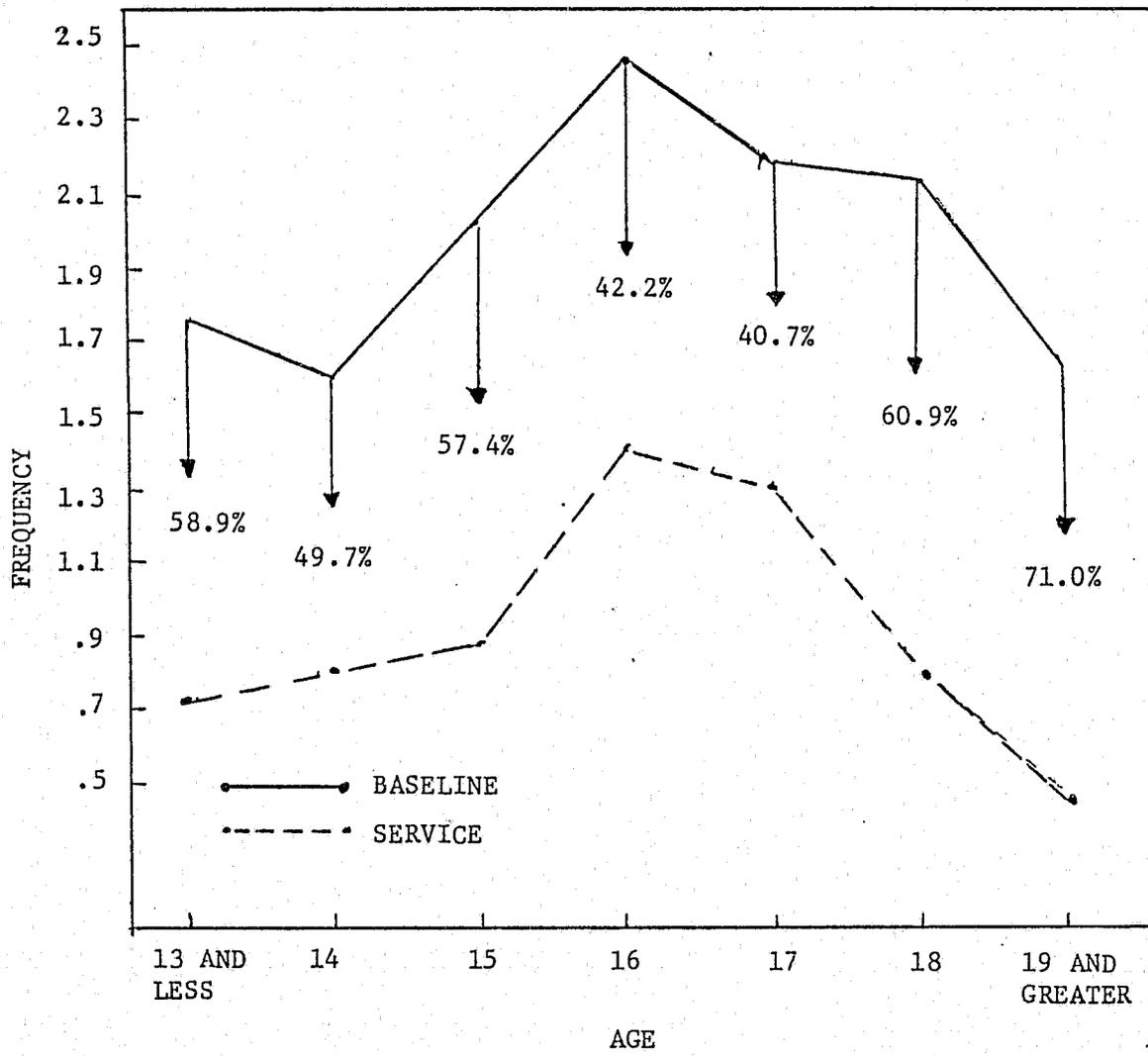
Severity Comparisons. In order to compare the qualitative nature or severity of offenses in the baseline and project periods, all offenses in the baseline and project periods were assigned severity ratings based on Warren's California Severity of Offense Scale.<sup>4</sup> Results indicate that there was little change from baseline to project periods in terms of the average severity of an offense. The overall reduction in severity was only 5.4% (from  $\bar{x} = 4.99$  to  $\bar{x} = 4.72$ ) for all projects with Case Management achieving the largest reduction (11.8%). Thus, it can be concluded that, although these projects achieved a 50 percent reduction in the quantity of offenses, the seriousness of these offenses was unchanged.

Client-Descriptive Variables as Predictors. Bivariate correlations were calculated between client-descriptive variables and frequency of offenses during the baseline and project periods.

Of greatest interest were the significant negative relationships between age and baseline and/or project frequency for New Pride, L.A. Aftercare, and Essex County. Despite these negative correlations, the correlations for all clients between age and baseline frequency ( $r = -.03$ ) and project frequency ( $r = -.06$ ) are almost zero. This pattern of results suggests that these significant negative correlations are derived from only part of the total range of age for all clients, and that at another part of this range, a positive correlation exists (thus, creating zero correlations over the total range). To explore this possibility further, baseline and project frequencies were plotted for each age group summing over the individual projects (see figure 1). As figure 1 clearly shows, there appears to be a positive relationship between age and baseline and project frequency for the lower half of the age group, and a negative relationship for the upper half. Because New Pride, Essex County, and L.A. Aftercare have the oldest clients, most of their clients fall in the upper half and, thus, the negative correlations. Figure 1 also indicates the percentage reduction in frequency for each of the age groups. These reductions were sizable for all groups, ranging from 40.7 percent for 17-year-olds to 71.0 percent for those 19 or older. Sixteen-year-olds were the worst recidivists (that is, had the highest frequency of offenses) in project, but also had the highest baseline frequency.

Criminal Offense Variables as Predictors. The Denver recidivism study has underlined the importance of prior contacts with the criminal justice system as a predictor of recidivism. In that study, prior referrals ( $r = .38$ ) and prior arrests ( $r = .36$ ) proved to be the best predictors of number of rearrests in a 1-year period. Similar to the results of the Denver study, the best predictors of project frequency of offense (recidivism) were baseline frequency ( $r = .30$ ,  $p < .01$ ), and preproject number of offenses ( $r = .29$ ,  $p < .01$ ) and preproject number of offenses ( $r = .26$ ,  $p < .01$ ). Thus, baseline frequency proved to be the best single predictor of recidivism in this study.

FIGURE 1.--Project frequency, baseline frequency, and percentage reduction in frequency by age



Both this study and the Denver study found criminal offense variables far more useful as predictors of recidivism than client-descriptive variables like age and ethnicity. It is possible, however, that these client-descriptive variables could significantly interact with other variables in the prediction of recidivism yet still prove (as they did) nonpredictive in a simple linear regression. To explore the possibility that baseline frequency could interact with one or more of the client-descriptive variables in the prediction of project frequency, clients were separated into two groups--those with one or two baseline offenses and those with three or more. Each of these two groups was then further divided into two groups in terms of four client-descriptive variables: age--under 16 or 16 and over; ethnicity--white or nonwhite; living situation--living with both parents or some other living situation; and educational lag--1 year or less, or more than 1 year.

Since only one of these variables, age, interacted with baseline frequency in the prediction of project frequency, this is the only interaction discussed. Table 3 reports the project frequencies for four groups of clients classified on the basis of baseline frequency and age. Statistical tests of these project frequencies (recidivism) show that, among clients with one or two baseline offenses, clients 16 and over recidivated more ( $t = 3.21, p < .01$ ). Among clients with three or more baseline offenses, clients under 16 recidivated more ( $t = 1.19, n.s.$ ). The difference in t-values for the comparison of the two baseline frequency groups for clients under 16 ( $t = 3.41, p < .01$  and 16 and over ( $t = 1.81, n.s.$ ) indicates that baseline frequency was a better predictor of project frequency for the younger clients. The two groups which appear to have benefited least from intensive supervision, then, were older juveniles with one or two baseline offenses and young juveniles with three or more baseline offenses. The latter group recidivated most, averaging almost two offenses per service year.

TABLE 3.--Project frequency by baseline frequency  
for two age groups

Number of offenses baseline	Variable		Number of clients	Baseline frequency	Project frequency	S.D.	Percent reduction in frequency
	Age group						
1, 2	Under 16		120	1.28	.54	.93	57.8%
	16 and over		160	1.35	.99	1.37	26.7%
3+	Under 16		31	3.97	1.87	2.11	52.9%
	16 and over		70	4.10	1.37	1.47	66.6%

## SUMMARY OF RESULTS AND CONCLUSION

The major findings of the present assumptions research are that all projects achieved significant reductions in recidivism in terms of a baseline to project period comparison. The percentage reductions in recidivism for the individual projects ranged from 28.4 percent to 61.9 percent. The overall percentage reduction was around 50 percent, reflecting an overall change in frequency from two offenses in the baseline year to one offense in the project year. The quality or seriousness of offenses, however, showed no change from baseline to project. In order to provide some basis for a comparison of the recidivism of juveniles under intensive supervision with those receiving traditional probation supervision or none at all, 1-year offense frequencies were presented for clients in this study and juvenile offenders in Denver matched on number of prior offenses. The data indicated that clients in intensive supervision recidivated less at every level of prior offenses.

Of the client-descriptive variables, age proved most useful in the prediction of various criminal offense measures. In a number of projects, age showed negative relationships with baseline and/or project frequency. Although there were no significant overall correlations between age and baseline and project frequency, the breakdown of baseline and project frequency by age level suggested curvilinear relationships. Both baseline and project frequency increased until age 16 and decreased thereafter. Sixteen-year-olds were the worst recidivists, but also had the highest baseline frequency. Sizable percentage reductions in recidivism (40.7 percent to 71.0 percent) were found at each age level.

Overall, the criminal offense predictors proved more useful than the client-descriptive predictors. The best predictors of recidivism were baseline frequency and preproject number of offenses. In terms of project frequency, the possibility of interactions between baseline frequency and client-descriptive variables was explored. The interaction between age and baseline frequency revealed that, for clients with one or two baseline offenses, older clients recidivated more. For clients with three or more baseline offenses, younger clients recidivated more. The most serious recidivist was clearly the young juvenile with numerous previous offenses.

Based on the analyses performed here, it would appear that intensive supervision, as a general strategy, was effective in terms of reducing recidivism. Both the baseline to project comparison and the comparisons with matched groups of juvenile offenders from Denver point to the effectiveness of intensive supervision. In addition to the significant reduction in recidivism for each of the five projects, reductions were found at every level of preproject number of offenses and baseline frequency. Also, the analysis of interactions between various client-descriptive variables and baseline frequency indicated that reductions in recidivism occurred for all levels of age group, ethnicity, educational lag, and living situation. In short, intensive supervision seemed to be beneficial for clients with different criminal backgrounds and different demographic characteristics, although some groups appeared to benefit more than others.

Given the results of this research, it would seem more work in the correctional caseload area is warranted. Other research in this area has generally been equivocal, but many studies have found that something seems to be working, at least for some types of offenders and under certain conditions. To isolate the effective factors so that policy and programmatic recommendations can be

made at a broad level, therefore, will require a renewed and redoubled commitment to the research spectrum of programmatic evaluation. Given the program costs involved in the abandonment of community corrections in favor of institutionalization, it is likely that, sooner or later, such a commitment will have to be subscribed.

## NOTES

1. The President's Commission on Law Enforcement and Administration of Justice. Task Force Report: Corrections. Washington, D.C.: U.S. Government Printing Office, 1967.
2. Martinson, Robert. "What Works"--Questions and Answers about Prison Reform," The Public Interest, No. 35, Spring, 1974, pp. 22-54.
3. Carr, J.D. and M.J. Molof. Juvenile Recidivism. Denver Anticrime Council, July 1974.
4. Warren, Marguerite Q. "Community Treatment Project: History and Prospects," California Youth Authority Annual Report, 1967.

OBJECTIVE SELECTION TECHNIQUES AS A LOGICAL PREREQUISITE,  
TO THE EVALUATION OF WORK RELEASE PROGRAMS<sup>1</sup>

Douglas E. Scoven  
Duane E. Brookhart  
J.B. Ruark  
Virginia Department of Corrections

In recent years, considerable emphasis has been placed upon the development of treatment-oriented programs which utilize community-based resources. It appears that a limited number of correctional administrators have come to agree that mere incarceration has been ineffective in achieving the goal of reducing deviant behavior. Crime is now viewed as a social problem, and the community-based approach dictates that the most effective method of dealing with that problem is within its social context (Cooper, 1973; Keller and Alper, 1970).

Ironically, recent changes in correctional philosophies and approaches have not necessarily been accompanied by the development of more reliable measures and innovative methods for evaluating community-based programs. Recidivism continues to be used as a major indicator of program success or failure, and it is often assumed that success implies the achievement of rehabilitative goals (cf. Martinson, 1974). This emphasis on postoutcome results undermines the importance of assuring that the program under investigation is both properly structured and operationally effective (Cavior and Cohen, 1975). It is argued that the common finding of no differences between treatment and comparison groups may be a function of the manner in which the program itself is administered. If this is in fact the case, then the treatment effectiveness of the program cannot be accurately assessed (Cressey, 1965). Thus, it is possible that some researchers are attempting to evaluate criminal justice programs at a level which is greater than is currently possible or justifiable. This implies the need for a careful reassessment of priorities.

Community-based programs cannot be all things to all offenders. Logically, there will be some proportion of every institutional population that could be classified into a poor risk category as far as treatment goes. If it is true that some offenders are better risks for treatment than others, then it must follow that a specific correctional program which proves to be effective with one type of inmate may not be equally effective in all other cases (Vold, 1958). To the extent that a program has been carefully planned and is oriented toward the attainment of specific treatment goals, the challenge faced by the evaluator is to identify objective screening techniques which will aid in the selection of those offenders who are most likely to benefit from program participation. As Vold (1958: 302) has stated, "an essential requisite to successful rehabilitation practice is clearly a system of intelligent and careful selection of good risks."

There will be a select group of criminal justice planners and researchers who strongly object to the line of reasoning discussed above. Martinson (1976: 181) contends that "to ask which methods work best for which types of offenders and under what conditions . . . is to impose the narrowest of questions on the search for knowledge." In the same article, Martinson (1976: 189) asserts that "it is nonsense to set out to answer this question if you are not sure that you have found a method that really works in the first place." The fact that Martinson offers few suggestions which would enable one to discover a method that really "works" is not surprising.

Each criminal justice evaluator, planner, and program administrator will hopefully assess both arguments and conclude for himself which of the two appears to be the better approach. In doing so, the following question must be considered: Is the placement of an inappropriate or unsuitable individual into a correctional program fair to the offender, the program, and society? Based upon the practical example discussed below, we maintain that it is not.

#### WORK RELEASE: AN ILLUSTRATION

To date, there appears to be little evidence that participation on work release adequately prepares or reintegrates the offender into society. Most investigations pertaining to the effectiveness of work release have utilized quasi-experimental designs in which the treatment and control groups are compared by recommitment rates. Studies conducted by Witte (1975), the Oklahoma Department of Corrections (1973), the District of Columbia Department of Corrections (1971), and Johnson (1969) indicate no substantial differences in rates of recommitment between groups of offenders who had and had not participated in work release.

The consistent lack of variation in postoutcome performance between the treatment and control groups was given a dual interpretation: either work release was of little benefit in reintegrating the offender into society, or unsuitable individuals were being placed on the program. Upon examining the in-program violation rate of the work release program operated by the Virginia Department of Corrections, cumulative terminations due to rule infractions, new convictions, and escapes were found to be as high as 30 percent per month. This clearly suggested the need for an investigation of the internal operation of the Virginia program.

Because the assignment of offenders to most correctional programs is based solely upon human discretion, unpredictability is substantially increased, as the grounds for making decisions are usually vague and often subjective. It was suggested that

". . . two types of errors substantially account for failures in the Work Release Program: either the participant was incorrectly selected for program participation; or, the participant was correctly selected, but he was placed in an inappropriately designed program. . . . Proper selection is of integral importance, as program structure changes remain essentially speculative until it can be assured that the appropriate individuals are entering the program (Brookhart, Scoven, and Ruark, 1975)."

Clearly, an extensive evaluation of the Virginia work release program would have been premature until its operational efficiency, as measured by the in-program violation rate, could be improved. Extensive research was undertaken in an attempt to identify those factors which could discriminate between potentially successful and unsuccessful program participants. The final instrument was pretested, and was formally implemented in February 1976. An assessment of this selection technique during its first year of operation, together with its implications for use in other correctional programs, is discussed below.<sup>2</sup>

## METHODOLOGY

Sample. Data on 21 preprogram characteristics were collected on a random sample of 250 adult male felons who had completed participation in the work release program of the Virginia Department of Corrections between 1973 and 1975. Of these participants, 119 were considered to have successfully completed the program, as evidenced by termination of work release due to parole or expiration of sentence. The remaining 131 participants were considered to have been unsuccessful on work release, with program participation terminated due to a violation of institutional or program regulations, commission of a new offense, or escape. All data were obtained from the participant's Central Criminal Record as maintained by the Division of Adult Services. The final selection instrument was developed utilizing information gathered for this sample.

Statistics Utilized. The data were analyzed through the use of linear discriminant analysis. This statistical technique assumes that the dependent variable can be partitioned into two or more mutually exclusive and exhaustive categories. In addition, discriminant analysis permits the researcher to compute individual probabilities of group membership, a feature which was considered to be of vital importance to the practical application of the instrument. A stepwise solution based upon the within-groups covariance matrix and employing a selection criterion equivalent to the inclusion of variables yielding the largest overall multivariate F ratio was used in the analysis of data.

Program Monitoring. Subsequent to February 1, 1976, data were gathered on all offenders as they became eligible to participate in the work release program. This information is applied to the discriminant function, thus yielding a standard index of participant suitability. The records of those who meet the suitability requirements are forwarded to the appropriate administrative officials for final approval. Those who do not meet the suitability requirements are reconsidered on a periodic basis. Thus, no inmate is placed on work release merely on the basis of a statistic. The suitability index is viewed as an aid to administrative decisionmaking rather than a replacement for it.

## FINDINGS

While discriminant analysis provides the option of utilizing all variables toward the explanation of predictable criterion variance, it is generally desirable to identify a subset of characteristics which provide for optimal separation between groups. Through a careful analysis of both the standardized discriminant function coefficients and the intercorrelation matrix, it was found that such a subset of variables did appear to exist. These characteristics were used in all further analyses. Of the original 21 variables, 8 were found to significantly discriminate between those who had been successful and unsuccessful while on the work release program. No two variables exhibited substantial

covariance, and all multivariate F ratios were significant at or beyond the .01 level of confidence. The variables entering the final analysis, together with their associated standardized discriminant function coefficients, are presented in table 1.

TABLE 1.--Variables entering the discriminant analysis and their relative contributions\*

Variable	Relative contribution
Level of emotional maturity Higher ratings on the Initial Classification Psychological Report indicate successful work release participation.	.6148
Relationship to parole eligibility The further past the date of parole eligibility, the greater the likelihood of success.	-.5595
Institutional adjustment reports Greater numbers of institutional adjustment reports tend to be indicative of unsuccessful program participation.	-.3833
Time remaining to discharge The greater the number of months to discharge, the greater the likelihood of program success.	.3545
Number of total convictions The greater the number of total convictions, the less the probability of successful program termination.	-.1702
Number of prior felony convictions The greater the number of prior felony convictions, the less the probability of successful participation.	-.1580
Type of offense Those convicted of less serious offenses tend to succeed in work release.	-.1496
Occupation The lower the skill-level prior to incarceration, the greater the chance of work release success.	-.1327

\*Standardized discriminant function coefficients. Multivariate F ratios and Wilk's Lambda for all coefficients significant beyond the  $\alpha$  .01 level.

As indicated in table 1, level of emotional maturity, as measured by the participant's rating on the Initial Classification Psychological Report, evidenced the greatest contribution toward discriminating between successful and

unsuccessful work releases. Relationship to parole eligibility entered on the second step of the analysis with a relative contribution of .5595: those who enter work release after their parole eligibility date tend to be successful on the program. Number of institutional adjustment reports and time remaining to discharge made significant contributions to the present analysis. The greater the number of prior adjustment reports, the fewer the number of months from date of assignment to date of discharge, the greater the probability of unsuccessful work release performance. Both number of total convictions and number of prior felony convictions are inversely related to successful program participation. Finally, the more serious the offense, and the greater the occupational skill-level prior to current conviction, the more likely is the participant to be unsuccessful on work release.

Followup Analysis. As mentioned above, 3 of every 10 offenders assigned to the work release program prior to February 1976 evidenced unsuccessful in-program terminations. If the instrument described in this report does in fact aid in the selection of suitable program participants, then the violation rate should have reduced significantly. Between February 1, 1976 and January 7, 1977, 104 offenders were assigned to and subsequently terminated from the Virginia Department of Corrections' work release program. In each case, the predictive instrument was used as an aid in the decisionmaking process.

As reported in table 2, the in-program violation rate has been reduced from 30 percent to approximately 11 percent within a 1-year period. Of the 104 offenders who were assigned to the program between February 1, 1976 and January 7, 1977, 84 were paroled (80.76%), and 8 (7.69%) were discharged. Twelve participants were unsuccessfully discharged from the program: eight (7.69%) were removed due to violation of program standards, two (1.93%) were terminated as a result of institutional charges or rule infractions, and two participants (1.93%) escaped. These data support the conclusion that the development and implementation of objective selection criteria have been successful in improving the operational effectiveness of the Virginia work release program.

## DISCUSSION

Work release is viewed as a program of special integrity, as it serves as a preliminary step in bridging the gap between the extremes of confinement within total institutions and the responsibilities which the offender must face when released into the free community. Because of its linkage with incarceration, the program is typically vulnerable to community sentiment. It is therefore understandable for correctional administrators to be cautious in assigning offenders to the program. However, the conclusion that emerges from the present research is that appropriate caution cannot be exercised until salient factors which distinguish between successful and unsuccessful participants are identified.

The predictive strategy discussed above has important practical implications for both program administration and evaluation. From an administrative standpoint, the entire program can be more closely monitored and controlled. Personnel who are currently charged with supervising work release participants could now focus upon those offenders who most need their services. This can be of vital importance in establishing relationships which foster program success. In addition, variables such as time to discharge give the program a unique dynamic quality. The suitability index which is computed does not remain static;

TABLE 2.--Predictive power of the selection instrument  
as indicated by followup data

Category	Number	Percent
Successful in-program terminations:		
Paroled	84	80.76
Discharged	8	7.69
Unsuccessful in-program terminations:		
Administrative removals	8	7.69
Institutional charges	2	1.93
Escapes	<u>2</u>	<u>1.93</u>
Totals*	104	100.00

\*February 1, 1976 - January 7, 1977: Cumulative success rate = 88.45%  
Cumulative violation rate = 11.55%

rather, it continues to reflect the eligible offender's status so that he can be reconsidered at the appropriate time. Finally, acceptable probabilities of success or failure can be lowered or raised so that the decisions can be made at any level of risk which circumstances dictate.

The impact of objective selection techniques upon the work of the criminal justice evaluator is substantial. To the extent that the results of objective selection techniques reflect the operational effectiveness of the program, the evaluator is alleviated of the responsibility for making assumptions which are frequently unjustifiable. The quality of any evaluation is enhanced through the quantification of in-program information. Knowledge of the internal effectiveness of the program should allow for more accurate interpretations of its treatment effectiveness. Because the salient factors which affect successful program participation are known, more relevant comparison groups can be selected.

The continued growth of community-based programs in general may well be dependent upon the extent to which the appropriate individuals are selected for participation. Objective screening criteria, when used in conjunction with subjective assessments, can be used to reassure the public that only those offenders who prove to be good risks are being considered for program participation. The public is entitled to such reassurance, and it is suggested that the development of reliable screening criteria is worthy of careful consideration among correctional administrators.

## NOTES

1. The authors wish to thank the Virginia Department of Corrections for the assistance and support of this research. Although the Department facilitated this project, the statements and conclusions contained in this report are those of the authors, and should not necessarily be interpreted as reflecting the position or carrying the endorsement of the Virginia Department of Corrections, the Division of Probation and Parole Services, or the Division of Adult Services.
2. It should be noted that an instrument similar to the one discussed herein was developed for use by the District of Columbia Department of Corrections. Unfortunately, this predictive tool was not implemented.

## BIBLIOGRAPHY

- Brookhart, Duane E., D. E. Scoven, and J. B. Ruark.  
"An Identification of Factors Discriminating Between Work Release Successes and Failures." Unpublished Report: Virginia Department of Corrections. 1975.
- Cavior, Helene E. and S. H. Cohen.  
"Evaluative research: Perspectives from a corrections setting." *Criminal Justice and Behavior* 2 (September): 237-257. 1975.
- Cooper, H. H. A.  
"Toward a rational doctrine of rehabilitation." *Crime and Delinquency* 19 (April): 228-240. 1973.
- Cressey, Donald H.  
"Prison Organizations." Pp. 1023-1070 in James G. March (ed.), *Handbook of Organizations*. Chicago: Rand McNally. 1965.
- District of Columbia Department of Corrections.  
"Performance of Institutional Releases by Release Category." Washington: Research Unit, District of Columbia Department of Corrections. 1971.
- Johnson, Elmer H.  
Work Release: Factors in Selection and Results. Carbondale: Southern Illinois University. 1969.
- Keller, Oliver J. and B. S. Alper.  
Halfway Houses: Community-Centered Correction and Treatment. Lexington: D.C. Heath. 1970.
- Martinson, Robert.  
"California research at the crossroads." *Crime and Delinquency* 22 (April): 180-191. 1976.  
  
"What works?--Questions and answers about prison reform." *Public Interest* (Spring): 22-54. 1974.
- Oklahoma Department of Corrections.  
"Oklahoma City Community Treatment Center: Research Report #4." Oklahoma City: Planning and Research Division, Oklahoma Department of Corrections. 1973.
- Vold, George B.  
*Theoretical Criminology*. New York: Oxford University Press. 1958.
- Witte, Ann Dryden.  
"Work Release in North Carolina: An Evaluation of its Effects After Release From Incarceration." Technical Report: North Carolina Department of Corrections. 1975.

THE ACTUALIZATION AND IMPACT OF TEAM CLASSIFICATION  
IN STATE CORRECTIONAL INSTITUTIONS\*

John R. Hepburn, Ph.D.  
Department of Sociology  
University of Missouri-St. Louis

Celesta A. Taylor, Director  
Criminal Justice Planning  
Missouri Division of Corrections

The concept of a treatment team working together in routine inmate classification decisions is a rather new and untested correctional innovation (Loveland, 1960; Burns, 1975). The treatment team concept, as an alternative to the conventional classification team, was first implemented in 1961 at the Federal Reformatory in El Reno, Oklahoma, modified somewhat when introduced at the Federal Youth Center in Englewood, Colorado, and refined further before its inauguration at the Federal Correctional Institution at Tallahassee, Florida (Hagan and Campbell, 1968). The major characteristics of Team Classification as implemented are (1) caseloads structured around housing units, (2) teams of a small number of specialists, and (3) assignment to the team of the full range of case management responsibility.

Team Classification in Missouri is designed to exemplify a team approach to the decentralization of decisionmaking. The team consists of the inmate and those institutional staff who theoretically are most closely and directly involved with the inmate and are most aware of his assets and needs: his caseworker, his parole officer, his immediate work supervisor, and the correctional officer supervisor or counselor assigned to the inmate's housing unit. The combined effort of these staff and the inmate is to formulate and implement a personalized plan for each inmate around which will be made decisions pertaining to such matters as cell, educational and job program assignments, furlough and work release, and disciplinary action.

The inmate's presence on the team as an active participant with equal voice and vote in deciding the immediate issue that concerns him is designed to facilitate an understanding of the different positions of staff and inmate. Furthermore, the inclusion of the inmate is designed to foster a greater understanding of the rationale underlying the decision, resulting in greater inmate cooperation with and conformity to that decision. The observations and combined

---

\*This project was jointly funded by the Missouri Division of Corrections and Grant #75-ED-07-0003 from the Law Enforcement Assistance Administration, U.S. Department of Justice, under the Omnibus Crime Control and Safe Streets Act of 1968, as amended. The findings and conclusions stated in this document do not necessarily represent the official position of the Law Enforcement Assistance Administration, the U.S. Department of Justice, or the Missouri Division of Corrections.

efforts of Team Classification are, by design, to culminate in an accurate, holistic view of the inmate's adjustment to the institution and provide a basis for more appropriate and less fragmented decisionmaking. Similarly, while each staff member makes a unique contribution to the decisionmaking process, it is felt that the inclusion of both treatment and custody staff will create fewer misunderstandings, will develop a spirit of fair play, and will develop mutual goals with other staff members of the team. Through the inclusion of all relevant persons in the decisionmaking process, therefore, team classification, as a conceptual model is designed to facilitate greater understanding among staff and between staff and inmates and thereby achieve more favorable attitudes between and among these two segments of the institution.

The team's central overall task is that of developing a personalized plan for each inmate. Relatedly, the team is to continually reevaluate the inmate's behaviors and attitudes in light of present staff observations. The team is to accomplish its tasks through collective decisionmaking with all five team members having an equal vote. By formulating an overall plan which theoretically reflects the needs and desires of the inmate and which takes into account the institution's ability to meet these defined needs, fair and appropriate decisions are to result. In summary, the team is structured and charged to provide for "continuity of responsibility, which reduces program fragmentation, increases the likelihood of meeting inmates' individual needs, and fosters improved interpersonal relationships among staff and between inmates and staff" (Directive on Team Classification, 1975:2).

#### RESEARCH OBJECTIVES<sup>1</sup>

This evaluation of Team Classification focuses on program actualization and program impact. The assessment of actualization will provide information pertaining to the extent to which Team Classification is being implemented according to its original conceptualization, for there is no available evidence that the existing operation of Team Classification adequately reflects the concept as previously delineated. Among the criteria for effective implementation of Team Classification are the following: (1) development of personalized plan which meets the needs of the inmate appearing before the team; (2) granting an equal role to participating members; (3) decisions based on inmates' present behavior rather than past behavior; (4) open discussions in team meetings; (5) familiarity of team members with inmate team members; (6) team responsiveness to inmate needs. Data pertaining to these criteria, in addition to data regarding the frequency of involvement, will permit an assessment of the degree to which Team Classification has been actualized in the Missouri Correctional System.

The evaluation of program impact is directed to (1) staff and inmate attitudes toward Team Classification as a decisionmaking process and (2) staff and inmate attitudes toward general conditions in the facility. Those attitudes toward Team Classification which are operationalized include: fairness of Team Classification, staff's perceived impact of Team Classification on their job, staff support of Team Classification, impact of Team Classification on inmates, and perceived effect of Team Classification on the relations among staff and between staff and inmates. The following general conditions in the facility are utilized to ascertain the broader impact of Team Classification among staff: attitude toward inmates, attitude toward work assignment, staff attitude toward other staff, treatment vs. punishment orientation, and role conflict. The

general attitudes among inmates are attitude toward living assignment, attitude toward program assignment, attitude toward staff, and alienation.

The underlying model guiding this research (see Suchman, 1969) assumes that greater actualization of Team Classification will lead to more positive attitudes toward Team Classification by both inmates and staff, which, in turn, will result in more positive staff and inmate attitudes toward their general conditions within the facility. More specifically, the evaluation is directed toward the assertions made by the proponents of Team Classification, assertions which explicitly state that a highly actualized implementation of Team Classification will gain the cooperation of inmates, promote inmate-staff and staff-staff understanding, and provide a fair vehicle of decisionmaking. It is reasoned, further, that if the implementation of Team Classification can realize these objectives it may also, as a result, have a larger impact within the facility, such as the reduction of role conflict among staff and alienation among inmates.

#### RESEARCH METHODS

Although an experimental research design would be desirable, Team Classification had been an operational program in the institutions for at least 2 years prior to the funding of this research project. This evaluation, consequently, had to be based on a cross sectional design (Suchman, 1970) without the more desirable pretest measure and control group.<sup>2</sup> A 90-item, anonymously completed survey was verbally administered to a random sample of inmates in groups of approximately 25 persons at each institution. A combined total of 1,297 inmates were surveyed from all of the correctional institutions. Table 1 reports the inmate population and the size of the inmate sample; each sample size allows for 98 percent precision in 99 out of 100 samples. A self-administered instrument was distributed in each correctional institution to all those staff who directly interact with inmates on a routine basis. Table 1 also reports the number of anonymous questionnaires returned by staff at each facility. The data analysis is thus performed on the responses to forced-choice items by 27.6 percent of Missouri's adult inmate population and 67.6 percent of all treatment and custody staff working within Missouri's adult correctional institutions.

Each variable is measured by means of a summated, Likert scale consisting of items to which the respondent indicates the extent of his agreement or disagreement on a 5-point continuum from Strongly Agree to Strongly Disagree. In each case the items of the scale are significantly ( $p < .001$ ) correlated with one another and with the total scale score value. Measures of split-half reliability indicate greater reliability with those scales measuring staff attitudes, but all scales meet the minimum reliability requirements. Actualization is operationalized as the summated score of the individual scale scores of its components: familiarity, personalized plan, equal role, present behavior, open discussions, and, for inmates, responsiveness. Perhaps a more reliable assessment of Team Classification's actualization could be obtained by means of systematic observation of Team Classification meetings; this would require a large sample of lengthy observations within each facility, however, which was precluded by restraints on time and resources. The absence of an objective criterion of the degree of actualization has mandated a reliance on the views of respondents, acknowledging that respondents may perceive a higher or lower level of actualization than actually (if measured in other ways) exists. From a

TABLE 1.--Inmate and staff samples by institution

Institution	Inmates			Staff		
	Total inmate population	Sample size	Percent of population	Number surveyed	Number of respondents	Percent return
Missouri State Penitentiary	2,341	356	15.2	325	198	60.9
Missouri Training Center for Men	1,025	199	19.4	178	120	67.4
Church Farm	386	169	47.8	34	27	79.4
Renz Farm	169	95	56.2	32	25	78.1
Fordland Honor Camp	153	101	66.0	34	27	79.4
State Correctional Center for Women	111	90	81.1	38	30	78.9
Missouri Intermediate Reformatory	<u>516</u>	<u>287</u>	<u>55.6</u>	<u>110</u>	<u>81</u>	<u>67.4</u>
Total	4,701	1,297	27.6	751	508	67.6

social-psychological view, however, it is the perception of actualization, rather than the actual degree of actualization, that will have whatever attitude impact is associated with Team Classification.

#### DATA ANALYSIS

Implementation and Actualization. Data presented in table 2 indicate the scale score range, scale score mean, and the proportion of the maximum value approached by the mean for each actualization component for staff and inmates who indicated membership in Team Classification. Since there is no a priori absolute value which, when reached, indicates that effective implementation is occurring or against which the observed level of actualization can be measured, another operational device was created. The scale score range for each actualization component indicates the minimum and maximum scores possible, the scale score mean indicates the average score, and the proportion reveals how close the observed score is to the maximum obtainable score (a proportion of 1.0 would indicate that the maximum and the average are the same). Familiarity among members and the examination of present rather than past behaviors are the most highly actualized among staff since in both cases the proportion indicates that the degree of actualization is well over half the (measurable) potential. According to inmate responses, responsiveness to inmate needs and familiarity of members are most highly actualized. Given the admittedly crude measurement device of a questionnaire, the results suggest that the components of Team Classification have been implemented to various degrees, and all of them appear to have been implemented to some degree.

Table 2 also reports the Pearsonian correlation coefficients of each component scale score to the total actualization scale score. The coefficients among component scales are all positive and range from .45 to .92 for staff and .16 to .55 among inmates, indicating a high level of concomitant variation among the components. This suggests that (1) those teams in which one or two components are actualized are also those teams which are likely to actualize all components and (2) the various components do not appear to be working at cross-purposes.

According to the survey results, 48.7 percent of the staff and 79.8 percent of the inmates have been involved with Team Classification. Difference of means t-tests which examine the difference in staff and inmate attitudes when distinctions are made on Team Classification membership, frequency of participation or length of participation are reported in table 3. Inmate attitude toward work assignment is significantly more favorable among members than nonmembers. Yet, inmate alienation is higher and attitude toward staff is less positive among those with a longer rather than shorter length of participation in the program. Similarly, inmates with a higher frequency of participation have a less favorable attitude toward staff than those with a lower frequency of participation. Furthermore, this apparent negative effect is not limited to inmates. Data for staff reveal that team members have a significantly lower attitude toward both inmates and their work assignment, a significantly more favorable attitude toward punishment of inmates, and a greater degree of role conflict than nonmembers. Less favorable attitudes toward their work assignments and toward inmates are also more characteristic of those staff who have participated for longer rather than shorter durations and frequently rather than infrequently.

TABLE 2.--Actualization of team classification: an analysis of its components

Team Classification components	Scale score range	Scale score mean	Mean to maximum value proportion	Pearsonian correlation to actualization scale score
<u>Staff:</u>				
a. familiarity	1- 5	3.426	.69	.73
b. personalized plan	4-20	12.402	.62	.87
c. equal role	6-30	14.653	.49	.91
d. present behavior	3-15	11.502	.77	.77
e. open discussions	9-45	23.246	.52	.92
<u>Inmates:</u>				
a. familiarity	1- 5	2.89	.58	.60
b. personalized plan	4-20	9.45	.47	.64
c. equal role	9-45	18.44	.41	.74
d. present behavior	2-10	4.28	.43	.61
e. open discussions	4-20	10.32	.52	.67
f. responsiveness	2-10	7.04	.70	.54

TABLE 3.--Mean differences on inmate and staff general attitudes between levels of Team Classification membership, participation and actualization

	Team membership			Length of participation			Frequency of participation <sup>2</sup>			Actualization level <sup>2</sup>		
	Nonmember	Member	T-Value	1 year or less	More than 1 year	T-Value	Low	High	T-Value	Low	High	T-Value
<b>Inmates:</b>												
a. Living assignment	6.05	6.11	.28	6.13	6.04	.64	6.08	6.15	.48	5.90	6.31	3.29*
b. Work assignment	2.97	3.62	2.06*	3.65	3.39	1.38	3.67	3.36	1.54	3.08	4.08	5.77*
c. Staff	11.17	10.71	1.08	11.11	10.10	3.80*	10.91	10.30	2.23*	9.09	12.53	15.36*
d. Alienation	14.99	15.42	1.20	15.23	15.63	1.86*	15.38	15.40	.11	16.28	14.40	9.88*
<b>Staff:</b>												
a. Inmates	6.40	5.53	5.14*	5.75	5.28	1.80*	5.83	5.38	1.54	5.29	5.82	2.02*
b. Work assignment	11.94	10.49	4.42*	11.18	10.03	2.17*	11.38	10.31	1.59	9.83	11.89	4.04*
c. Other staff	13.17	12.89	.80	13.75	12.77	1.63	13.92	13.04	1.27	12.44	14.68	3.77*
d. Punishment	10.69	11.86	2.44*	12.50	11.93	1.08	12.63	12.05	.96	12.71	11.22	2.84*
e. Role conflict	17.14	18.50	2.89*	17.99	19.05	1.42	17.85	18.73	1.03	19.60	16.60	4.10*

1. Since staff had a higher average frequency of participation than did inmates, "low" inmate frequency is four or less meetings and "high" inmate frequency is five or more meetings while "low" and "high" frequency for staff are nine or less meetings and ten or more meetings, respectively.

2. Actualization level is dichotomized into "low" and "high" around the mean actualization level score. The mean for inmates is 15.55 and the mean for staff is 15.15.

\*Indicates the T-Value is significant at or greater than  $p < .05$ .

There is little evidence to suggest that involvement per se improves the general attitudes of inmates or officers. Indeed, it appears that, where any such attitudinal differences do exist, involvement is more likely to result in less favorable rather than more favorable attitudes. Table 3 also reports the mean differences among staff and inmate attitudes by the level of the respondent's perception of the actualization of Team Classification. The results clearly and consistently illustrate a positive effect of higher actualization on the attitudes of both inmates and staff. It is not the existence of the program, then, nor the involvement of personnel in that program that appear to affect attitudes. What appears to be the crucial factor in its effect upon these attitudes is the degree to which those who are involved feel that the program is being implemented in a credible and efficacious manner consistent with the program's design.

Assessing the Impact of Team Classification. Does the degree to which Team Classification is felt to be actualized affect the attitudes of inmates and staff toward Team Classification, which in turn affects certain attitudes about the facility? The matrix of Pearsonian correlation coefficients presented in table 4 demonstrates the relationship among the measures of actualization, attitudes toward Team Classification, and general attitudes for those staff who are or have been members.<sup>3</sup> With regard to the relationship between actualization and staff attitudes toward Team Classification, two important facts emerge. First, actualization is not significantly related to: improved staff understanding, improved inmate understanding, or positive job impact. Second, actualization is negatively associated with staff attitude pertaining to Team Classification's impact on inmates, effect on inmate-staff understanding, and support by inmates. When combined with the positive associations between actualization and fairness, staff support and warden support, it appears that a highly actualized Team Classification procedure has no relation to staff understanding of inmates or other staff but is nonetheless favorably received by the staff, who themselves feel it is not favorably received by the inmates.

Table 4 also reports the coefficients of correlation among staff actualization scores and staff attitude toward inmates, work assignment, other staff, punishment, and role conflict. These data suggest that actualization is positively associated with staff attitude toward inmates, work assignment, and other staff, and negatively associated with staff attitude toward punishment and role conflict. The relationships existent between the various measures of attitudes toward Team Classification and general attitudes are also presented in table 4. Attitude toward inmates is positively associated with inmate support for Team Classification, improved inmate-staff understanding, and impact on inmates. Staff attitudes toward work assignment, other staff, punishment, and role conflict, with few exceptions, are significantly associated with each of the measures of attitudes toward Team Classification. It was assumed that a highly actualized Team Classification procedure would reduce role conflict among staff and result in a less punitive attitude toward inmates; the negative relationships reported in table 4 provide support for that assumption.

Table 5 presents the matrix of coefficients between inmate measures of actualization, attitudes toward Team Classification, and general attitudes. The relationships between Team Classification actualization and all inmate attitudes toward Team Classification are at a level of statistical significance and, with one exception, indicate a strong, positive relationship between actualization

TABLE 4.--Correlation matrix of actualization,  
attitude toward Team Classification and general attitudes, for staff

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<u>Actualization of program:</u>															
1. Actualization score	1.00	.15	.11	-.16	.17	.16	-.04	-.04	-.21	-.22	.17	.19	.19	-.09	-.15
<u>Team Classification assessment:</u>															
2. Fairness		1.00	.90	.40	.29	.26	.87	.84	.49	.48	.08	.18	.08	-.20	-.19
3. Positive job impact			1.00	.29	.20	.19	.93	.93	.40	.44	.09	.17	-.02	-.27	-.19
4. Inmate support				1.00	.43	.23	.27	.20	.52	.52	.19	.27	.33	-.20	-.30
5. Staff support					1.00	.43	.15	.07	.55	.51	.04	.27	.44	-.04	-.36
6. Warden support						1.00	.13	.12	.32	.35	.05	.17	.22	-.08	-.24
7. Improve staff understanding							1.00	.96	.34	.31	.09	.12	-.01	-.19	-.12
8. Improve inmate understanding								1.00	.22	.24	.07	.10	-.10	-.19	-.06
9. Improve inmate/staff									1.00	.69	.17	.34	.40	-.21	-.42
10. Impact on inmates										1.00	.25	.33	.28	-.34	-.47
<u>General attitudes:</u>															
11. Inmates											1.00	.37	.15	-.22	-.13
12. Work assignment												1.00	.47	-.16	-.45
13. Other staff													1.00	.01	-.51
14. Punishment														1.00	.28
15. Role conflict															1.00

TABLE 5.--Correlation matrix of actualization  
attitude toward Team Classification and general attitudes, for inmates

	1	2	3	4	5	6	7	8	9	10	11	12	13
<u>Actualization of program:</u>													
1. Actualization score	1.00	.55	.51	.15	.18	.57	.46	-.24	.50	.09	.14	.46	-.31
<u>Team Classification assessment:</u>													
2. Fairness		1.00	.54	.17	.33	.59	.50	-.09	.70	.08	.13	.44	-.22
3. Positive job impact			1.00	.19	.27	.51	.54	.01	.52	.08	.13	.34	-.21
4. Staff support				1.00	-.03	.09	.18	.15	.16	.03	.02	.08	.01
5. Inmate support					1.00	.24	.28	.03	.33	.03	.08	.18	-.04
6. Improve inmate/staff relations						1.00	.46	-.17	.51	.11	.06	.47	-.26
7. Improve inmate/staff understanding							1.00	.01	.46	.06	.10	.32	-.17
8. Impact on inmates								1.00	-.10	.04	-.02	-.16	.24
9. Helpfulness									1.00	.09	.18	.37	-.23
<u>General attitudes:</u>													
10. Living assignment										1.00	.09	.07	-.02
11. Work assignment											1.00	.20	-.11
12. Staff												1.00	-.29
13. Alienation													1.00

and support for Team Classification, especially the improvement of inmate-staff understanding and relations.

The correlations between actualization and general inmate attitudes indicate that the higher the actualization, (1) the more favorable the attitude toward work assignment, cell assignment, and staff and (2) the lower the inmate's feeling of alienation. The associations between these general attitudes and each of the measures of inmate attitudes toward Team Classification are also presented in table 5. Team Classification's fairness, impact on job, improvement of inmate-staff relations and understanding, impact on inmates, and helpfulness are all strongly associated with the inmate's attitude toward staff (positive) as well as his feeling of alienation (negative). While attitudes toward Team Classification are largely unrelated to attitude toward both cell and work assignment, their strong association with alienation and attitude toward staff suggest that a favorably received Team Classification procedure reduces some sources of stress and conflict among inmates.

#### CAUSAL INFERENCES

In summary, data from inmate respondents reveal that (1) actualization is significantly associated with attitudes toward Team Classification, (2) actualization is significantly associated with inmate alienation and attitude toward both staff and work assignment, and (3) nearly all attitudes toward Team Classification are significantly associated with inmate alienation and attitude toward staff. Covariation, however, is not causation and the causal order among these factors cannot be established within the limitations of these data. Yet inferences of causation can be made on the basis of partial correlation coefficients. Subsequent analysis of inmate responses examines (1) the relationship between actualization and attitudes toward Team Classification when the effects of general attitudes are partialled out and (2) the relationship between actualization and general attitudes when the effects of attitudes toward Team Classification are partialled.

Performing the first set of partials, the findings reported in table 6 indicate little observed change in the relationship between actualization and each of the attitudes toward Team Classification when inmate alienation, attitude toward cell assignment, attitude toward work assignment, or attitude toward staff is controlled. This suggests that these general attitudes do not intervene between actualization and attitudes toward Team Classification; apparently, then, it is not the case that more favorable general attitudes lead to more favorable attitudes toward Team Classification. Furthermore, the coefficients between actualization and general attitudes, when partialling on attitudes toward Team Classification, remain quite strong. Yet some of these measures should reduce the bivariate relation to near-zero if attitude toward Team Classification intervenes completely in this relationship. Apparently there is a strong relationship between actualization of Team Classification and general inmate attitudes which is not dependent upon a favorable attitude toward Team Classification. Finally, the bivariate relation between attitude toward Team Classification and general attitudes is reduced but strong when actualization is partialled. This suggests that although actualization is related to both of the other measures, the relation between those other measures is not spurious.

The emergent conclusion is that (1) actualization has an effect on both attitudes toward Team Classification and on general attitudes, and (2) attitude

TABLE 6.--Partial correlation coefficients of attitudes toward Team Classification, general attitudes, and actualization, for inmates

Team Classification	Coefficient of actualization and attitudes toward Team Classification when general attitudes partialled				Coefficient of attitudes toward Team Classification and general attitudes when actualization partialled				Coefficient of actualization and general attitudes when attitudes toward Team Classification partialled			
	General attitudes				General attitudes				General attitudes			
	Living assignment	Work assignment	Staff	Alienation	Living assignment	Work assignment	Staff	Alienation	Living assignment	Work assignment	Staff	Alienation
Fairness	.40	.39	.28	.37	.04	.06	.26	-.22	.05	.08	.28	-.24
Positive job impact	.38	.37	.29	.35	.05	.07	.14	-.06	.05	.08	.35	-.24
Staff support	.08	.09	.07	.10	.07	-.01	.01	.07	.08	.14	.45	-.32
Inmate support	.10	.09	.04	.09	.02	.06	.11	.02	.08	.13	.44	-.31
Improve inmate/staff relations	.44	.44	.32	.40	.07	-.02	.28	-.10	.03	.12	.26	-.21
Improve inmate/staff understanding	.33	.32	.24	.30	.02	.04	.14	-.03	.07	.11	.37	-.27
Impact on inmates	-.26	-.25	-.21	-.21	.06	.01	-.06	.17	.10	.14	.44	-.27
Helpfulness	.37	.36	.28	.33	.06	.13	.18	-.09	.05	.06	.34	-.24

toward Team Classification has an effect on general attitudes. The multiple correlation coefficients of all Team Classification attitudes and actualization on attitude to living assignment, work assignment, staff, and alienation are .15, .20, .55, and .38, respectively, indicating that a significant amount of the variation in inmate alienation and attitude toward staff is explained.

Summarizing the analysis of staff data, the following associations have already been noted: (1) actualization is positively associated with Team Classification fairness, impact on job, and both staff and warden support, and negatively associated with improved inmate-staff understanding, impact on inmates, and inmate support; (2) actualization is positively associated with staff attitude toward inmates, work assignment, and other staff and negatively associated with punishment of inmates and role conflict; (3) nearly all staff attitudes toward Team Classification are significantly related to staff attitude toward work assignment, other staff, punishment of inmates, and role conflict. In addition, analyses not presented herein demonstrate that (4) whatever effect attitudes toward Team Classification has on general attitudes is largely limited to those staff that are members of Team Classification and (5) among members, actualization is more strongly associated with attitudes toward Team Classification among treatment staff but more strongly associated with general attitudes among custody staff.

The partialling technique was utilized to make inferences about the causal order of actualization, Team Classification attitudes, and general attitudes; and these data are provided in table 7. The bivariate relation between actualization and each attitude toward Team Classification remains unchanged when each general attitude is partialled, indicating that general attitudes do not intervene in the relation of actualization and Team Classification attitudes, and demonstrating that general attitudes are not a cause of both actualization and Team Classification attitudes. When partialled by attitudes to Team Classification, the covariation between actualization and general attitudes remains rather constant, suggesting that the effect of actualization on general attitudes is more direct than indirect (through attitudes toward Team Classification) and that attitude toward Team Classification is not the cause of both its actualization level and general attitudes. Finally, the coefficients of attitudes toward Team Classification and general attitude remain quite strong when partialled by actualization and, consequently, it cannot be said that general attitudes affect the (perceived) degree of actualization, which in turn affects attitudes toward Team Classification.

The results of the partial correlations suggest that general staff attitudes are directly affected by both the level of actualization and attitude toward Team Classification and that attitude toward Team Classification is affected by level of actualization. The multiple correlation coefficients of .35, .41, .56, .43, and .53 are obtained when actualization and Team Classification attitudes are correlated with staff attitudes toward inmates, work assignment, other staff, punishment of inmates, and role conflict, respectively. It is noteworthy that a significant amount of the variation in each of the five general attitudes is explained by actualization level and attitudes toward Team Classification.

As a final note, the relationships between actualization, Team Classification attitudes, and general attitudes for both inmates and staff were examined by each of the seven correctional facilities in the state. No systematic

TABLE 7.--Partial correlation coefficients of attitudes toward Team Classification, general attitudes, and actualization, for staff

Team Classification	Coefficient of actualization and attitudes toward Team Classification when general attitudes partialled					Coefficient of attitudes toward Team Classification and general attitudes when actualization partialled					Coefficient of actualization and general attitudes when attitudes toward Team Classification partialled				
	General attitudes					General attitudes					General attitudes				
	Inmates	Work assign-ment	Other staff	Punish-ment	Role conflict	Inmates	Work assign-ment	Other staff	Punish-ment	Role conflict	Inmates	Work assign-ment	Other staff	Punish-ment	Role conflict
Fairness	.28	.26	.28	.26	.25	.04	.10	.00	-.16	-.12	.12	.25	.25	-.13	-.24
Positive impact on job	.18	.15	.20	.15	.14	.07	.13	-.07	-.24	-.15	.12	.26	.27	-.14	-.25
Inmate support	.25	.21	.20	.24	.20	.16	.21	.28	-.16	-.24	.09	.23	.19	-.14	-.21
Staff support	.25	.19	.16	.25	.17	.01	.21	.41	.01	-.31	.13	.23	.17	-.18	-.21
Warden support	.20	.16	.15	.19	.15	.03	.12	.18	-.04	-.19	.13	.26	.22	-.17	-.24
Improved staff understanding	.11	.09	.13	.09	.10	.07	.09	-.03	-.18	-.09	.13	.27	.26	-.16	-.27
Improved inmate understanding	.06	.04	.10	.03	.06	.07	.08	-.12	-.18	-.04	.08	.17	-.12	-.11	-.13
Improved inmate/staff understanding	.38	.33	.32	.37	.31	.13	.26	.34	-.15	-.36	.13	.28	.27	-.17	-.27
Impact on inmates	.29	.25	.26	.27	.22	.22	.26	.22	-.31	-.42	.06	.20	.19	-.08	-.15

variation was found according to size or security level. In general, the impact of actualization level varies somewhat among the facilities and the higher the level of actualization, the more favorable the inmate and staff attitude toward Team Classification and general conditions.

## IMPLICATIONS

Numerous implications emerge from the findings to inform and instruct those who would administer such a classification program. Among the major factors to be considered are the following: (1) membership and participation per se are not likely to produce the given results; decisive measures must be taken to maximize the degree to which the operation of Team Classification adheres to the tenets and guidelines around which the concept has developed; (2) those tenets and guidelines can be put into operation and they are not mutually incompatible; (3) when well implemented, Team Classification appears to improve inmate understanding of and relations to staff and reduce inmate alienation; (4) when well implemented, Team Classification appears to improve staff relationships, reduce role conflict among staff, and improve their outlook on their job and on inmates; (5) the effects of Team Classification will be greater on members than on nonmembers, and treatment staff will respond in a different manner than custody staff. As a caveat, it should be noted that this evaluation focused on desired or positive consequences of Team Classification; little can be said of the negative consequences or costs of such a program.

Perhaps the most obvious implication to emerge from this evaluative effort is the need for a more systematic and longitudinal evaluation, an experimental design initiated prior to the program's implementation to randomly assign subjects, assemble preprogram data, monitor the development and operation of the program, and gather postprogram data for comparison purposes. The utilization of cross-sectional research designs provides inadequate and, at best, inconclusive results regardless of the sophistication of the data collection and data analysis processes. Because of the inherent limitations within this evaluation of Team Classification, the implications derived from the data must be viewed as more suggestive than informative. Some associations have been observed in this, the "Reconnaissance Phase," and the impetus now must be to use an experimental design to measure the extent of the effect (Rossi, 1972).

## NOTES

1. The research objectives and data analysis are necessarily abbreviated here. The complete report is available upon request to the Missouri Division of Corrections.
2. There is no doubt that an experimental design with pretest and control group is desirable. In its absence, however, quasi-controls can be made by examining the observed relationships according to membership and participation in the program.
3. The relationships among the measures of attitudes toward Team Classification and those among the general attitudes presented in tables 4 and 5 are worthy of examination. Limitations of space, however, prohibit such a discussion and the reader is encouraged to utilize all the data reported rather than only that explicitly discussed.

## BIBLIOGRAPHY

- Burns, Henry.  
Corrections: Organization and Administration. Minnesota: West Publishing Company. 1975.
- Directive on Team Classification.  
Missouri Division of Corrections, 101-080. Jefferson City, Missouri. 1975.
- Hagan, Charles R., and Charles R. Campbell.  
"Team Classification in Federal Institutions." Federal Probation 23 (March):30-36. 1968.
- Loveland, Frank.  
"The Classification Program in the Federal Prison System: 1934-1960." 15 Federal Probation (June):7-12. 1960.
- Rossi, Peter H.  
"Boobytraps and Pitfalls in the Evaluation of Social Action Programs," in Weiss (ed.) Evaluating Action Programs. Boston: Allyn and Bacon, Inc. 1972.
- Suchman, Edward.  
"Evaluating Educational Programs: A Symposium." Urban Review 3 (February):15-17. 1969.
- "Action For What? A Critique of Evaluative Research." M. O'Toole (ed.) The Organization, Management, and Tactics of Social Research. Cambridge: Schenkman Publishing Company. 1970.

RESIDENTIAL CORRECTIONS PROGRAMS IN MINNESOTA:  
AN EVALUATION REPORT\*

Michael J. McMahon, Ph.D.  
Governor's Commission on Crime Prevention and Control

A. INTRODUCTION

Under the direction of the Governor's Commission on Crime Prevention and Control, the Evaluation Unit is charged with providing the Commission with the kinds of evaluation information which may be used to make policy decisions about the dispersement of LEAA funds. The major policy at issue was whether the Commission should continue to provide funds for the development and implementation of new halfway houses for parolees.<sup>1</sup> Given this policy orientation for evaluation research, the decision was made to analyze halfway houses as a group and focus on their common goals, instead of developing individual evaluations for individual projects.<sup>2</sup>

For purposes of this evaluation, the term "halfway house" refers to a "residential facility designed to facilitate the transition of paroled, adult ex-offenders who are returning to society from institutional confinement." The limitation to adults serves to distinguish halfway houses from juvenile residences which serve juveniles. The identification of paroled ex-offenders as the target population of halfway houses distinguishes the primary intervention stage of these projects from that of P.O.R.T. projects<sup>3</sup> in which the primary intervention stage is probation.

B. HALFWAY HOUSES

Halfway houses are funded to achieve specific goals by implementing treatment programs for their clients. A review of the goals and objectives of these programs helps to present an overview of the halfway houses included in this evaluation.<sup>4</sup> There are two purposes for discussing the goals and objectives of halfway houses. First, statements of goals and objectives provide a basis for describing what the projects are attempting to accomplish. Second, goals and objectives are the standards by which projects are held accountable. The LEAA program is based on a management-by-objective approach. This approach requires grantees to focus on and to articulate what they plan to accomplish, rather than

---

\*This research was supported by grants #0319007375 and #0321007376 awarded to the Evaluation Unit by the Governor's Commission on Crime Prevention and Control. Points of view and opinions stated in this report are those of the author and do not necessarily represent the official position or policies of the Governor's Crime Commission.

simply stating what they plan to do. Thus, the accountability of recipients of LEAA funds is based, in part, upon their achievement of stated goals.

As operationalized by the Governor's Commission on Crime Prevention and Control, the term "goal" refers to a statement of the impact, or effect, the project should have if it is successful. Although there are a number of differences among halfway houses, the eight programs included in this evaluation share the following goals:<sup>5</sup>

- (A) To reduce the recidivism rate of the client population relative to the rate of a comparable group of parolees who do not participate in halfway house programs.
- (B) To increase (i) employment, (ii) educational level, and/or (iii) vocational skills of their clients.
- (C) To rehabilitate ex-offenders placed in the projects.
- (D) To demonstrate that ex-offenders can be effective staff members of halfway houses.

The immediate focus of halfway house programs is to resolve economic, psychological, and social problems of their clients. Project planners believe criminal behavior is a result of or response to problems of these types. The objectives of halfway houses encompass a number of basic approaches developed to resolve these problems and to achieve program goals. First, most halfway houses begin providing counseling to clients prior to their release from correctional institutions. Prerelease counseling concentrates on helping inmates develop realistic parole plans which will be implemented when the client begins residence in the community project. Second, halfway houses provide either services and counseling for all types of problems their clients have, or arrangements for needed services via referrals to appropriate agencies. Third, many halfway houses employ ex-offenders, particularly in counseling positions. Project planners believe the effectiveness of halfway houses depends on the development of trust between the client and his counselor. Project planners also believe that this trust is more easily and effectively developed when the counselor is an ex-offender who has had to face the same types of problems and obstacles the client faces. Fourth, most halfway houses encourage clients who have completed residence to return to the projects for counseling and support. Finally, halfway houses hope to become community projects which are effective in helping ex-offenders avoid returning to criminal behavior. Both individual and group counseling techniques are used throughout halfway house programs.<sup>6</sup>

The ultimate goal of halfway house programs is to reduce the recidivism rate of the client population, i.e., achievement of Goal (A). The "philosophy" underlying these programs is that through the achievement of Goals (B) and (C), achievement of Goal (A) will follow. Consequently, these programs concentrate on solving the problems of individual clients and increasing the client's education and/or employment prospects. The combination of individual and group counseling within the project and the use of existing community agencies outside the project is thought to be the most realistic approach toward achieving Goals (B) and (C).

There is no single theory of criminal behavior and rehabilitation on which residential corrections programs are based. From the goals and objectives of these programs a number of possible causes of criminal behavior can be inferred: lack of education, lack of employable skills, poor employment histories, drug or alcohol abuse, immaturity, family problems, and so on. The programs of halfway houses are designed to help clients resolve these problems. Consequently, the residential programs are hybrids of theories of criminal behavior and rehabilitation. None of these programs is specifically designed to test those theories.

### C. EVALUATION OF EFFORTS

Two types of evaluation measures were used to analyze the results for halfway houses: measures of efforts and measures of effects. Two aspects of the efforts projects expend serving clients are presented. First, the flow of clients through projects and the occupancy rates of projects are discussed in order to provide a rough measure of the efficiency with which projects are used by the criminal justice system. Second, the costs which projects expend in serving clients are considered.

The average length of residence in halfway houses varied, by project, from 2 months to 6½ months. In each project, clients who completed the residential program tended to remain in residence longer than clients terminated prior to completing their residential programs. The average length of residence for those who completed residence exceeded 9 months at one program, but was less than 6 months at the other halfway houses. Thus, halfway houses have relatively short periods of residence for their clients.

Occupancy rates measured from the time these projects began accepting residents varied from 40.0% to 84.4%. However, during 1975, the occupancy rates of halfway houses increased substantially and varied from 60.0% to 94.0%, with an average rate of 74.6% per project. Although occupancy rates of halfway houses increased in 1975, these results do not suggest a need for new residential projects. Few halfway houses have consistently operated at or near capacity. Moreover, at least part of the increase in the occupancy rates of halfway houses is due to increased admissions of probationers in 1975. Thus, these results indicate that there are sufficient residential programs for parolees.

A cost/client/day measure was based on the expenditure of the project and the average number of clients/day. During 1975, the costs of halfway houses serving males varied from \$13.59 to \$38.37/client/day. As a comparison, the costs/inmate/day were \$26.99 at the State Prison and \$31.03 at the Reformatory for Men. Three of the halfway houses serving males had costs/client/day lower than those of both the Prison and the Reformatory, whereas only two halfway houses had costs/client/day appreciably greater than that of the Reformatory. Halfway houses serving females had costs which varied from \$14.57 to \$31.41/client/day. These costs were less than that of the Correctional Institution for Women, which had a cost/inmate/day of \$65.02. Consequently, halfway houses do operate at costs comparable to or less than those of corresponding state correctional institutions.<sup>7</sup>

#### D. EFFECTS OF RESIDENCE

Evaluation of the effects of these projects employed two types of measures. First, measures were taken to compare socioeconomic status at intake and at termination from residence to determine whether halfway houses were affecting the problems with which their clients entered the programs. Second, measures of the recidivism of clients during and after residence were made.

The first measure of the effectiveness of halfway house programs is the proportion of clients who satisfactorily completed their residential programs. Because halfway house programs use either a phase progression system or a contract to determine when a client has completed the residential program, "satisfactorily completed residence" is defined as "completion of the phased residential program or residential contract." Clients who have satisfactorily completed residence are those clients who, in the judgment of project staff, have achieved the objectives with which they began residence. Clients who failed to satisfactorily complete residence are those clients who were terminated from the program for reasons of lack of cooperation, poor adjustment, absconding, re-arrested, convicted of a new offense, or revocation of parole or probation. Clients who neither satisfactorily completed nor failed to satisfactorily complete the residential programs were terminated for neutral reasons, which include voluntary termination, withdrawn by referring agency, transferred to another program, and death.

Based on this classification, there are three classes of reasons for which a client may have been terminated from residence: satisfactorily completed residence, failed to satisfactorily complete residence, and neutral. Of 625 clients on whom termination data were available, 32.8% satisfactorily completed residence. However, 45.0% of the halfway house clients failed to satisfactorily complete their residential programs. The primary reasons for failing to satisfactorily complete residence were "absconded" (19.8%) and "lack of cooperation/poor adjustment" (17.8%). The remaining 22.2% were terminated from residence for neutral reasons. The fact that so few clients satisfactorily complete the program suggests that, for a variety of reasons, residential corrections programs are an inappropriate form of rehabilitation for a majority of the persons for whom these programs are now being used.<sup>8</sup>

The data provide evidence that halfway house programs are helping to increase employment among their residents: there was an increase of 24% from intake to termination among all halfway house clients. Moreover, clients who satisfactorily completed residence had an increase of 42%. However, changes in educational level and vocational skills were slight. Even clients who satisfactorily completed residence in halfway houses showed little increase in education and vocational training, although 9% of the satisfactory clients completed high school level education during residence. Consequently, progress toward fulfilling the goal of increasing education, vocational skill, and employment is due primarily to increased employment of halfway house residents.

Halfway house residents had a 13% reduction in perceived financial problems and a 24% reduction in clients relying on governmental assistance for support. As was the case with employment, clients who satisfactorily completed residence were much more likely to resolve their financial problems and become independent of governmental assistance than were those who failed to satisfactorily complete residence.

Although clients who satisfactorily complete residence, in general, have more favorable changes in socioeconomic variables than do clients who fail to satisfactorily complete residence, these results should not be misinterpreted. The relationship between satisfactory completion of the program and favorable change in socioeconomic variables is, in part, definitional because such changes may be a part of the contracts for the residential program. However, if residents who satisfactorily complete the program did not show greater improvement than unsatisfactory clients, the value of remaining in and satisfactorily completing residence could be questioned. Yet only about one-third of the halfway house clients satisfactorily complete residence. If these programs are to have an increased impact, halfway houses will have to increase the number of persons who satisfactorily complete residence, since it is these persons who show the greatest improvement while in the program. With the exception of employment, the overall impact of these programs has been slight.

#### E. RECIDIVISM

The analysis of recidivism of halfway house clients used two measures of recidivism: total convictions and revocations, and felony convictions and revocations. This analysis also looked at recidivism of clients during residence, at recidivism of clients who satisfactorily completed residence and clients who failed to satisfactorily complete residence following termination from residence, and at recidivism of halfway house clients and a comparison group.<sup>9</sup>

The recidivism of clients who satisfactorily completed residence was compared to the recidivism of clients who failed to satisfactorily complete residence.<sup>10</sup> The results show that during the first 6 months and 12 months following termination from residence, clients who satisfactorily completed residence had a significantly lower recidivism rate than did clients who failed to satisfactorily complete residence. This conclusion holds for recidivism measured in terms of felonies and revocations and in terms of total convictions and revocations. However, by the end of the 24-month followup period, there were no significant differences between the recidivism of satisfactory clients and of unsatisfactory clients. By the end of the 24-month period, 21.1% of the clients who satisfactorily completed residence were recidivists and 24.5% of the clients who failed to satisfactorily complete residence were recidivists. Consequently, while these results show that satisfactory clients are less likely to recidivate than are unsatisfactory clients during the first year following termination from residence, they also show that satisfactorily completing residence in a halfway house does not have a significant long-term effect on recidivism.

The recidivism of halfway house clients was compared to that of a comparison group in two overlapping time periods. First, comparisons were made for an "at-risk" period which is concerned with the recidivism of halfway house clients from intake to residence and with the recidivism of comparison group members from placement on parole. Second, the recidivism of halfway house clients was measured from termination from residence and compared to that of the comparison group members.

The results show that during the 6-month at-risk period there were no significant differences between the recidivism rates of halfway house clients and comparison group members. This conclusion holds for recidivism measured in terms of felonies and revocations and in terms of total convictions and revocations. Since this is also the period during which the influence of halfway

house programs ought to be most effective, it appears that participation in halfway house programs has no detectable effect on recidivism. However, this conclusion must be viewed in relation to the differences between halfway house clients and comparison group members.<sup>11</sup>

By the end of the 12-month at-risk period, the results show that the comparison group had significantly more total recidivism than did the halfway house group, although there was no significant difference between the groups when recidivism was measured in terms of felonies and revocations. By the end of the 24-month at-risk period, the recidivism of comparison group members was significantly higher than that of halfway house clients on both measures of recidivism. However, given the results from the 6-month at-risk period--when the influence of halfway house programs should be strongest--differences between the groups in later followup periods cannot be attributed solely to participation in halfway house programs.

But at-risk recidivism information includes the recidivism of halfway house clients who entered programs for a short time and recidivated during residence. Consequently, the recidivism of halfway house clients following termination from residence (and, thus, not including recidivism during residence) was compared to the recidivism of the comparison group.<sup>12</sup> The results of this comparison were essentially the same as the results for the at-risk periods. During the first 6 months of the postresidence followup period, there were no significant differences between the recidivism of halfway house clients and that of comparison group members. However, in the 12-month and 24-month postresidence followup periods, the recidivism of comparison group members was significantly higher--on both measures of recidivism--than that of halfway house clients. Again, because there were no significant differences in the 6-month postresidence period, these results indicate that participation in halfway house programs does not significantly affect recidivism.

The analysis of the recidivism of halfway house clients sought answers to two questions. First, what effect does satisfactorily completing a halfway house program have on recidivism? On the basis of the data, satisfactorily completing a halfway house program decreases the likelihood of recidivism in the first year following termination from residence but has no long-term effect on recidivism. By the end of the first 2 years following residence, there were no significant differences between the recidivism of those clients who satisfactorily completed residence and of clients who failed to satisfactorily complete residence. Second, what effect does participation in a halfway house program have on recidivism? Subject to the comparability of the comparison group members and halfway house clients on those characteristics relevant to recidivism,<sup>13</sup> the results show no detectable short-run differences in recidivism which may be attributable to participation in halfway house programs. Long-term differences, which were detected, cannot be attributed solely to participation in halfway house programs.

#### F. MUSINGS ABOUT EVALUATION

Where does this type of evaluation lead? To answer this question, I want to note some of the advantages and disadvantages of this approach to evaluation and to suggest that, while it fits the needs of the Governor's Commission on Crime Prevention and Control of Minnesota, it may not represent an approach which would be useful in other agencies.

Among the advantages of this type of approach are the following: First, by analyzing the effectiveness of a number of residential programs by measuring their ability to achieve common goals, the Evaluation Unit is able to provide information on a number of approaches to postinstitutional residential programming. Second, this approach allows the Unit to evaluate a number of programs with limited personnel. Third, this approach provides the Governor's Commission on Crime Prevention and Control with the kind of evaluation information needed to make policy decisions about LEAA funding in Minnesota. Fourth, the evaluation results are available for other states considering similar policy decisions. Fifth, this type of approach to evaluation can be implemented in those contexts of criminal justice research and evaluation in which evaluators do not have control of program variables and/or access to control groups.

Among the disadvantages of this approach are the following: First, evaluators do not have control of program variables and/or access to control groups. Consequently, many program variations are implemented within single programs which may affect effectiveness. This also hinders generalizability of results. Second, this type of approach does not allow us to fully analyze program components and strategies. Consequently, although we may be able to say a great deal about halfway houses as a group, we do not have detailed information on individual projects. Thus, while we are able to tell a project whether it is doing better or worse than other projects, we cannot say why this is so. Third, although the results of this type of evaluation are available to other agencies considering funding halfway houses, no results are available which indicate what the specific program structures of those projects should be.

Despite these limitations, policy-oriented evaluations which group projects according to common goals and target populations do serve the needs of the Governor's Crime Commission. Policy decisions form the basis for the funding decisions of the Commission. In this context, the major policy decision for this report is whether the Commission should continue to fund new halfway houses. Our approach to the evaluation of halfway houses provides information about the effectiveness of halfway houses. At a more specific level, the Commission must make decisions about funding individual programs for particular target populations. But the Commission has traditionally allowed applicants wide leeway in the development and implementation of treatment modalities for specific target populations--provided that applicants can demonstrate the need for a program for a specific target population. Refunding decisions for second- and third-year grants do use evaluations, but primarily in a monitoring role not as measures of effectiveness. Finally, Minnesota is not a large criminal justice state in terms of potential target populations. Although the implementation and evaluation of experimental treatment programs are important, replication of individual programs is not a major need in Minnesota. For example, there are variations among the eight halfway houses included in this report both in terms of program structure and in terms of target populations. This reflects the diversity of subpopulations within the parolee population. Indeed, although a number of halfway houses were quite similar in their original applications, they have each developed in such a way that they are now recognized as being adept at handling particular types of offenders.

In summary, these evaluations lead to fulfillment of the needs of the Governor's Commission on Crime Prevention and Control relative to evaluation information. This agency has a strong commitment to evaluation and seeks information which will help it form policies for the dispersement of LEAA funds.

But it views these funds as experimental funds for new treatment programs and allows grantees to develop and implement their own programs. That is, it is not committed to funding any particular types of treatment programs. An agency which is primarily concerned with the effectiveness of treatment modalities and for selecting among these modalities would require a different approach to evaluation research.

## NOTES

1. It should be emphasized that this evaluation was directed by the policies of the Governor's Commission on Crime Prevention and Control relevant to the allocation of LEAA funds for corrections programs in Minnesota. In addition to the policy on funding halfway house programs--for which effectiveness and recidivism results are most important--policy recommendations were based on the need for new programs (using occupancy data), on the costs of programs (using cost/client/day data), and on the Commission's policy that LEAA monies should be used in Minnesota to develop and test innovative criminal justice programs. Finally, it must be noted that such policy decisions are political decisions for which evaluation results are only one type of information available to the decisionmakers.
2. The Governor's Commission on Crime Prevention and Control has a firm policy of funding individual programs for a maximum of 36 months. Given the amount of time required to collect and analyze data on the effectiveness of individual programs, the 3-year time constraint also makes evaluations of individual programs impractical from the standpoint of funding decisions on individual grants. However, this does not mean that individual programs are ignored. The Evaluation Unit normally provides short reports on projects being considered for continuation funding. The report on the first year of operations is generally a descriptive report which covers such topics as program structure, staff organization, staff background and training, startup problems, and descriptions of the clients. Second-year reports are presented in the form of data summaries, including descriptions of clients, effectiveness of residence, and recidivism. Both reports are used to inform Commission members about the project and to describe the progress which has been made. After 3 years of LEAA funding, continuation funding must be obtained from state and local sources. The Evaluation Unit has frequently prepared and presented reports on individual projects seeking continuation funding from other agencies.
3. "P.O.R.T." stands for "Probationed Offenders Rehabilitation and Training." Whereas halfway houses normally accept residents following incarceration, P.O.R.T. projects normally accept residents as an alternative to incarceration. Kay Knapp reports on P.O.R.T. projects in Minnesota in "P.O.R.T. Projects: Alternatives to Incarceration?" at this conference.
4. One of the advantages of policy-oriented evaluation research is that one can concentrate on a number of different programs which share the same goals. This allows one to increase the data base for measures of effectiveness. However, there are corresponding disadvantages to this approach. For example, some programs may have unique goals which are not shared by other programs. These may not receive the attention they deserve in policy-oriented evaluations.
5. The eight programs are: Alpha House, a program for male, adult parolees, which has recently developed a program for sex offenders; Anishinabe Longhouse, a program for Indian men; Anishinabe Waki-igan, a program for Indian men released from the Minneapolis Workhouse (Waki-igan closed in January 1975); Freedom House, a program for male and female adults, particularly

offenders with chemical abuse problems; Pi House, a program for female parolees which closed in January 1976; Reshape, a program for chemically dependent parolees from the Reformatory for Men; Retreat House, a program for male parolees; and 180 Degrees, a program for male and female parolees.

6. A complete description of halfway house programs is presented in Residential Community Corrections Programs: A Preliminary Evaluation, Governor's Commission on Crime Prevention and Control, (April 1975), chapter 4, pp. 69-129.
7. This conclusion depends on whether these costs are comparable. Because halfway houses usually accept clients paroled from state institutions, actual treatment costs include costs of incarceration and of halfway house placement. If residents are paroled to halfway houses in lieu of further incarceration, these data indicate that halfway houses do not increase and may even decrease treatment costs. However, if halfway houses are used as supplements for paroles which would have been granted anyway, costs of halfway house residence represent increases in costs of treatment. Cost effectiveness analysis would help to resolve this problem. Charles M. Gray and Chris Johnston-Conover present a model for such analyses in "Cost Effectiveness of Residential Community Corrections: An Analytical Prototype" at this conference.
8. Reasons for termination from residence in halfway houses were reviewed for 1975 and for prior years. Prior to 1975, 31.2% of the halfway house clients satisfactorily completed residence, 43.7% failed to satisfactorily complete residence, and 25.1% were terminated for neutral reasons. During 1975, 35.2% satisfactorily completed residence, 47.0% failed to do so, and 17.8% terminated for neutral reasons. The increase in clients who satisfactorily completed residence does not appear to be significant. Overall, only 3.3% of the halfway house clients were terminated for reasons of new convictions and revocations.

Further data on reasons for termination and other program effects are presented in Residential Community Corrections Programs in Minnesota: An Evaluation Report, Governor's Commission on Crime Prevention and Control, (November 1976), chapter 5, pp. 116-138. Appendix G of this report presents a brief literature review which indicates that the low rates of program completion in Minnesota's residential community corrections programs appear to be typical of most residential programs.

9. Comparison group members were randomly selected among persons who were released from the State Prison, State Reformatory for Men, and Correctional Institution for Women and who met the formal, objective selection criteria for admission to halfway house programs; i.e., who were adults who had been convicted of two or more offenses and were released from state correctional institutions. Appendix D of Residential Community Corrections Programs in Minnesota: An Evaluation Report presents a comparison of the characteristics of halfway house clients and comparison group members. Appendix F reviews the methodology for the analysis of adult recidivism.
10. Recidivism data on clients who were terminated from residence for neutral reasons were not included in this part of the recidivism analysis.

11. Comparisons of halfway house clients and comparison group members revealed major differences between these two groups in terms of sex, ethnic background, and months incarcerated for most recent conviction. However, preliminary analysis found no relationship among background characteristics, program participation, and recidivism. In particular, no relationships were found for sex, ethnic background, and months in institutions for most recent conviction. Furthermore, the recidivism of halfway house clients was reanalyzed using only those clients who had been released from state institutions. Comparisons between this group of halfway house clients and comparison group members led to the same results as the comparisons using all halfway house clients.
12. In effect, the at-risk analysis assumes that the "treatment" program is incarceration and views regular parole and parole plus halfway house placement as two posttreatment alternatives. In contrast, the postresidence analysis views incarceration and incarceration plus halfway house placement as two forms of "treatment" with nonresidential parole as the normal post-treatment modality.
13. Although preliminary analysis does not indicate that differences between these two groups do significantly affect recidivism, uncontrolled differences might have effects. In particular, information on actual chemical abuse/dependency and on juvenile correctional histories--which are thought to be relevant to recidivism--was not available on members of both groups. Moreover, it is not evident that we have knowledge of all those factors which are relevant to recidivism.

82922

SALT LAKE CITY:  
PROTOTYPE EVALUATION OF DES MOINES REPLICATION\*

William M. Rhodes  
Thomas Blomberg  
School of Criminology  
Florida State University  
Tallahassee, Florida

Steven T. Seitz  
Department of Political Science  
University of Illinois  
Urbana, Illinois

I. BACKGROUND

In the last several years the National Institute of Law Enforcement and Criminal Justice has attempted to discover exemplary criminal justice programs that demonstrate both a measurable and marked level of effectiveness. The criteria used in the selection of exemplary projects includes: goal achievement, replicability, measurability, efficiency, project accessibility, and a willingness to work with other communities in the establishment of similar projects. The Community-Based Corrections Program of Polk County (Des Moines), Iowa was selected as such an exemplary program.

Essentially the Des Moines program offered four alternatives to formal criminal justice processing: release on own recognizance, pretrial supervised release, probation, and residence in a residential center which offers work and educational release. In an effort to facilitate transfer and further evaluate the Des Moines Program, the Technology Transfer Division of the National Institute selected five cities for the replication of the project. In addition, the Institute contracted with the School of Criminology at Florida State University to evaluate the five sites' implementation of the Des Moines program. The evaluation has several purposes. Included among these are: providing the local jurisdictions data feedback on various project problems; questions and impact findings; determination of each jurisdiction's replication success or failure; and specification and explanation of variations in the replication success or failure. Overall, the purpose of the evaluation is to generate an analysis that provides a systematic and comparative basis from which sound and justifiable decisions can be made concerning continuation of the Des Moines program concept and the efficient transfer of the program to other local jurisdictions throughout the nation.

---

\*Prepared for delivery at the National Conference on Criminal Justice Evaluation, February 22, 1977. We would like to thank Charles Massey who administered the Salt Lake City evaluation office and wrote early drafts of the organizational analysis section. Bernard Gropper of the Office of Evaluation and Robert Burkhart of the Office of Technology Transfer made invaluable comments on earlier drafts of this paper. This evaluation was funded by LEAA grant No. 77-NI-99-0020.

Preceding page blank

This paper is a condensed version of a manuscript presented at the National Conference on Criminal Justice Evaluation, February 22, 1977. As such, it represents only part of the research findings pertaining to the replication in Salt Lake City, and completely omits findings relevant to other sites. A complete version of the Salt Lake City findings (including tables and research methodology) is available on request.

## II. SITE SELECTION

By 1974 criminal justice in Salt Lake County faced some serious problems. First, the crime rate was high and getting worse; second, public intoxication was and is handled as a criminal matter; third, weekend arrests were often held in jail for court arraignment on Monday; fourth, and as a consequence of the first three, the county jail was overcrowded; and fifth, budgetary pressures threatened to eliminate special alcohol and probation programs, thus placing further burdens on the county jail.

In late 1973, the National Institute of Law Enforcement and Criminal Justice announced monetary support for replications of the court services program developed in Des Moines, Iowa. Salt Lake officials learned about the project through the LEAA Denver Regional Office and the Utah Law Enforcement Planning Agency. LEAA officials reported that \$250,000 of Institute grant money and up to \$400,000 of LEAA discretionary grant money would be made available to selected communities interested in replicating the Des Moines Community Corrections program.

State planning officials were of the opinion that the Des Moines program could help Salt Lake reduce jail overcrowding by providing acceptable alternatives for selected offenders. The State Planning Agency endorsed Salt Lake County as a replication site and the county began preparing a Des Moines replication grant proposal.

After preliminary discussion among state and local officials, a member of the Regional Planning Agency requested support from the Salt Lake County Commissioners. The Commissioners were unwilling to firmly commit the county to the new program, but it did ask the Regional Planning Agency to prepare a Federal grant application. The application was prepared with the assistance of Urban Rural Systems Associates (a technical assistance consulting firm under contract to NILECJ) and a steering committee comprised of a Salt Lake City Judge, a member of the Utah Board of Corrections, and several representatives from local criminal justice and service agencies.

The major purpose of the proposed grant was to reduce jail population and operating costs by providing safe community alternatives. These alternatives included: pretrial release (release on own recognizance), pretrial services (supervised release), misdemeanor probation, and a community rehabilitative facility. These replication components were to be administered by the new Department of Court Services.

The grant proposal aimed to do more than relieve the overcrowded jail. It proposed delivery of services to needy clients, and it proposed increased criminal justice coordination in Salt Lake County. In his report to the county commissioners on July 1, 1974, the county auditor argued that the proponents of

the Des Moines project did not adequately take into account these other functions which had not been publicly aired or clarified by professional staff members within the local law enforcement and criminal justice community.

Most political observers in Salt Lake agreed that the jails were overcrowded. They believed a program to address this need would meet with little political opposition, particularly if it promised to reduce jail operating costs. However, interviews with CJS staff indicated that an attempt to increase coordination among the various criminal justice agencies would seem likely to raise opposition. In addition, community surveys indicate that a project emphasizing "rehabilitative" as opposed to "retribute" justice would also incur considerable popular objection. At least one member of the three-man Board of Commissioners was notably opposed to the replication project (because it would expand county correctional responsibilities and would probably result in increased financial obligations). A second commissioner was, for the most part, indifferent. Proponents of the project chose to emphasize the jail reduction function which had the greatest political support. Had all three functions been publicly aired, the chances of gaining commissioner approval for the proposal might have been considerably less.

### III. SUPPORT AND OPPOSITION

Criminal justice personnel in Salt Lake County generally supported the grant proposal. However, law enforcement officials objected to the project's general philosophy. Their criticism centered upon the plan to release defendants from jail soon after admission. (Ultimately, the jail staff did support the project, issuing favorable press releases during a bail bondsmen strike.) Neither law enforcement nor jail personnel organized to oppose the proposal.

Some judges and court personnel were actively supportive. (One judge served on the steering committee and also assisted the project during implementation. He actively intervened in a serious problem involving staff dissatisfaction.) The judges and court personnel generally agreed that a project like Des Moines should be tried in Salt Lake County. However, they viewed the project as a temporary solution to the overcrowded jail and thought more time was needed to develop a more permanent solution.

The bail bondsmen offered the most vocal and organized resistance to the grant proposal. An attorney was even retained to represent them. They especially criticized the ROR program, claiming that it would attract "safe risks" clientele, leaving them with clients more likely to forfeit bond. (Their fears were not unfounded. Once the project began, three bonding offices closed, and the bondsmen went on strike for 18 days.)

The State Department of Corrections also opposed the grant proposal. State officials publicly argued that resource constraints prevented their supplying promised match money. According to some observers, the Department of Corrections felt that the intensive probation program and the community corrections facility interfered with state authority in corrections. (This jurisdictional dispute had a serious impact on later attempts to implement the project.)

The county commissioner asked the county auditor to review the grant proposal, and he in turn prepared a detailed analysis of the project and its likely impact. This report, presented at a July 1, 1974 Commission meeting, concluded

that a local correctional alternative and a bail reform program would indirectly reduce the period of incarceration and hence the need for future jail space. However, the auditor also recommended that the state/local jurisdictional dispute be resolved before a final decision was made.

At their July 1 meeting, the commissioners voted (two against one) to approve the proposal and accept responsibility for administering the project. Support for the project was based on the need to reduce jail overcrowding. However, they did not resolve the jurisdictional questions. Dissenting Commissioner McClure cited this failure and also argued that the Commission had ignored the question of long range follow-on funding.

#### IV. IMPLEMENTATION

On July 10 the county commissioners placed the Des Moines project under the general supervision of the Director of Social Services. The commissioners also agreed to advertise for a project director, who was selected on September 16.

A Pretrial Release on Recognizance (ROR) program existed in Salt Lake before the Des Moines replication. The new project director, therefore, was asked to incorporate this program into the new court services project, and the ROR program director and the eight screening officials were retained. (Each screener worked 20 hours a week, with interviews held 7 days a week.)

Pretrial services, which provided Supervised Pretrial Release, was scheduled to begin in February 1975. More clients were accepted than anticipated, however, and consequently an attempt was made to begin the program earlier.

During November and December 1974, the service program director sought to establish working relations with community service agencies. Many of these agencies were not particularly cooperative, arguing that they were already operating at full capacity and could not afford to accept new cases. Most of these objections were generally resolved, and the first case referral was made on January 9, 1975.

The Misdemeanant Probation Program fared less well. Two problems delayed implementation of this program: money and a jurisdictional dispute. The Utah Division of Corrections decided it could not honor its commitment to supply matching funds. Eventually, an agreement between the Utah State Division of Corrections and the Court Services Project resolved the funding and jurisdictional crisis. The State Division of Corrections agreed to provide "in kind" assistance to misdemeanor probation, sharing with it state personnel and office space. On July 1, 1975, the Misdemeanant Probation Program finally began, staffed by three state probation officers, one of whom served as Director of Misdemeanant Probation.

Two problems plagued the rehabilitation facility: location and legal custody of its residents. The rehabilitation facility needed space to house 40 male and 10 female residents. The facility was eventually housed in a wing of the old county hospital using space made available by the county Commissioners. The last major problem centered upon legal custody of the facility's residents. The sheriff had custody over jail prisoners, but his legal responsibility for facility residents was unclear. To resolve the custody issue, the sentencing

judge was required to order the release of jailed inmates to the rehabilitative facility.

The program director for the rehabilitative facility was appointed in March 1975. The program director had a staff of 15 counselors, an assistant director, and several volunteers. The rehabilitative facility received its first clients on June 1, 1975.

A year had passed before all project programs were in operation. The county auditor, several local criminal justice personnel, and some politicians blamed the project director for inadequate administration. (He was later removed.) But other sources also impeded the Des Moines replication: (1) The State Division of Corrections, through the jurisdictional dispute and the decision not to honor its funding commitments; (2) the county auditor's hiring freeze; (3) an extremely complex funding formula with consequent uncertainty in financing; (4) lingering questions on the legality of the county's correctional responsibilities. Finally, the project did not have a strong political base in Salt Lake. The jurisdictional dispute brought attention to the project's role in developing a county-level alternative to corrections, and Salt Lake officials did not want to publicly debate or to assume the financial burden of a county-level corrections program.

None of these impediments prevented implementation of the replication project. Salt Lake County's version of the Des Moines project did work, but its functions extended far beyond reducing the jail population. The Court Services Project was innovative and it caused several adjustments within the criminal justice process in Salt Lake County.

#### V. OPERATION

The administrative division of Salt Lake's Court Services Project performs several functions. It prepares the budget, oversees staff selection, produces a quarterly report for local justice agencies, and compiles a semimonthly report for the judges. The staff training function fares less well. In-service training relies upon joint volunteer-staff programs. Staff attendance at these sessions is poor. Finally, the administrative division gathers cost information similar to that compiled in Des Moines. The operating costs in Salt Lake appear consistent with those reported for the program in Des Moines.

Pretrial release screeners are physically housed in the county jail. They attempt to interview 75% of those booked into the jail. The interviewers explain pretrial release to the prospective client and administer a profile questionnaire. The questionnaire awards points for stable community ties, which are based on such factors as residency, employment, and prior record. Recognizance release is immediately granted to misdemeanants who score at least 5 points and to third degree felons who score at least 7 points. Grant guidelines prohibit release of public intoxicants arrested within the city limits, and the courts prohibit release of defendants held for nonjudicial reasons, including military and immigration charges and agency "holds." Those not qualifying for immediate release are informed of the pretrial services program and given the opportunity to make bail following booking.

Pretrial supervised release counselors may interview those not released on ROR or bail. This interview follows a request by a judge or by a PTS counselor.

A recommendation is then made to pretrial services based on the verified ROR questionnaire and the counselor's subjective opinion of the prisoner. If Pretrial Services accepts a recommendation for release under supervision, the county attorney is given written notice at least 24 hours before a bonding hearing. Then, if the recommendation is not opposed, pretrial services outlines a tentative service program to the presiding judge at the hearing. If the judge approves, the prisoner is released to the custody of a pretrial services staff member. The prisoner remains under staff supervision pending the judicial disposition, receiving psychological testing and referrals to community agencies. At the time of sentencing, pretrial services prepares a "progress report" for the consideration of the presiding judge.

Considerably fewer than 65 people per month are released to the supervision of pretrial services. Several months into the program, about 18 people per month were so released, and during the last quarter of 1975, pretrial services supervised 113 cases (65 were felonies).

Misdemeanant probation is now an adjunct to the Utah State Adult Probation and Parole Department. It does not function like its counterpart in Des Moines. In Des Moines, the probation component supervises parolees from Fort Des Moines (the rehabilitation facility) as well as defendants granted probation. In Salt Lake, the Misdemeanant Probation Program does not deliver posttrial services beyond those delivered by state probation and parole.

The residential facility receives referrals from attorneys, probation officers, friends of the defendant, jail officials, and, infrequently, judges. One of two counselors interviews prospective inmates for an initial screening. Candidates must: (1) have at least 1 month of their jail sentence remaining or, (2) if not already serving a custody sentence, be convicted of a nonviolent offense. Eligible subjects who pass screening are recommended for admission. The judge presiding at sentencing then determines whether or not placement is in order.

The rehabilitation facility has no security devices, and all subjects have private rooms. The average time of residency is 60 days. The facility serves as an alternative to jail confinement and offers its clients specialized treatment and counseling services. The facility is primarily an educational and work release center.

## VI. COMMUNITY IMPACT

Political observers in Salt Lake argue that the Des Moines project does not have a strong philosophical base in the community. They claim cultural norms in Salt Lake favor retribution over rehabilitation. Project proponents do not favor public debate and tend to emphasize the need to reduce jail overcrowding rather than the provision of special services.

The project received considerable attention in the press and electronic media. Attention was generally focused on relieving jail pressures, less frequently on rehabilitation. In January 1975, 17.4% of the community had heard of the project and another 9.0% thought they may have heard. In January 1976, 20.0% of the community had heard of the project.

Our data suggest that the formal media coverage did not highlight the service functions of the Des Moines project. Informal contacts among friends, neighbors, relatives, spouse, and work associates, however, did include discussion of these functions. Most of those who discussed the project were recipients of services or acquaintances of recipients (directly or indirectly), and these people (approximately 41% of those who heard about the replication) were considerably more favorable toward the project than the community as a whole.

A majority of the negative comments address the "rehabilitative" function of the project. There is a sharp drop in positive comments from 1975 to 1976, and there is a sharp increase in qualifying statements like "it depends upon the crime or criminal."

Comments provided on project goals are particularly illuminating. Approximately 10% of the community, many of whom were recipients or acquaintances of recipients, favored the service function. About 12% disapproved the service function, and another 10% indicated that services should be restricted to certain types of offenders.

The project may have had an indirect impact on community. Community attitudes toward the police remained relatively consistent from 1975 to 1976, but community attitudes toward the courts significantly improved during this period. This result may not be directly related to the project, however, as in January 1975 several people chose to withhold comment on the courts due to bitterness over a mass murder case in another Utah city. Attitudes toward probation improved from 1975 to 1976, particularly among probation recipients and their acquaintances. There was no appreciable change in the community's attitude toward crime. Consistent with the unusually high crime rate, Salt Lake City residents considered crime a serious problem.

## VII. ORGANIZATIONAL IMPACT

The Des Moines replication is an innovative project in Salt Lake, but it appears to have brought minimal structural changes to established criminal justice agencies. Pretrial release was incorporated under the project umbrella, but the State Board of Corrections maintained control over local probation. Still, the project did affect changes in the Salt Lake County criminal justice system, and these changes were reflected in our organizational questionnaire.

All criminal justice agencies agree that the Court Services Project received considerable support from the courts, probation ("in-kind" services), and local government. All agencies appeared to share information with the replication project, although law enforcement was somewhat less willing when the project was implemented. All agencies reported that contacts were productive; and law enforcement, the courts, and community corrections show further improvement in productive contacts from 1975 to 1976.

Initially, law enforcement and community corrections had difficult times getting information from project personnel, but this dramatically improved during the first year of operation. Similarly, law enforcement and community corrections questioned their ability to influence project decisions in 1975, but their assessment improved by 1976. In 1976, law enforcement and community corrections felt that the project staff kept agreements, a distinct improvement

from 1975. Overall, attitudes of law enforcement officials toward the project greatly improved during the first year of operation.

Unlike the early relations with law enforcement, project relations with the courts were more comfortable in 1975. From 1975 to 1976, contacts between the courts and the project became more procedural and more productive. Throughout the year we monitored the project, the city court personnel were highly supportive of the project, and the courts continued to assist the project. (The justices' of the peace were not favorably disposed.)

Tensions between community corrections and the other court service programs eased from 1975 to 1976. Community corrections personnel felt that local government became more supportive, found it easier to get information from other project components, and felt that contacts with the project became less guarded. During the first year, contacts between community corrections and the other project components became more regular and considerably more procedural. (The improvement in procedural contacts is relative, for the community corrections personnel did feel that contacts were not sufficiently procedural.)

By the end of 1975, the project had definitely improved its relations with other criminal justice agencies. The early antagonisms between law enforcement and project personnel seemed to subside. Still, state efforts to isolate the community corrections facility from the rest of the project did remain a stumbling block in January 1976. All told, the project appeared to settle into the criminal justice system in Salt Lake County during its first year (January 1975 to January 1976).

#### VIII. CLIENT IMPACT

Every community generates a client flow into criminal justice that generally differs from the client flow in other communities. These differences result from: (1) varying demographic characteristics; (2) different cultural factors; (3) different statutory provisions; and (4) different emphases in law enforcement. The client impact of a Des Moines replication project centers upon the types of clients coming into the criminal justice system.

Relative to the general Salt Lake client flows, several factors distinguished those defendants recommended for ROR and SR. ROR and SR recommendations tended to favor young Caucasians. Clients who live alone were also over represented, while American Indians and defendants living with parents were under represented.

Compared to the SR recommendations, ROR recommendations (i.e., release without supervision) favored defendants living with a spouse and children, those from upper occupational strata, those with stable residential patterns, those who appeared less likely to change jobs, those with higher incomes, those not likely to be unemployed, and those with higher educational achievements. Release on recognizance recommendations tended toward clients with favorable social and economic circumstances, when compared to the total client profile and when compared to SR recommendations.

Compared to ROR, SR was more typically recommended for those living alone, those from lower occupational strata, those with less stable residential and employment characteristics, those who were poorer, and those who were less

educated. Compared to total client profiles and the ROR recommendations, SR tended toward the socially and economically disadvantaged.

Over 60% of criminal justice clients in Salt Lake City are young adults (under 30), most are unmarried, and most have highly unstable residential and employment patterns. These young adults are most likely found at the lower levels of the occupational strata. Although most have some form of employment, their income is considerably lower than the city as a whole. These young adults are not well integrated into the social structure of Salt Lake County, and the SR program tended to impact heavily on this particular group.

Some additional comparisons highlight this point. Of those living in the community less than 4 months, 40% were recommended for SR compared with only 16% of those living in Salt Lake more than 1 year. Of those unemployed, 34% were recommended for SR, compared with 27% of the intermittently employed and only 9% of the steadily employed. Of those with no income, 54% were referred to SR, compared with 28% of those citing other sources of income and 18% of those deriving income from their own employment.

Recommendations for OR release include disproportionately more charges for prostitution, driving while intoxicated, liquor law violations, disorderly conduct, driving without a license, and petty larceny. Burglary and public intoxication are noticeably under represented. (By policy, OR release is not immediately granted to public inebriates, although the courts can allow such release at arraignment.) Recommendations for SR include disproportionately more charges for aggravated assault, burglary, grand larceny, auto theft, forgery, traffic in stolen property, and vandalism. Public drunkenness and prostitution are noticeably under represented, and compared to ROR recommendations, recommendations for SR involve higher proportions of serious crime. (30% of all SR referrals are charged with burglary, 18% with narcotics violations, and 15% with auto theft.)

A sizable minority of all recommendations for OR release involve liquor offenses. Given the overcrowded jails, it would appear that OR release before arraignment could be profitably employed for certain inebriates.

Besides releasing suspects who were likely to remain in jail pending trial, the SR project delivered services to clients in need of rehabilitative help. Almost all (94%) received some services, and 75% received multiple services. (These services were primarily job training or placement and psychological counseling: 46% of the SR clients received employment counseling; 44% received vocational counseling; and 85% received some psychological counseling.) Psychological counseling and transportation services were provided directly by SR staff, while 30% of employment counseling and 95% of vocational counseling services were provided by other community agencies.

Most SR subjects do not become clients of the traditional service agencies in Salt Lake until they break the law. SR acts as a direct referral in these cases, identifying those needing services. Clients completing the program received an average of 26.6 hours of service, 17 of these hours being provided by staff personnel. Almost half of the clients completing the program (45%) secured new jobs. However, delivery of such services does not always prevent recidivism. Seven percent of the SR clients were found to have used drugs; 10%

used alcohol excessively; and 9% were rearrested--most for serious crimes. (Of 17 rearrested, 3 were charged with rape, 7 with burglary, 4 auto theft, and 1 fraud.)

Defendants placed on Supervised Release in Salt Lake have the more serious criminal charges and tend to have serious social problems. Under these circumstances, the SR record may be considered as not too unfavorable. About 30% of all SR subjects failed to complete the program, with a third of these failures charged with new crimes and a quarter found to be using drugs or alcohol excessively.

Commitment to the residential facility and placement on intensive probation are court-ordered sentences. Like supervised release, both the residential facility and intensive probation seek to provide services for their clients. About half of those sentenced to the residential facility were convicted of lesser crimes (e.g. driving while intoxicated and shoplifting.)

Of those sentenced to the residential facility, over three-quarters were unemployed. In addition, 78% were from the lowest occupational strata, although only 20% of the clients had families on public assistance. Most were young adult offenders, without families, with low educational achievement, a low or modest income, and few job skills.

Most (65%) received employment services and another 19% received vocational training. Most of the clients were also provided transportation to work. Employment counseling and transportation were primarily provided by facility staff, while other services were provided on a referral basis by other agencies. In the latter area, 63% of the residents received special education services, 60% psychological counseling, and 47% alcohol counseling.

The rehabilitation facility appears to have had a significant impact on its clients. Of the study sample of 98 clients, 75 were unemployed when they entered the facility and 43 of these were employed when they left. All those employed when entering the program were employed when leaving. By termination, 5 part-time students were registered full time, and 13 other clients became students. At termination, only 10 subjects were on public assistance. However, there was no appreciable improvement in job skills. In fact, counselors reassessed job skills downward as they became more familiar with actual client capabilities.

Drug and alcohol violations on the part of residents were high. Some improvement in the use of alcohol was offset by a slight increase in the use of drugs. About 6% of the residents were charged with serious offenses while at the facility (two with burglary, one with larceny, and three for narcotic violations). An additional 17% were charged with various misdemeanors. (Clients of the residential facility appear to have lower recidivism rates than those released from prison or jail. Considering only the more serious offenses, 14% of residential facility clients were accused of a new serious crime within a year from their referral, compared with 30% of jail or prison releases, and 21% of probationers over the same period of time.)

Unlike the residential facility, intensive probation never became an important sentencing alternative in Salt Lake County. As the intensive probation

staff did not provide adequate information about their program, this study sample is small (25 clients) and interpretation must be limited.

Unlike the supervised release and residential programs, 70% of intensive probation clients were charged with less serious offenses. Probationers were slightly older than SR or rehabilitation facility subjects and also had a higher proportion of Caucasians. Many of the probationers were unemployed. Services delivered to intensive probationers were also noticeably less than those delivered to residents of the rehabilitation facility. Still, 44% received employment related services, 24% vocational counseling, 24% psychological counseling, and 32% alcohol counseling. Approximately 25% of the probationers received no services and 68% were provided with more than one service.

At least 16% of the probationers were rearrested while under supervision. This rate compares favorably with the residential subjects and those released from jail or prison. However, intensive probationers appear to be "safer risks" in that they were sentenced for relatively minor offenses.

#### IX. SYSTEM IMPACT

The Des Moines replication effort in Salt Lake seems to have altered client flows through the criminal justice system. Points of impact include dispositions before arraignment, after arraignment, after trial, and after sentencing. The Des Moines programs may also have affected judicial decisionmaking at the points of arraignment, trial, and sentencing. In addition, the programs appear to have altered failure-to-appear rates and recidivism.

Case flows and jail status for more serious offenders were analyzed for a 40-week period between November 24, 1974 and September 14, 1975. Only cases from city criminal courts are included in the analysis (J.P. and traffic are excluded).

The first potential impact of the Des Moines programs concerns client disposition before arraignment. This impact is dependent upon the ROR program inasmuch as bail commissioners have authority to grant ROR release to all misdemeanants with the exception of public inebriates.

Approximately 10% of all defendants were released on their own recognizance (prior to arraignment) during the first 13 weeks of the grant. Eventually about 16% were released on their own recognizance prior to arraignment. Of the 6% gained over 41 weeks, many came from the bail release category (about 2%), but a significant number came from those otherwise remaining in jail. The 6% gain was inflated by the bondsmen strike. During the strike the ROR program was expanded, although later it was somewhat reduced.

The proportion of dismissals before arraignment increased from 11% to 18% during the time period monitored. Most of these dismissals involved first offense public inebriates. City judges informally agreed to release public inebriates after booking, providing the inebriate had not faced a similar charge within the past year. During the bondsmen strike, this proportion rose a couple percentage points, thus freeing more jail space for those not released on ROR. (The jail proportion jumped, then declined, during the strike. After the strike there was a further reduction in jail population for a few weeks while the bondsmen made up for lost time.)

Excluding dismissals before arraignment, the total proportion of suspects kept in jail until arraignment never fell below two-thirds. There was also very little change in the proportion of clients remaining in jail for more than 1 day before arraignment. (A significant minority were public inebriates with prior arrests for intoxication.) However, there was a significant change in the average number of pre-arraignment jail days during the 41-week study period: .96 to .77 (significant .05). Some of this decline occurred following the bondsmen strike.

Before arraignment, the ROR program had a greater impact on bail release than it had on jail retention. Initially, the ROR subjects tended to be the better bail risks, but over time the program was cautiously extended to include higher risk cases, including those making bond through personal indebtedness and those considered bail risks. Its impact on these groups, however, was relatively slight. Two factors that limited the impact of ROR on pre-arraignment jail proportions were: limited authority to release defendants before arraignment and the exclusion of public inebriates.

Comparisons of April 1974 data with those of April 1975 highlight the long run impact of the new ROR program. (OR at arraignment existed in Salt Lake before the Des Moines replication.) In April 1974, 44% of court defendants were released on bail while 16% were freed on personal recognizance. In April 1975, after the pre-arraignment ROR program was introduced, 33% were released on bail and 29% on recognizance. The increase in ROR releases consisted of those released on bail (11%) and those who might otherwise have remained in jail (2%). In short, the ROR program impacted heavily on the traditional bail release group and only had slight impact on those ordinarily remaining in jail.

The release rates are specific to the crime charged and the accused offender's past record. For example, felons are virtually never released prior to arraignment. For prostitution/commercialized vice, 66% were released--two out of every three released were on personal recognizance (half of the remaining cases were never arraigned). The release rate was also high for narcotics violations (54%) but here almost three of every four releases were by bond. Considering suspects charged with public intoxication, almost half are dismissed prior to arraignment. Most of the remainder are detained in jail, although a minority are released on bail. Finally, about one-third of shoplifters are released on recognizance; one-sixth are released on bail. Overall, release rates are crime specific.

On the other hand, the ROR program did little to reduce the traditional biases of bail against the socially and economically disadvantaged. An unemployed misdemeanor was 17% more likely to remain in jail than a man with a \$5,000 yearly income, and 28% more likely than a man with a \$10,000 yearly income. (The overall likelihood of a misdemeanor remaining in jail is 64%; hence these figures are quite significant.)

ROR clients were considerably more advantaged than those remaining in jail and were also usually more advantaged than those released on bail before arraignment. The bias appears to rest upon the Vera-Manhattan scale used to determine ROR eligibility and the current handling of public intoxication cases. This scale is not likely to reduce differential treatment before arraignment, given its emphasis on strong community ties.

The Des Moines project had greater and more favorable impact after arraignment. Judges may continue ROR, grant ROR, continue or grant bail, grant SR, or remand to jail at arraignment. The proportion of defendants receiving ROR at arraignment increased from 12% to 22% during the 41-week study period. SR accepted about 7% of all defendants at arraignment, taking most of its clients from those who would otherwise remain in jail. In addition, over the 41 weeks monitored, the total proportion of subjects on SR increased from 5% to 9%. The proportion of suspects in jail fell from 56% to 43%.

Over the time period examined, judges made earnest attempts to reduce jail populations by shifting some of those who would have remained in jail into SR and onto bail. Some of those who would normally receive bail release were in turn granted ROR. Second, the bondsmen strike made judges less cautious about using the SR program, which expanded by about 50% during the strike.

An examination of individual level data suggests that SR compensates for some of the biases found in the ROR program. Those ineligible for ROR before arraignment (and likely to have difficulty securing bail) do have an alternative to incarceration at arraignment. Hence the Supervised Pretrial Release Program does address the differential release status of Salt Lake City offenders. SR tends to serve recent arrivals to the community, those with prior adult convictions, the unemployed, and those without a source of income.

In contrast, ROR recommendations concentrate among "safer risks." ROR seeks to release people who are not likely to commit a crime while awaiting trial, and people who are likely to appear for scheduled court dates. To effect this policy, ROR release recommendations are concentrated among the more stable residents who are employed and without prior convictions. (A man with one prior arrest is 9% more likely to remain in jail, and 7% less likely to receive ROR. These figures are relatively high because the likelihood of receiving ROR is only 17% for all clients combined.)

The Des Moines project also reduced failure to appear rates in Salt Lake City. This appears to have been due, at least partly, to ROR interviewers supplying better defendant information to the judges and the SR staff closely monitoring the appearance of their SR subjects. Table 1 summarizes the failure to appear rates for misdemeanants before arraignment, after arraignment, and felons after arraignment. The proportion of defendants actually charged with failure to appear, or convicted of failure to appear, is considerably smaller than the failure to appear rates. Note that suspects released on bail are more likely to be charged with FTA, an indication of a higher rate of willful failures to appear among defendants released on bail.

Program reports from Salt Lake City indicate that 2% of the ROR subjects and 16% of the SR subjects were rearrested during the pretrial period. Our data support these figures and indicate that the rearrest rates tended to rise for both ROR and SR subjects over the course of the 41-week followup period. Part of this slight increase was due to the expanded use of the programs, and part was due to program changes resulting from the bondsmen strike.

The SR program, and to a lesser extent the ROR program, reduced the probability that a defendant would remain in jail after arraignment. However, the programs were not entirely successful in erasing the differential impact on offenders. Among misdemeanants, for example, SR received a poorer clientele than

those remaining in jail. (ROR received more economically advantaged clientele, followed by bail.) Unemployed misdemeanants are more likely to remain in jail and are less apt to receive ROR. Recidivists facing misdemeanor charges are also more likely to remain in jail and are less likely to receive SR. In the case of felony charges, the economically advantaged are more likely to be released on bail, and the economically disadvantaged are more likely to be placed on SR. Felons placed on SR tend to be unemployed, while those receiving bail tend not to be unemployed. Again, recidivists are more likely to remain in jail.

TABLE 1

	FTA	SR	BAIL
Misdemeanants prearraignment			
FTA	6%	N/A	15%
Companion Ch.	1%	N/A	6%
Conv. Comp. Ch.	1%	N/A	4%
Misdemeanants postarraignment			
FTA	15%	18%	18%
Companion Ch.	1%	0%	1%
Conv. Comp. Ch.	1%	0%	1%
Felony postarraignment			
FTA	15%	10%	16%
Companion Ch.	1%	2%	1%
Conv. Comp. Ch.	1%	2%	1%

We found evidence in Salt Lake that the Des Moines programs affected the client's likelihood of pleading guilty or demanding trial. Economic factors also appear to be unrelated to this decision.

The Des Moines programs (as well as certain social and economic factors) also appear to have had some effect on sentencing decisions. The most important determinants of sentence are charge and prior criminal history. For example, incarceration following conviction varies widely: burglary (25%), narcotics (12%), public intoxication (80%), shoplifting (22%), and prostitution/commercialized vice (8%) as does a sentence for probation. Across all crime categories examined, recidivists are more likely to receive a jail term (for narcotics, the likelihood of going to jail is increased by 12%, for larceny 28%, and for drunkenness 11%).

Our evidence indicates that those who remain in jail are more likely to receive jail sentences, independent of charge or criminal history, since those remaining in jail tend to be more disadvantaged than those released on bail or ROR (and in some instances, SR), the social and economic release bias carries into the sentencing decision. There are at least three possible explanations for the program impact at sentencing. First, the ROR interviewers provide

judges information that allows some "filtering" of clients at arraignment. Hence the release status at arraignment may be an indicator of things to come. Second, the SR referrals are sometimes credited for performance in service programs while awaiting adjudication. And, third, some judges may use release status as an indicator of a defendant's general overall status, i.e., compared to other defendants facing the same charge who are released before adjudication.

Finally, we examined the possible impact of the Des Moines programs on recidivism. Our sample for recidivism contained 261 cases collected 1 year after the close of our tracking data. Recidivist rates were generally high for all sentencing categories, except intensive probation (37%) and fines (43%). Of the three forms of incarceration, those sentenced to jail had the highest recidivist rate (77%). The residential facility followed with 65%, and the prison releases had a rate of 58%. The recidivism for the residential facility was quite high, given that the community corrections staff were sometimes quite selective in their recommendations.

However, if only subsequent serious crime (felonies) are examined, the recidivism picture changes. Here the residential facility fares better than prison or jail. Intensive probation falls between those incarcerated and those receiving regular probation or a fine. The intensive probation subjects tend to be the more serious offenders, including those released from the rehabilitation facility and those who received SR before adjudication. Hence this figure (16%) is also promising.

#### X. COSTS AND BENEFITS

When the Des Moines replication was proposed in Salt Lake County, its proponents argued that it would be cost effective. By releasing clients from jail, cost savings were expected to accrue both to the county government as well as to defendants whose alternative was incarceration. The former was expected to benefit as there would be less need to use the jail for pre- and posttrial confinement. The latter would benefit by not being required to post bail and not suffering the opportunity costs associated with jail. In addition to these short run savings, it was argued that the replication project would reduce recidivism, and thereby decrease the necessity of jail and the criminal process in the long run.

Thus, the replication project was expected to be cost effective. An evaluation of the replication must attempt to assess whether this intention was actually realized. This estimation is especially crucial, since limited public resources may indicate that continuation of the Court Services Project may hinge on the demonstration of cost savings, or at least, a suitable return to the county's investment.

An estimation of cost effectiveness is necessary to the evaluation. Unfortunately, given the present state of the arts, precise cost estimates are impossible. It is necessary to settle for "ball park" estimates, and though these estimates are not always as satisfactory as precise dollar figures, they are dictated by several considerations. First, cost analysis depends crucially on an accurate assessment of program impact, e.g., the number of jail days saved, the number of criminals rehabilitated, etc. However, the estimates of program impact that have been presented here are subject to errors, especially when the estimates are extended beyond criminal misdemeanants and felons to

justice of the peace and traffic cases.\* Additionally, many costs are subjectively measured. The cost to a defendant of remaining in jail is one example. The best that the evaluation team can do with such estimates is to indicate how they were calculated and allow the reader to make adjustments as he sees fit.

With these two caveats in mind, this chapter now turns to cost estimation. There are several savings that can be attributed to the Des Moines Replication:

1. savings in the reduction of prearraignment jail usage
2. savings in the reduction of postarraignment jail usage
3. savings to the defendant in the following:
  - a. reduced requirement to post bail
  - b. fewer days spent in jail pending trial and the associated lost income
4. savings in the reduced use of jail for a correctional alternative

$$S_1 = \$2,297$$

$$S_2 = \$13,502$$

$$S_3 = \$67,800$$

$$S_4 = \$121,545$$

These calculations are subject to two assumptions. First, the numbers used compare release rates at the beginning of the evaluation period with that at the end. The cost savings assume no effective transitional period, thereby over-emphasizing the actual savings. Second, the calculations ignore the fact that the project was in operation prior to the evaluation (the ROR component). Thus, the incremental savings may not reflect the dollar savings from starting the program from "scratch." According to the county auditor, the rate of releases increased even prior to the project.

	Bail	ROR	Releases
January/February 1974	-Δ325	+Δ454	+Δ129
January/February 1975			

\*The sample from justice of the peace and traffic courts indicated a much weaker trend than did the evidence from city and county courts. In fact, it is reasonable to suppose that the replication had very little additional impact on these cases. More importantly, use of the jail is subject to vagaries that cannot be identified, let alone controlled. It is just a presumption that observed changes can be attributed to the Des Moines Project.

If this increase in releases is indicative of decreased jail use prior to the project, the reduction in jail use is 774 suspects beyond that estimated above. The county auditor did not give any information about how this total should be allocated between prearrest and postarrest incarceration. But if we assume 10 days per defendant, then this would generate an additional \$50,000 in savings ( $S_1 + S_2$ ).

As crude as these estimates are, they indicate that the replication project was not cost effective if we calculate cost savings as the sum of  $S_1$ ,  $S_2$ , and  $S_4$ . However, it is evident that the savings covered a significant proportion of the cost of the grant, and perhaps reduced the necessity of constructing a new jail. In addition, there is evidence that the quality of justice was improved, that the delivery of services was enhanced, that recidivism may have been reduced, and that real costs and opportunity costs were reduced for defendants. These savings were significant and cannot be ignored by cost analysis.

In addition, evaluating the cost of a jail day at \$6.66 appears low. Evaluating jail days at closer to \$8.00 per day would yield cost savings approximately equal to incremental program costs. Finally, whether these results were "worth" the expense cannot be determined objectively.

#### SUMMARY

Salt Lake City has both a crime problem and an overcrowded jail. However, the political clientele of Salt Lake does not favor experimental programs such as the Des Moines project. The program was successfully introduced, but primarily due to the Federal monetary incentive. During its first year of operation, the project gained some support among the disadvantaged neighborhoods of Salt Lake and the project did introduce both formal and informal changes to the criminal justice system. The impact of Supervised Release was particularly impressive, and the delivery of services after adjudication was noteworthy. The project also helped to introduce a number of changes in pretrial release procedures with ROR impacting heavily on bail release and SR impacting on many who might otherwise have remained in jail. In addition, the project's impact on both failure to appear rates and recidivism shows some promise. Finally, the programs have returned some benefits to the county in terms of jail day monetary costs.

SECTION IV  
EVALUATING COMMUNITY PROGRAMS

Preceding page blank

## EVALUATING CITIZEN CRIME PREVENTION PROGRAMS<sup>1</sup>

Robert K. Yin  
The Rand Corporation  
Washington, D.C. 20037

### THE POTENTIAL IMPORTANCE OF CITIZEN CRIME PREVENTION EFFORTS

From a policy perspective, there are four broad approaches<sup>2</sup> for reducing the vulnerability of residential areas to crime:

- Paid public policing activities, in which individuals--i.e., police officers--are supported at public expense and are specifically sworn to enforce the law;
- Environmental design, in which planners and builders incorporate public safety concerns into the design of residential areas and new housing;
- Private-minded actions, in which citizens act to protect their private domains (their persons or their homes)--e.g., through the purchase of security devices and alarms or through training in the martial arts; and
- Public-minded actions, in which citizens act to protect public domains (public areas or the public interest)--e.g., through the development of neighborhood norms for behavior in public places and through surveillance activities conducted by residents.

On the whole, there is still little definitive evidence concerning the relative effectiveness of these four approaches or their combinations. Recent research, however, suggests that effective crime prevention--to the point that a neighborhood or set of residences can be said to be "safe"--may require some component of the fourth approach.

As for the effectiveness of the first approach, Wilson (1975), for instance, reviewed several key studies of the effects of police preventive patrol and concluded that, with the possible exception of saturation patrol (a condition that may be socially undesirable and fiscally impossible for long-term application in any neighborhood), there was no clear evidence that increases in preventive patrol alone led to a reduction in crime.<sup>3</sup> This conclusion was based mainly on several post-hoc studies (e.g., Press, 1971) as well as on the results of an actual field experiment--the Kansas City Preventive Patrol experiment (Kelling et al., 1974). Levine (1975), in conducting his own analysis of reported robbery and murder as a function

of different levels of police manpower in 26 cities, also found no relationship between the amount of police and crime prevention.

As for the second approach, there has been preliminary evidence of some effective measures that can be taken (e.g., Newman, 1972; and Reppetto, 1976). However, it should be noted that much of the crime prevention effect occurs through behavioral changes among residents, who become more vigilant and develop a greater sense of territoriality. Thus, the appropriate environmental designs do not work by themselves but depend upon the type of activity that is characteristic of the fourth approach below.

As for the third approach, it is well known that residents who are fearful of crime take many steps to make their homes and daily routine more secure, e.g., the purchase of locks, alarms, and window bars. There is a limit, however, to the effectiveness of crime prevention through precautions of this sort. Heller et al. (1975) have helped to identify some of the difficulties with private-minded actions by citizens. In their review of Operation-Ident projects, which require residents to mark their valuables in order to facilitate recovery if stolen, a key finding was that the enrollment rate for such efforts was very low (around 10 percent). Thus, although those who participated in the program appeared to be better protected from crime, only a minimal portion of the citizenry was involved in the program. Further, the authors found it difficult to identify ways of increasing the enrollment rate. Private-minded actions may therefore always suffer from a low participation rate, and even though they may be effective in theory, difficulties in implementation will pose a chronic problem. Although low participation rates may also characterize other crime prevention activities, the low rates are particularly debilitating for any strategy based on private-minded actions because, by definition, everyone must participate in such actions in order to establish full coverage (see also Schneider and Eagle, 1975).

In contrast, the admittedly rudimentary evidence concerning the potential importance of public-minded citizen crime prevention activities--in which residents take an interest in each other's activities and therefore use their own eyes and ears to monitor their neighborhood--has led to increased interest on the part of policymakers in exploring new activities by citizens (Washnis, 1976). Public-minded activities may be carried out in a number of ways--e.g., the formation of a resident patrol (Yin et al., 1976) or the establishment of various citizen crime reporting systems (Bickman et al., 1976). Such activities may also be one component of a mixed approach. Reppetto (1974), for instance, conducted a study of residential crime, including interviews of offenders involved in burglary cases. He concluded with the suggestion that the most fruitful course of future action might be the development of a crime prevention approach that would (1974, p. 87):

" . . . blend the deterrent effects of the criminal justice system with citizens' anticrime efforts . . . It is possible, for example, that the "rapid response" techniques of the police could become a more meaningful deterrent to residential crime if environmental characteristics could be modified to maximize surveillance possibilities and encourage a sense of territorial concern among residents; citizens would take a

few more precautions aimed at 'slowing down' the prospective burglar so that his suspicious activities might attract the attention of neighbors; and observing neighbors might feel a 'social commitment' sufficient to prompt them to summon the police."

Thus, the provisional findings from the public policy perspective point toward the need to assess resident-based, public-minded actions in promoting residential crime prevention.<sup>4</sup> In addition, it may also be argued that such actions may lead to the further cohesiveness of a neighborhood, in which residents also help each other in dealing with other everyday functions, such as child care and supervision, shopping, schooling, and emergency assistance. In contrast, neighborhood interactions and cohesion might even diminish as a result of two of the other approaches to crime prevention, i.e., increasing preventive patrol by police or stimulating private-minded actions by citizens to protect themselves and their homes.

#### WHY THE EVALUATION OF CITIZEN CRIME PREVENTION EFFORTS IS DIFFICULT

At the same time, citizen crime prevention efforts present some of the most difficult circumstances under which a policy evaluation must take place. Normally, the evaluation of a public policy intervention must surmount five general obstacles:

- The identification of measurable objectives;
- The identification of a target population;
- Control over the intervention program, so that it can be applied or withheld according to a specific research design (e.g., experimental vs. control groups);
- The ability to measure the key features of the intervention process; and
- The availability of sufficient time so that the short- and long-term effects of the intervention can be assessed.

These obstacles have only occasionally been completely surmounted in existing evaluations of such programs as a Head Start program, a manpower program, or a drug rehabilitation program (e.g., methadone maintenance). For instance, a common problem is that members of the target population may not be easily assigned on a random basis to experimental and control groups (Boruch, 1976). This and many other methodological problems have been adequately described by the existing literature on evaluation research (e.g., Campbell and Stanley, 1963; Suchman, 1967; Rossi and Williams, 1972; Weiss, 1972; Caporaso and Roos, 1973; and Bernstein, 1976).

Any evaluation of crime control programs initiated by law enforcement agencies must face these five general obstacles as well as others. Regarding the five obstacles, the difficulties of evaluating crime control programs are well known (e.g., Maltz, 1972; and Chaiken, 1976): (1) the identification of measurable objectives usually calls for the use of "crimes reported to the police," which is a highly inadequate data base; (2) the

target population is usually a geographic area as well as a set of individuals, creating difficulties in selecting control or comparison groups; (3) and (4) the intervention process is usually controllable and measurable but may have unavoidable complications<sup>5</sup>; and (5) the pressure to produce results usually means that the long-term effects of a crime control program are ignored by the evaluation. However, evaluations of crime control programs are also confronted by three additional obstacles:

- The actual crime control objective is to prevent an event from occurring, and such an absence of events is difficult to assess;
- The full assessment of a crime control program requires the measurement of possible displacement effects--to different geographic areas,<sup>6</sup> to different times of the day, or to different types of crime; and
- Within most realistic ranges of activity, any single crime control program that is the subject of evaluation may be expected to have only a weak (and hence more difficult to measure) effect on the incidence of crime, and such effects may also not be readily separable from the effects of other crime control activities.

Difficult as these conditions appear, evaluations of citizen crime prevention programs face all of these as well as two other obstacles:

- By definition, many citizen activities are voluntary acts and hence cannot be manipulated by policymakers as they might manipulate other intervention programs; and
- The most effective forms of citizen crime prevention activities--e.g., maintaining strong informal relationships with other residents or maintaining frequent informal surveillance over behavior in public places--may be the most difficult to measure. It would probably be extremely difficult to determine, for instance, when such informal patterns first emerged and hence when the "treatment" actually began.

The full combination of all these obstacles often leads to awkward dilemmas in evaluating citizen crime prevention efforts. For instance, if the number of crimes reported to the police is the only outcome measure available, it cannot even be hypothesized that an effective effort will lead to a reduction in reported crime; reported crime may very well increase, at least in the short-term, as a result of an effective program (Schneider, 1975; and Bickman et al., 1976). As a second example, police coverage is a factor that should be made constant in comparing two geographic areas, one with a citizen crime prevention activity and the other without; yet such coverage may also be affected by the very existence of such an activity (Yin et al., 1976).

This brief sketch of the evaluation problems to be resolved should suggest the enormous difficulties posed by trying to answer the question of whether a particular citizen crime prevention activity is effective or not. As a result, it is not surprising that attempts to evaluate such activities

have either culminated with an absence of conclusive findings or been described under such hedged conditions that the casual reader believes the activity was probably ineffective. Such a negative connotation is easily captured by the mass media, and further citizen voluntarism may even be somewhat reduced. This is a most unfortunate occurrence, because what is at fault is clearly our research methodology, and not necessarily the crime prevention activity. To the extent that methodological progress cannot be made, then it is incumbent on researchers and policymakers to state clearly the conditions under which evaluative questions may be fairly asked.

#### STUDY OF CITIZEN PATROLS

A recent study of citizen patrols attempted to deal with some of these evaluation problems (Yin et al., 1976). Our main objective here will be to identify the main possibilities for any evaluation of patrol activities. However, the basic findings of the study should also be summarized.

First, the study began with a fourfold definition of patrols. Such activities had to:

- Include a specific patrol or surveillance routine that was not just a part of another full-time activity such as driving a taxicab;
- Be aimed at preventing criminal acts;
- Be controlled by a citizens' or residents' organization or a public housing authority; and
- Be directed at residential rather than commercial areas.

Second, the study's main contribution to the state of knowledge was descriptive rather than evaluative. As a result of a research design that sampled cities across the country and personal or telephone interviews with over 100 such patrols, the study concluded that:<sup>7</sup>

- There are about 800-900 patrols in the U.S., lasting an average of 4-5 years;
- Patrols emerge in a variety of neighborhoods at all income levels and racial mixes (including areas in which residents wish to preserve a previously crime-free environment);
- For most patrols (over two-thirds were operated by volunteers), voluntary efforts and contributions comprise the main resources, with few cases of funding from any public agency such as the U.S. Law Enforcement Assistance Administration; and
- Patrol operations are facilitated if a patrol is affiliated with a larger community or neighborhood organization; if some resources are used to support rudimentary bureaucratic procedures such as maintaining formal membership lists, schedules, and substitution procedures; if a patrol has a cooperative relationship (where relevant--see below) with the local police; and if a patrol

allows for a flexible range of activities so that even noncrime prevention activities can be pursued when patrolling appears unnecessary (with the patrol capability being maintained, however, should patrolling subsequently become necessary).

Put simply, the study found that citizen patrols had become part of the normal repertoire of residents concerned with crime prevention, and that these contemporary patrols are far different from the riot-related patrols that have dominated earlier research (e.g., Marx and Archer, 1972). To the extent that the patrols have positive effects, they represent an important additional resource because of their low cost.

The study also developed a typology of patrols. For the purpose of the present paper, two of these types are especially relevant: building patrols,<sup>8</sup> in which volunteers or paid guards maintain surveillance over and neighborhood patrols, in which volunteers or paid guards maintain surveillance over a small geographic area, usually by driving around in a car and maintaining radio communication with a base station (e.g., one of the resident's homes). These two types of patrols appeared to develop different organizational histories and to produce different outcomes. For instance, building patrols are easier to staff on a 24-hour basis, because a much more limited area is involved and can be covered with fewer personnel than is the case with a neighborhood patrol. As another example, building patrols are simpler to operate than neighborhood patrols because there need be little or no interaction with the local police, who do not usually cover specific buildings and who therefore do not feel the need to know about (much less coordinate with) the activities of a building patrol. As a final example, the work of the building patrol may be made easier by the fact that such patrols can concentrate on screening strangers and keeping them off the premises; a neighborhood patrol can only focus on observed behaviors that appear suspicious, which is a more judgmental task that may lead to false alarms and unnecessary provocations.

#### EVALUATING CITIZEN PATROLS<sup>9</sup>

The distinction between these two types of patrols also serves as a starting point for differentiating between those evaluation questions that can and cannot be asked. To begin with, it is probably true that, at least in conventional terms, the effectiveness of a neighborhood patrol cannot be determined by research evaluations. This is because:

- The area covered and residents protected by a neighborhood patrol are too poorly defined--i.e., the selection of respondents for a victimization survey, for instance, would not precisely reflect the target population, and control or comparison areas would also be difficult to define;
- Any effects of the neighborhood patrol cannot be sufficiently distinguished from those related to police and other crime control activities--i.e., a neighborhood patrol is only a weak treatment, and conclusive evidence of its effect could probably only be established through an unrealistic experiment in which one condition was to have some neighborhoods with no crime control activities other than a neighborhood patrol; and

- The quality and quantity of local police activities could change as a result of the presence of a neighborhood patrol (e.g., police coverage might be reduced in the face of a neighborhood patrol that was perceived to be effective), and even if such changes could be measured accurately, the overall outcomes would be affected in a very complex manner.

However, even though the ultimate crime control effects of a neighborhood patrol might not be determinable, other important questions about the patrol might be asked. A policymaker might be interested, for instance, in the extent to which the residents of an area (even if it did not exactly coincide with the patrol area): (1) were aware of the patrol, (2) had encounters with the patrol, (3) felt any differently about safety in the neighborhood as a result of the patrol, or (4) made greater or lesser use of public areas or otherwise changed their crime prevention behavior because of the patrol. Such outcomes have become important because of a contemporary concern with citizen alienation (e.g., Yin and Lucas, 1973), even though it must be made clear that little would be known about crime deterrence.

In contrast, the effectiveness question can probably be addressed for building patrols, given certain limitations. Because building patrols serve a well defined population and do not usually compete with the police in providing preventive patrol, an evaluation could probably be established whereby similar buildings (or building projects) with and without patrols are compared. Victimization rates among the residents would have to be the major outcome measures. Such an evaluation could not, however, follow the classic pre-post, experimental-control design. This is because building patrols are a voluntary citizen activity that can neither be initiated at a time and place of the evaluator's choosing, nor prevented from forming at any preestablished control sites. Furthermore, as a voluntary activity that often emerges quickly and unexpectedly, there would probably be no time to collect baseline data, unless the evaluator had the uncommon luxury of having monitored the victimization rates at numerous sites for a period of time, in the hope that patrols would emerge at some meaningful subset of these sites.

One feasible evaluation design would be a matched, posttreatment design, in which sites with existing patrols are compared with comparable sites with no patrols. This design would obviously be limited by the nature of the matching procedure. In addition, there might be little record of the early patrol activities (i.e., before the evaluator came upon the scene), so that the exact nature of the intervention could not be established.

An alternative design could overcome these obstacles but would involve other tradeoffs. In this alternative, a public agency (e.g., local law enforcement agency, public housing authority, or Federal criminal justice agency) could initiate a program to assist residents in establishing new building patrols. The classical pre-post, experimental-control design could then be followed, given the stipulation that patrol activities could not begin until baseline data had been collected. The evaluation would clearly not be of a purely citizen-initiated activity, and strictly speaking, the results could only be generalized to other resident patrols willing to accept external funding under the same prespecified conditions.

However, there is reason to believe that such patrols might not differ substantially from truly citizen-initiated patrols, and that the evaluation results might therefore be of considerable value. Nevertheless, although such a design might make an evaluator more confident of his results, this alternative presents major risks:

- A resident group might be responding to a serious crime problem and might thus be unable to honor its intentions to await baseline data collection--there could be intense pressure on the evaluator to complete the job, residents' attitudes may have changed anyway with the knowledge that a patrol is about to begin, or informal patrolling could actually be initiated, even at the risk of losing the external funds, if the situation became sufficiently critical. Under any of these conditions, both the external funding agency and the evaluator would be forced into an embarrassing public relations position, because the purity of science would appear to be of higher priority than the public interest.
- Resident groups in the predesignated control sites could still initiate their own patrols and upset the research design.
- The costs of implementing and conducting the research project might be substantially greater than the funds disbursed to the patrol groups and might generate public criticism. For instance, a building patrol might satisfactorily operate for a year with a grant of less than \$10,000 from the external agency; the costs of a victimization survey at that site alone (i.e., not counting the implementation of the project, the support needed for the research design and preparation, or even the surveys to be done at the control sites) could be \$50/respondent for 100 respondents in each of at least two waves--or \$10,000 by itself.

In short, the decision to conduct a classical evaluation and to attempt to answer the effectiveness question must be weighed against the financial and possible political costs of doing the evaluation.

#### IMPLICATIONS FOR CITIZEN CRIME PREVENTION ACTIVITIES

The above examples should make clear that most evaluations of citizen crime prevention activities are likely to be severely limited. In the face of such limitations, it is extremely important that any negative connotations--which at present stem from the methodological state of the art--not be allowed to reflect unfairly on the actual citizen crime prevention activity. Voluntary efforts are difficult to initiate and yet are low in cost to public agencies; such efforts should probably be actively encouraged unless there is clear evidence of a serious negative outcome.

One task of the policymaker (e.g., local law enforcement official, mayor, or Federal agency official) is therefore to create a different climate for viewing citizen crime prevention activities--one in which the burden of evidence is placed on those who would like to show that the patrols are ineffective. Note that ineffectiveness measures are not merely the observation of null effects for effectiveness measures. For a citizen crime

prevention activity, ineffectiveness measures might include errors (e.g., false arrests, unnecessary injuries, etc.), low participation rates, complaints by participants, complaints by residents, or failure to gain cooperation from the police. Most of these ineffectiveness measures are more easily monitored and assessed than the typical array of effectiveness measures. Yet we know of no evaluation that has deliberately assumed the burden of demonstrating ineffectiveness.

Other tasks may be considered as well. First, methodological research should be encouraged so that citizen activities might be more accurately assessed in the future. Second, the ultimate limitations of research on this topic may still have to be recognized. Here, it is important to remember that evaluation research is only one mechanism (and a fairly unimportant one, at that) for setting public priorities. The political process, as reflected in voting behavior and the priorities of legislators and other elected officials, is in fact the more common way of setting public priorities, and Federal agency policymakers, for instance, should enthusiastically mount new programs mandated by the Congress<sup>10</sup> without unduly straddling the programs with evaluative questions that we have shown cannot be answered. The worthiness of a citizen crime prevention activity might thus have to be judged simply by input measures--e.g., how many participated (or were employed) at what cost, and how the activity was received by other residents and the police--as well as by the absence of compelling evidence concerning negative outcomes.<sup>11</sup> Finally, policymakers could encourage citizens themselves to make more intelligent choices. Agency-sponsored public information programs could assist citizens in different cities and neighborhoods to become more aware of the experiences of others. For instance, residents--who, incidentally, may have already decided on a course of action independent of any formal effectiveness evidence--could benefit by knowing about the implementation problems and solutions of other groups.

## NOTES

1. The author wishes to thank Dr. Jan Chaiken for helping to develop and refine many of the ideas in this paper.
2. The full array of crime prevention techniques is best summarized in Sagalyn et al. (1973). See also Wilson and Boland (1976) for a discussion of interventions at other points in the criminal justice system (e.g., a changed sentencing policy) that might affect residential crime.
3. Similar conclusions were also reached in a review by Chaiken (1976), who also points out, however, that practically any kind of increase in police manpower will produce increases in the number of arrests made by the police (where arrest rates are used as an alternative outcome measure).
4. Although it is always difficult to make comparisons between different countries, a recent article (Bayley, 1976) suggests that crime prevention associations exist in most neighborhoods in Japan and appear to be one reason for Japan's low crime rate. (The article makes a compelling case that most other conditions that might affect crime rates--e.g., a tradition of violence or lenient sentencing policies--are either similar or less favorable in Japan than in the United States.)
5. An excellent example may have occurred in the Kansas City Preventive Patrol Experiment, where patrol cars for the reactive beats--i.e., beats that supposedly received minimal preventive patrol--may have inadvertently produced a significant preventive effect by their longer and more visible drives to intervene in incidents (Larson, 1975). The point here is that the desired treatment--variation in preventive patrol but constancy in response to actual calls for assistance--may be difficult to implement without some contaminating factors.
6. The role of a rudimentary typology of neighborhoods is usually overlooked in research on geographic displacement. Such research (e.g., Press, 1971) usually examines the geographically contiguous areas for displacement effects, whereas it can be equally argued that if crime is suppressed in one neighborhood and is thereby displaced to another neighborhood, the target of the displaced activity will be a neighborhood of the same type as the first, not necessarily the next closest neighborhood.
7. See Yin et al. (1976) for a full elaboration of the research methods, conclusions, and recommendations of the study.
8. In a large city, a block patrol (usually given the specific assignment of covering two, four, or eight block faces) may be regarded in the same manner as a building patrol.

9. This section draws in part from portions of Yin et al. (1976) that were drafted by Dr. Jan Chaiken.
10. For example, Congress passed the Community Anti-Crime Assistance Act (S.3337) in 1976, which calls for the disbursement of \$15 million to citizen groups across the country.
11. It should be noted that, before the advent in the mid-1960's of the new concern with evaluation research, many social programs were initiated with just these requirements. However, we are not advocating that evaluation of effectiveness be dropped indiscriminately--only that certain activities that cannot be evaluated not be restrained unless there is negative evidence.

## BIBLIOGRAPHY

- Bayley, David H., "Learning about Crime--the Japanese Experience," The Public Interest, No. 44, Summer 1976, pp. 55-68.
- Bernstein, Irene N. (ed.), Validity Issues in Evaluative Research, Sage Publications, Beverly Hills, Calif., 1976.
- Bickman, Leonard, et al., Citizen Crime Reporting Projects, Loyola University of Chicago, Chicago, Ill., 1976.
- Boruch, Robert F., "On Common Contentions about Randomized Field Experiments," in Gene V. Glass (ed.), Evaluation Studies Review Annual, Vol. 1, 1976, pp. 158-194.
- Campbell, Donald T., and Julian C. Stanley, Experimental and Quasi-Experimental Designs for Research, Rand McNally, Chicago, Ill., 1963.
- Caporaso, James A., and Leslie L. Roos, Jr., Quasi-Experimental Approaches, Northwestern University Press, Evanston, Ill., 1973.
- Chaiken, Jan M., "What's Known about Deterrent Effects of Police Activities," The Rand Corporation, P-5735, paper presented at the annual meeting of the Operations Research Society of America, Miami, Fla., November 1976.
- Heller, Nelson, et al., Operation Identification Projects, U.S. Department of Justice, Washington, D.C., 1975.
- Kelling, George, et al., The Kansas City Preventive Patrol Experiment: A Summary Report, Police Foundation, Washington, D.C., 1974.
- Larson, Richard C., "What Happened to Patrol Operations in Kansas City? A Review of the Kansas City Preventive Patrol Experiment," Journal of Criminal Justice, Vol. 3, 1975, pp. 267-298.
- Levine, James P., "The Ineffectiveness of Adding Police to Prevent Crime," Public Policy, Vol. 23, Fall 1975, pp. 523-545.
- Maltz, Michael D., Evaluation of Crime Control Programs, U.S. Department of Justice, Washington, D.C., April 1972.
- Marx, Gary T., and Dane Archer, "Community Police: An Exploratory Inquiry," unpublished paper, Harvard-MIT Joint Center for Urban Studies, Cambridge, Mass., June 1972.
- Newman, Oscar, Defensible Space, Macmillan, New York, N.Y., 1972.
- Press, S. James, Some Effects of an Increase in Police Manpower in the 20th Precinct of New York City, The Rand Corporation, Santa Monica, Calif., R-704-NYC, 1971.

- Repetto, Thomas A., "Crime Prevention through Environmental Policy," American Behavioral Scientist, Vol. 20, November/December 1976, pp. 275-288.
- Repetto, Thomas A., Residential Crime, Ballinger, Cambridge, Mass., 1974.
- Rossi, Peter, and Walter Williams (eds.), Evaluating Social Programs, Seminar Press, New York, N.Y., 1972.
- Sagalyn, Arnold, et al., Residential Security, U.S. Department of Justice, Washington, D.C., 1973.
- Schneider, Anne L., "Evaluation of the Portland Neighborhood-Based Anti-Burglary Program," Oregon Research Institute, Eugene, Ore., March 1975.
- Schneider, Anne L., and Jerry Eagle, "The Effectiveness of Citizen Participation in Crime Prevention: A Random Outlaw Model," Oregon Research Institute, Eugene, Ore., 1975.
- Suchman, Edward, Evaluative Research, Russell Sage, New York, N.Y., 1967.
- Washnis, George J., Citizen Involvement in Crime Prevention, Praeger, New York, N.Y., 1976.
- Weiss, Carol (ed.), Evaluating Action Programs: Readings in Social Action and Education, Allyn and Bacon, Boston, Mass., 1972.
- Wilson, James Q., Thinking about Crime, Basic Books, New York, N.Y., 1975.
- Wilson, James Q., and Barbara Boland, "Crime," in William Gorham and Nathan Glazer (eds.), The Urban Predicament, The Urban Institute, Washington, D.C., 1976, pp. 179-230.
- Yin, Robert K., and William A. Lucas, "Decentralization and Alienation," Policy Sciences, Vol. 4, September 1973, pp. 327-336.
- Yin, Robert K., and Douglas Yates, Street-Level Governments: Assessing Decentralization and Urban Services, D. C. Heath and Company, Lexington, Mass., 1975.
- Yin, Robert K., Mary Vogel, Jan Chaiken, and Deborah Both, Patrolling the Neighborhood Beat: Residents and Residential Security, The Rand Corporation, Santa Monica, Calif., R-1196-HEW, March 1976.

AN EVALUATION OF THE SEATTLE COMMUNITY CRIME PREVENTION PROGRAM

Kenneth E. Mathews, Jr., Ph.D.  
Senior Researcher and Evaluator  
Seattle Law and Justice Planning Office

The Community Crime Prevention Program (CCPP) is an attempt to use citizen-based action as a burglary reduction strategy. The project consists of a number of community organizers who help neighborhood residents plan and carry out burglary prevention and reporting measures consisting of three primary activities. First, organizers offer to engrave an identifying number on certain types of easily stolen property. Second, the organizers offer to conduct a security inspection of the residence, looking at common points of entry by burglars and suggesting various improvements in security devices. Third, a system of block watch organizations is set up in which members of the neighborhood agree to watch their neighbors' residences and to report any suspicious circumstances or probable burglaries.

The basic hypothesis tested by CCPP activities is that the establishment of block watch organizations and the provision of home security inspections and property marking will significantly reduce the occurrence of residential burglary for those residences and areas receiving CCPP services, when compared to residences and areas not receiving these services.

During the first 34½ months of operation, CCPP services provided by the project were as follows: 7,630 home security inspections, 8,245 residences with property marked, and 8,034 block watch organizations. In addition, 28,962 information and education contacts were made. The estimated cost per service unit during this period (including education and information contacts) was \$7.14. If only primary services are considered, the cost per service unit was approximately \$18.39.

EVALUATION OBJECTIVES

The evaluation was performed to assess the following crime impact objectives.

Objective One. To produce a statistically significant decrease in the number of residential burglaries in the project-treated areas as compared with those areas before project operation and with other nontreated areas during similar time periods.

Objective Two. To produce a statistically significant increase in the number of burglary-in-progress calls received by the police department from citizens living in the target areas as compared with other nontreated areas of the city.

Preceding page blank

## EVALUATION MEASURES AND OUTCOMES

Objective One. Reduction of residential burglary (hereafter referred to as burglary) was evaluated using victimization survey data and official Seattle Police Department statistics on reported burglaries. Victimization data were obtained through three major sources: CCPP data, SEA-KING, and an LJPO telephone survey. CCPP data consist of project-collected victimization data. At the time a household joins the project, project staff ask participants whether they have been burglarized within the preceding 6 months (CCPP predata). Six months following program entry, participants are recontacted and asked if they have been burglarized since becoming involved in the project (CCPP postdata).

SEA-KING data consisted of two sets of in-person victimization surveys in West Seattle (Federal census tracts 96, 97, 98, and 105). The surveys were conducted under the supervision of the Seattle Law and Justice Planning Office. The first survey conducted in mid-1975 dealt with crime victimization in calendar year 1974 and interviewed 1,474 residences (SEA-KING pre-). The second survey conducted in mid-1976 inquired about crime victimization in calendar year 1975 within 937 residences (SEA-KING post-). Within all four census tracts, approximately half were re-interviews of residences interviewed in the prior year. Within the two treated tracts (97 and 98), half of the residences interviewed were randomly chosen on the basis that they had received CCPP services, while the remaining half had not received such services (e.g., had refused, had not been offered services, were unaware, had recently moved into the area).

The Seattle LJPO telephone survey was conducted in August and September 1976. The survey sought victimization data for the preceding 6 months from both program participants and nonparticipants ( $n = 3,292$ ) in five census tracts (Federal tracts 87, 89, 95, 97, and 98). The five tracts were chosen on the basis of being recently treated (having been completed no more than 18 months nor less than 6 months prior to August 1976) and having met CCPP criteria for successful treatment (i.e., 30 percent or more of potential single and duplex residences received burglary reduction services). Telephone numbers of program participants were obtained from CCPP records; an equivalent number of non-CCPP telephone numbers (excluding businesses and apartments within the census tract boundaries) were randomly selected from the Pacific Northwest Bell Street Address Directory for Seattle.

### VICTIMIZATION DATA

CCPP Data. Project victimization data indicate a significant decrease ( $p < .05$ ) in burglary victimization, from 4.46 burglaries per 100 households per 6 months prior to program entry to 2.34 burglaries per 100 households per 6 months. This represents 407 burglaries within 9,129 residences interviewed at program entry between September 1973 and December 1976, and 138 burglaries within 5,903 residences after 6 months of program involvement interviewed between February 1974 and June 1976 ( $\chi^2 = 46.13$ ,  $df = 1$ ,  $p < .001$ ).

SEA-KING Data. For the total areas (see table 1), when treated census tracts are compared pre- and post-CCPP treatment, there is a marginally significant decline in the burglary rate (-36.3 percent, from 6.34 burglaries

TABLE 1.--SEA-KING victimization data

Table 1--SEA-KING Victimization Data

Area	Burglary Victimization per 12 Months									
	Pre-Treatment (Jan.-Dec., 1974)					Post-Treatment (Jan.-Dec., 1975)				
	Total	Not Burglarized	Burglarized <sup>1</sup>	Reported		Total	Burglarized	Not Burglarized <sup>1</sup>	Reported	
Yes				No	Yes				No	
<u>Control</u> (federal tract 96 and 105)	575	515	60 (10.43%)	28 (47%)	32	442	380	42 (9.95%)	24 (57%)	18
<u>Experimental</u> (federal tract 97 and 98)										
--CCPP:	356	334	22 ( 6.18%)	15 (68%)	7	247	241	6 (2.43%)	6 (100%)	0
--Non-CCPP:	543	508	35 ( 6.45%)	14 (40%)	21	248	234	14 (5.65%)	7 ( 64%)	4 <sup>2</sup>
<b>Total</b>	<b>899</b>	<b>842</b>	<b>57 ( 6.34%)</b>	<b>29 (51%)</b>	<b>28</b>	<b>495</b>	<b>475</b>	<b>20 (4.04%)</b>	<b>13 ( 77%)</b>	<b>4<sup>2</sup></b>

<sup>1</sup>Burglarized one or more times

<sup>2</sup>Does not include three cases where reporting data were unknown

Burglary victimization per 12 months

Area	Pretreatment (Jan.-Dec. 1974)					Posttreatment (Jan.-Dec. 1975)				
	Total	Not burglarized	Burglarized <sup>1</sup>	Reported		Total	Burglarized	Not burglarized <sup>1</sup>	Reported	
				Yes	No				Yes	No
<u>Control</u> (Federal tract 96 and 105)	575	515	60 (10.43%)	28 (47%)	32	442	380	42 (9.95%)	24 (57%)	18
<u>Experimental</u> (Federal tract 97 and 98)										
CCPP	356	334	22 (6.18%)	15 (68%)	7	247	241	6 (2.43%)	6 (100%)	0
Non-CCPP	543	508	35 (6.45%)	14 (40%)	21	248	234	14 (5.65%)	7 (64%)	4 <sup>2</sup>
Total	899	842	57 (6.34%)	29 (51%)	28	495	475	20 (4.04%)	13 (77%)	4 <sup>2</sup>

<sup>1</sup>Burglarized one or more times.

<sup>2</sup>Does not include three cases where reporting data were unknown.

per 100 per year to 4.04 per 100 per year), with a corresponding marginally significant increase in reporting rates (52.1 percent, from 50.1 percent to 76.5 percent). During the same period, adjacent areas (tracts 96 and 105) experienced a nonsignificant and minimal change in burglary rates (-4.6 percent, from 10.43 burglaries per 100 per year to 9.95 per 100 per year), with a nonsignificant increase in reporting (22.3 percent, from 46.7 percent to 57.1 percent).

To determine the precise effect of CCPP on burglary, interviews conducted in the treated area for both 1974 and 1975 were separately tallied on the basis of whether interviewed households joined the project in 1975. Pretreatment data (1974) show that 356 (39.6 percent) of the 899 interviews conducted in the treatment area were of residences that were to join the project in 1975. Of these, 22 (6.18 percent) had been burglarized and 15 (68.2 percent) were reported to the police. Of the 508 residences that did not join CCPP in 1975, 35 (6.45 percent) had been burglarized and 14 (40.0 percent) were reported to the police. A comparison of pretreatment burglary rates for the two groups indicates that they were virtually identical ( $\chi^2 = 0.03$ ,  $df = 1$ ,  $p = .86$ ).

A comparison of reporting rates for the 1974 data indicates that those persons who eventually were to join CCPP had a significantly higher reporting rate ( $\chi^2 = 4.29$ ,  $df = 1$ ,  $p < .05$ ) than those who did not join.

A comparison of 1975 data for the two groups (CCPP and non-CCPP members) in the treated area shows a statistically significant lower burglary rate for CCPP members ( $z = 1.818$ ,  $p = .03$ , one-tailed test). In terms of burglary rates, this reflects a 61 percent decrease in the risk rate for treated residences (6.18 per 100 per year to 2.43 per 100 per year) versus a 12 percent decrease in non-CCPP residences (to 5.65 percent). Given the extremely small number of burglary cases occurring for these two groups in 1975, it was not possible to perform a valid statistical test to determine if CCPP reporting rates were significantly higher than non-CCPP reporting rates. However, it should be noted that all burglaries ( $n = 6$ ) occurring in CCPP member residences were reported.

LJPO Telephone Survey. An extensive telephone victimization survey was conducted by the Seattle Law and Justice Planning Office during August and September 1976 to determine the burglary rates for both CCPP members and nonmembers within treated census tracts. This was believed necessary because of the relatively small number of CCPP participants (247) and CCPP "refusers" (248) included in the SEA-KING postsurvey. In addition to increasing the numbers of interviews, the telephone survey allowed data to be gathered from three additional CCPP tracts (for a total of 5 out of 17, or 29.4 percent of all treated tracts as of June 30, 1976).

When the raw data for the 1,970 CCPP members and 1,322 non-CCPP members are converted to a yearly rate and added to the SEA-KING postsurvey, there were significantly fewer burglary occurrences in the CCPP membership than non-CCPP members ( $z = 1.90$ ,  $p < .05$ , one-tailed test) or 9.21 per 100 households per year for treated versus 11.09 for nontreated.

Summary of Victimization Data. All three sources of victimization data collected at different times for different areas and in different

manners indicated a statistically significant ( $p \leq .05$ ) decrease in the occurrence of residential burglary.

Data from the SEA-KING surveys indicate that burglary displacement did not occur to any detectable extent, and that CCPP members do not appear to be self-selected on the basis of either higher or lower than average victimization rates. There does appear to be some self-selection of those that join CCPP on the basis of tendency to report to police if victimized. However, this tendency to report is further increased by CCPP involvement and tends to occur also for non-CCPP members who are given the opportunity to join.

#### OFFICIALLY REPORTED DATA (SPD)

Analysis of the number of residential burglaries reported to the Seattle Police Department on a census tract basis for areas serviced by CCPP in the first and second years (insufficient time having passed to evaluate the effect of third-year areas) indicated a significant relative decrease ( $p < .05$ ) in treated areas of Seattle versus nontreated areas for first-year areas (a -2 percent decrease for treated versus 11 percent increase for untreated areas). However, first-year results may have been confounded by the presence of other burglary related projects operating within the same areas. Areas treated in the second year showed a nonsignificant relative increase compared to the rest of Seattle (-9 percent for treated, -16 percent for nontreated). Combined first- and second-year data indicate an overall nonsignificant decrease favoring CCPP (-6.2 percent decrease in reported residential burglaries for treated areas versus -3.9 percent for untreated areas).

The failure of official SPD data to show the same consistent decrease of burglary victimization data is most likely due to increased reporting rates of residential burglary for areas treated by CCPP.

Objective two, to increase significantly the number of Burglary-in-Progress (BIP) calls, was evaluated using SPD computerized dispatch records (SELECT system). A nonequivalent control group design was used to examine BIP calls as a proportion of all burglary calls received by the SPD between September 30, 1974, and August 8, 1976. As in the case of official SPD residential burglary data, a nontreated area and a treated area were identified, and pre- and postdata were separately determined.

From the pre- to postperiod in nontreated areas, the BIP rate increased 4 percent, or from 8.5 percent to 8.8 percent of the total calls (see table 2). For the treated tracts, the BIP rate increased 27 percent, or from 9.1 to 11.6 percent of the total calls. When the posttreated data are adjusted to exclude the 4 percent increase observed in the S- area, the 9.1 percent to an adjusted 11.2 percent BIP rate is statistically significant ( $\chi^2 = 4.82$ ,  $df = 1$ ,  $p < .05$ ).

In addition, the quality of Burglary-in-Progress calls did not significantly change as a result of more burglary calls being received by the police. Within treated areas, both the number of calls including suspect information and resulting in arrests increased, although nonsignificantly.

TABLE 2.--Burglary-in-progress to total burglary calls

Number of calls classified	Treated car beats				S- car beats			
	Pre-		Post-		Pre-		Post-	
Burglary-in-progress (BIP)	160	9.1%	276	11.6%	431	8.5%	540	8.8%
Not BIP	1,592		2,109		4,634		5,583	
Total	1,752		2,385		5,065		6,123	
Calls per carbeat month	17.88		17.16		13.73		11.27	
Number of carbeat months	98		139		369		543	

## IMPLICATIONS

The data of the present report strongly suggest that the evaluation of crime reduction programs solely on the basis of offenses reported to the police may produce misleading results. This is especially true for those projects that involve a large degree of citizen involvement and as one of their goals include increased citizen observation and reporting of crimes. In the present instance, a significant decrease in the crime of residential burglary was associated with a marginally significant increase in victim reporting. The net effect of these two trends was to produce nonsignificant changes in the number of burglaries reported to the police in treated areas. It may be argued that any program which attempts to enlist citizen action against crime may obtain similar results. Therefore, it would be highly desirable to include victimization surveys as a component of such program evaluations.

A second implication of the data is that crime displacement, at least for residential burglary programs with relatively high citizen involvement (30 to 40 percent), may not be as serious a problem as has been believed. Present data show that burglary incidents were not prevented for program participants at the expense of their neighbors' victimization. In fact, immediate neighbors of program participants (nonparticipants in treated tracts) experienced a nonsignificantly larger reduction in victimization rates than residents in nontreated adjacent census tracts.

ATLANTA'S HIGH IMPACT ANTI-CRIME PROGRAM: NEW DIRECTIONS  
IN EVALUATING PROJECT PERFORMANCE

Samit Roy  
Atlanta Crime Analysis Team

INTRODUCTION

The High Impact Anti-Crime Program got underway in Atlanta in 1972 with concentrated effects for the next 5 years to reduce impact crimes in Atlanta. During the calendar year of 1971, there were a total of 18,398 impact crimes in the City of Atlanta. Regression analysis with monthly historical data showed that total impact crimes were increasing at an average annual rate of 25.1 percent.

The impact goals stated that reduction of 5 percent at the end of 2 years and of 20 percent at the end of 5 years in total impact crimes from the 1971 annual figure should be realized. The progress toward this goal would be used as a yardstick for measuring the success or failure of the Atlanta Impact Program.

One readily apparent drawback of this quantitative goal was that it totally ignored the increasing trend impact crimes were following at the inception of the program. In other words, it was assumed that the number of impact crimes, in absence of the program, would remain at the 1971 level, and that the program would effect reductions from this static figure.

At the end of the 5-year program period, December of 1976, the total number of impact crimes reported in Atlanta were 19,985. This represents an 8.6 percent increase over the 1971 base figure.

The Atlanta Impact Program has come under severe criticism for its failure to achieve the stipulated goal. However, if evaluation of the program is approached by incorporating the upward trend that impact crimes were following at the inception, then the effectiveness of the program becomes apparent.

The evaluation design most suited for observing the effectiveness of crime reduction programs over a period of time is the trend analysis approach. Accordingly, time series models for three distinct periods of the program were calculated and the following table shows the results:

Period	Rate of change	Remarks
1970-71	25.1%	Two years prior to program implementation
1972-74	13.5%	First 3 years of the program
1975-76	-9.3%	Last 2 years of the program

Figure 1 shows the plot of two trend lines for 1970-71 and 1975-76 to illustrate the reverse in impact crime trend brought about by the program. From a strict analytical point of view the impact program has decreased impact crimes by almost 35 percent in Atlanta.

While on a macrolevel the effectiveness of the program was underestimated, at a microlevel the effectiveness of some of the projects was grossly overestimated. This fault is illustrated in the following sections which discuss Atlanta's Anti-Robbery Project evaluation. By virtue of its quantitative goal achievement the Anti-Robbery Project has appealed to many as an exemplary crime reduction project when in fact the contrary may be true due to poor project planning and lack of sound evaluation design.

#### PROJECT DESCRIPTION

The selection of Atlanta as an impact city and the subsequent availability of LEAA funding provided the necessary financial assistance to plan and implement a viable robbery reduction program. Consequently in November of 1974, the Atlanta Bureau of Police Services implemented the antirobbery program aimed primarily at the reduction of commercial and open space robberies.

The antirobbery program incorporated two interrelated law enforcement concepts--apprehension and deterrence. As a short-range response, the antirobbery approach effectively removes the criminal offender from the street through the process of apprehension and conviction. The resultant benefit is an absolute decrease in the offender population in the target areas. The long-range benefit that is derived from the antirobbery approach is the deterrent effect on criminal behavior. The deterrence concept simply implies that as the risk of apprehension is increased, the potential offenders' incentive for criminal activity is reduced.

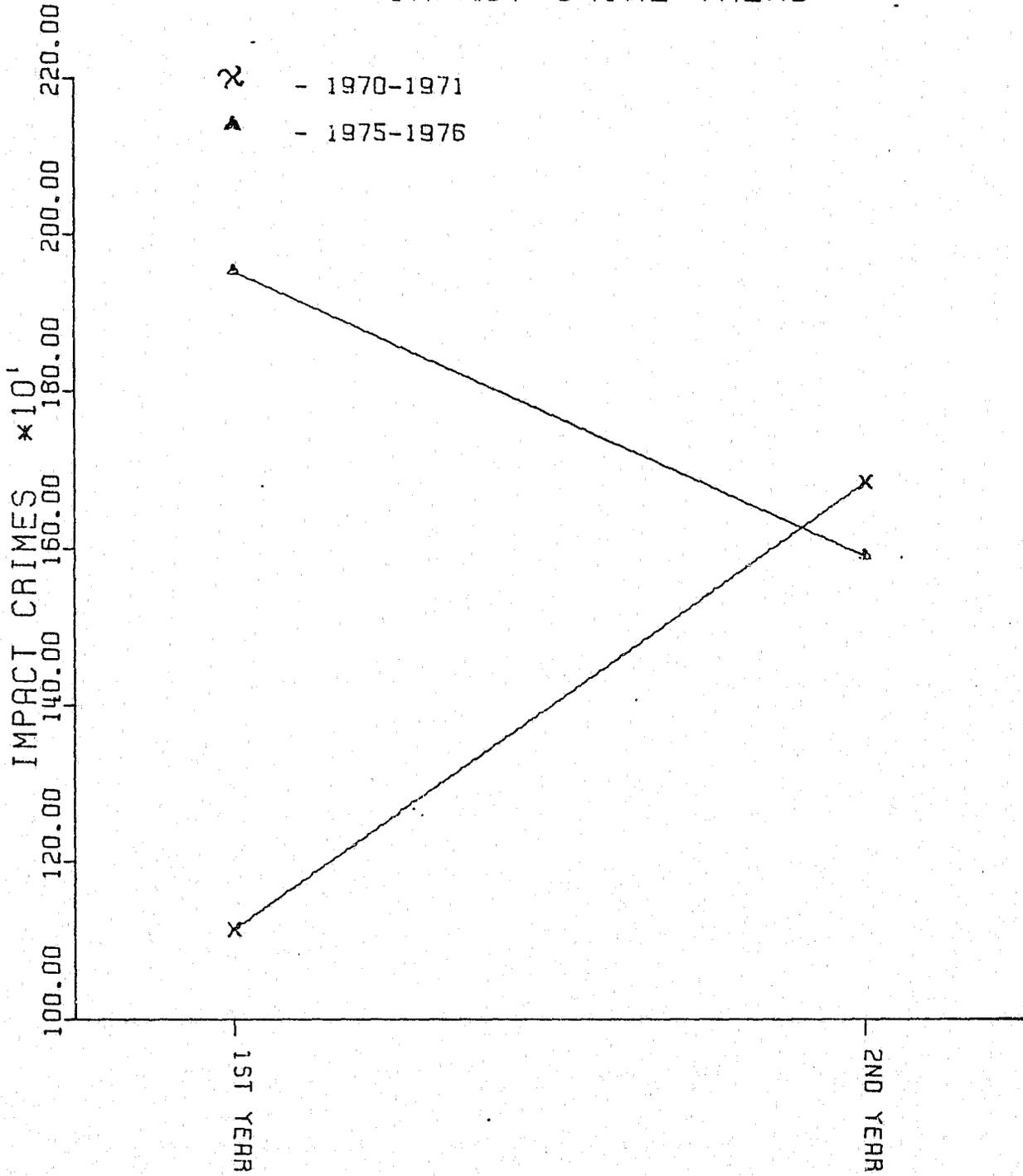
In combination, the apprehension of offenders and the deterrence of population offenders was intended to provide both immediate and long-term solutions for robbery reduction.

To achieve its robbery reduction goals, the AR project employed two basic field techniques. For commercial robberies, stakeout teams were assigned in or near high risk commercial establishments. For open space robberies, decoy teams were placed in areas which displayed a high rate of open space or pedestrian robberies.

The stakeout component of the antirobbery program was typically a two-person team assigned to a commercial establishment. In order to conceal

FIGURE 1

IMPACT CRIME TREND



themselves, the team used storerooms or other areas that were barred from customer traffic. As a prime necessity, the room was equipped with a one-way mirror. In those instances, the stakeout team relied upon their own visual observation of anticipated robbery site, usually the cash register.

The decoy component, which consisted of a team of five or six officers, placed antirobbery detectives in areas that had a high rate of pedestrian robberies. In conducting decoy operations, one member of the team assumed the role of a potential robbery victim. To accomplish this deceptive task, the antirobbery project provided a variety of suitable street clothing and makeup kits whereby the decoy could portray various social and economic character roles. Other members of the team were used as a cover capacity and were responsible for insuring the protection of the decoy victim should a robbery occur. The cover persons were also responsible for apprehending the robbery assailant. To accomplish this, they placed themselves in locations strategic to the reference point of the decoy. When possible they were positioned in such a way as to block avenues of escape.

As an addition, the second phase of the AR operation deployed the Tactical Anti-Crime units. The TAC consisted of four mobile units that had radio controlled holdup alarms which were activated from selected commercial establishments in the course of a holdup. The mobile units, which patrolled in close proximity to the areas that had TAC alarms, provided immediate response to holdup alarms, thus increasing the probability of onsite apprehension.

#### PROJECT GOALS AND EVALUATION ANALYSIS

The goals of the antirobbery project were an absolute reduction of 20 percent in commercial robberies and 15 percent in open space robberies by the end of 2 years of project operation, measured against an annualized base line figure of the year prior to project operation (December 1973 to November 1974). However, careful review of this goal and companion evaluation design revealed a number of shortcomings, most notable of which was the use of the annual base line figure as a yardstick for measuring the success of the project. Comparison of the incidents of robbery from 1 year to the next does not reflect changing conditions over time, and hence, any evaluation of the antirobbery project based on such a static concept could possibly obscure actual performance.

For example, if commercial and open space robberies were increasing or decreasing at a certain annual rate, the project's success should have been measured in terms of changes in the rate of increase or decrease rather than an absolute change from the base figure. Therefore, in order to make the evaluation design as realistic as possible, the following modifications were made:

1. The project performance was measured with emphasis on the rate of change of incidents, and
2. Evaluation of project was done on a 6-month rather than yearly basis in order to pick up the short-term changes in the trend.

## DATA ANALYSIS

The antirobbery project was implemented in Atlanta in December of 1974. In order to estimate the trends in commercial and open space robberies at this point, regression analyses on monthly data going back to January of 1972 for commercial robberies and January of 1973 for open space robberies were done. Trends calculated from these regression models showed that commercial and open space robberies were decreasing at an annual average rate of 7.1 and .9 percent, respectively.

Given this information, it was imperative to estimate the probable number of incidents that would have occurred had the project not been in operation. This approach involves forecasting of time series data for evaluating crime reduction programs. Several time series models were formulated to forecast the number of occurrences during the project period. However, prior to implementation of the project, discrepancies in the classification of robberies into the various categories were a common feature resulting in abnormal fluctuations in the data. Therefore, attempts to forecast with some of the well known time series models did not provide sufficient statistical significance. As an alternative method, forecast estimates were obtained by inflating (deflating) the number of incidents for the comparable period of the previous year by the current rate of change. Thus, for example, the number of commercial and open space robberies during the 6 months from December 1973 to May 1974 were 594 and 1,063 incidents and the average annual rates of decrease (as calculated from the regression models) were 7.1 and 9 percent, respectively. The forecasted number of incidents of the project operation (December 1974 to May 1975) were (594), (92.9) or 552 and (1,063), (99.1) or 1,053, respectively.

These numbers 552 and 1,053 can be interpreted as the number of incidents that would have taken place without any additional stimulus to the robbery reduction efforts and was used as baseline figures against which the project's first 6 months of operation were measured.

During the first 6 months of the project period, 350 decoys and 2,725 stakeout operations were undertaken. The number of commercial and open space robberies recorded at the end of this period were 450 and 839 incidents, respectively. These numbers provided two distinct measures for evaluating the project's performance.

1. From the forecasted number of incidents, commercial robberies decreased by 18.2 percent and open space robberies by 20.2 percent.
2. Adding the monthly number of incidents to the time series models showed that the rate of decrease for commercial robberies increased from 7.1 to 10.3 percent annually and open space robberies from .9 to 7.3 percent annually. The results are summarized in the following table.

	Number of operations	Forecasted number of incidents	Actual number of incidents	Change %	Annual rate decrease *prior	Annual rate decrease *after
Commercial	2,725	552	450	-18.2	7.1	10.3
Open space	350	1,053	839	-20.2	.9	7.3

As mentioned earlier, the dynamic nature of robbery incidents warranted forecasting of baseline data, for each project period using the most recent available data. Therefore, on the assumption that the level of operations undertaken during the first 6 months would at least continue during the second half of the year, forecast of the probable number of incidents was done by deflating the number of incidents for the comparable period prior to the second 6 months of the present project operation (June 1974 to November 1974) by the current rate of decrease. Thus, for the second 6 months of the project operation the forecasted base figures were (565), (89.7) or 507 commercial robberies and 931 (92.7) or 863 open space robberies.

During these 6 months, the antirobbery project conducted 485 decoys and 3,453 stakeout operations. The actual number of incidents during this period were 932 open space and 295 commercial robberies, or an increase of 8.0 percent and a decrease of 41.8 from the forecasted values in open space and commercial robberies. Additionally, update of the time series models showed that for open space robberies the rate of decrease dropped from 7.3 to 5.2 percent annually and the rate of decrease for commercial robberies increased from 10.3 to 13.1 percent annually. The results are summarized in the following tables.

	Number of operations	Forecasted number for the period	Actual number of incidents	Change %	Annual rate decrease *prior	Annual rate decrease *after
Commercial	2,359	507	295	-41.8	-10.3	-13.1
Open space	485	863	932	8.0	-7.3	-5.2

Similar calculations are done for the last two periods of project operation (December 1975-May 1976 and June 1976-November 1976) and the results are summarized in the following tables.

Commercial robberies						
	Number of operations	Forecast number for the period	Actual number period	Change %	Annual rate prior period	Annual rate after period
1st period	2,725	552	450	-18.2	-7.1%	-10.3%
2nd period	3,453	507	295	-41.8	-10.3%	-13.1%
3rd period	1,421-S 4,882-T	391	459	17.4	-13.1%	-11.2%
4th period	1,052-S 3,352-T	262	413	57.6	-11.2%	-10.1%
Open space robberies						
1st period	350	1,053	839	-20.27	-.9%	-7.3%
2nd period	485	863	932	8.0	-7.3%	-5.2%
3rd period	379	795	877	10.3	-5.2%	-5.1%
4th period	699	884	928	5.0	-5.1%	-3.8%

The captioned table shows that the net effects of the project were increasing the rate of decrease of commercial robberies from 7.1 to 10.1 percent and open space robberies from .9 to 3.8 percent annually. If the absolute changes between the base period and the project periods are calculated (as was in the original grant evaluation component) the rates are as follows:

<u>Commercial robberies</u>		
<u>Base period</u>	<u>1st year operation</u>	<u>2nd year operation</u>
1,159	745	878

Change between base period and 2nd year:

$$\frac{(878 - 1,159)}{1,159} \times 100 = -24.2\%$$

<u>Open space robberies</u>		
<u>Base period</u>	<u>1st year operation</u>	<u>2nd year operation</u>
1,194	1,771	1,805

FIGURE 2

COMMERCIAL ROBBERIES

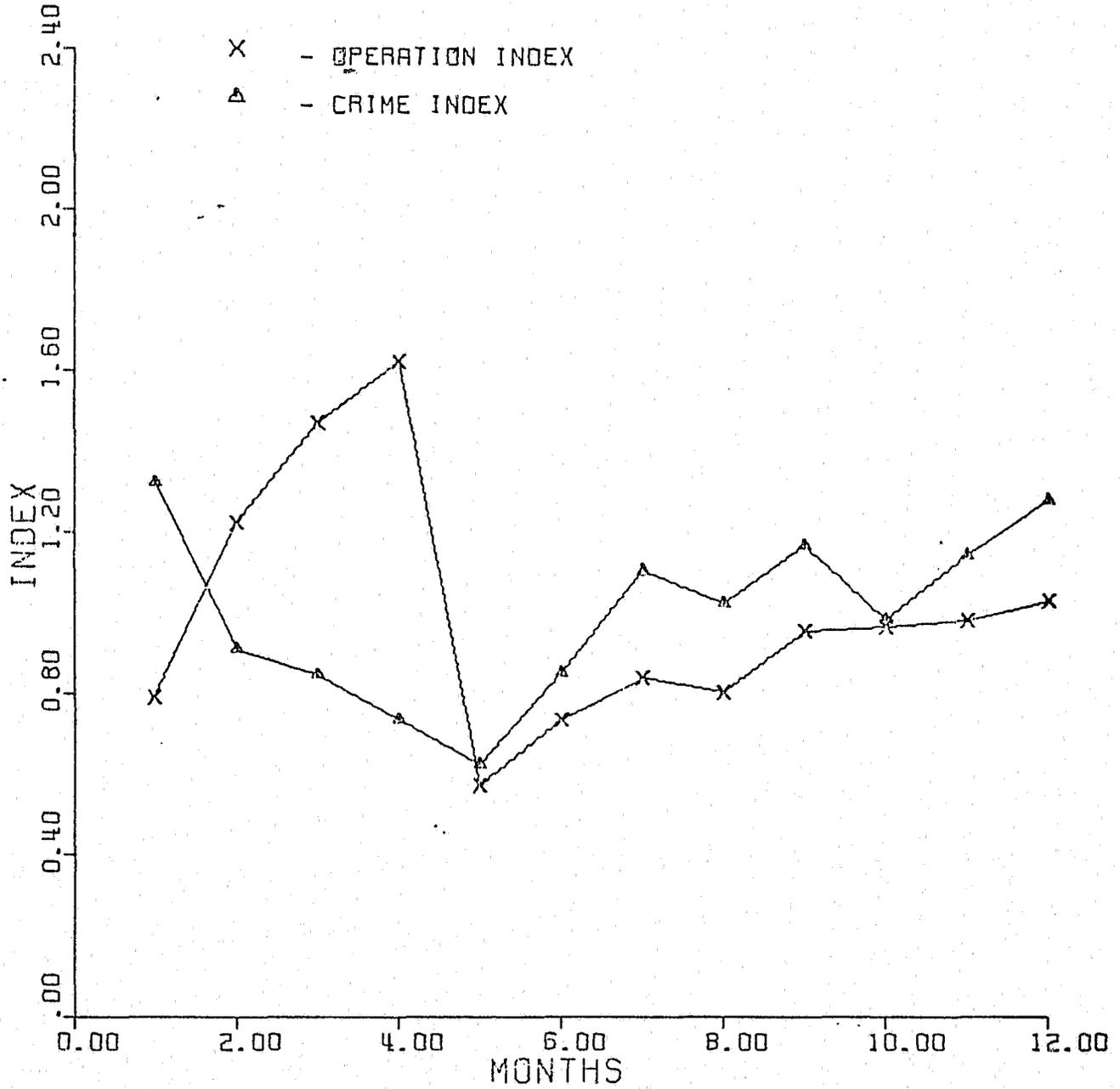
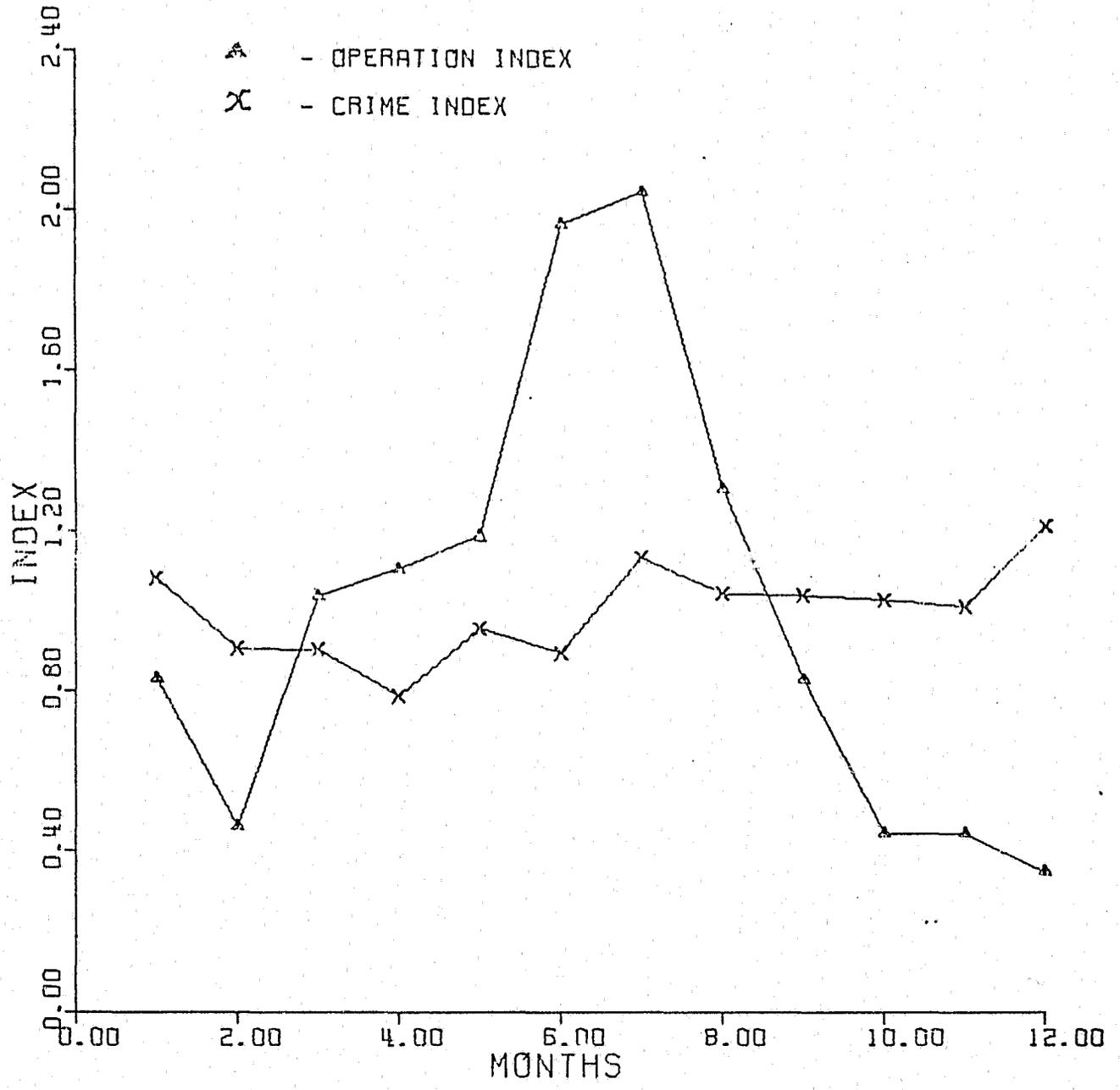


FIGURE 3

OPEN SPACE ROBBERIES



Change between base period and 2nd year operation

$$\frac{(1,850 - 1,994)}{1,994} \times 100 = 9.5\%$$

Conclusions that can be drawn from the two methods of evaluation are diversified. If the performance of the project is based on the static year-to-year changes, then it can be concluded that the project did make sufficient impact in the reduction of commercial and open space robberies. However, if reference is made to the fact that both these categories were decreasing at the inception of the project, and that some reduction in robberies would have been expected without the project, then this conclusion becomes rather optimistic.

The more realistic conclusion can be drawn from the evaluation approach of the first section. This analysis shows that during the initial periods the project made considerable impact in the reduction of commercial and open space robberies. However, in the subsequent periods the project had little or no impact in the reduction of robberies.

The latter conclusion about considerable impact during the first periods and not sufficient impact in the subsequent periods raises one important question: Was the antirobbery concept ineffective as a crime reduction measure or was the planning and operational aspect poor?

One answer to the question can be found in the deployment pattern followed during the 24 months of project operation.

Both commercial and open space robberies follow extreme seasonal patterns.

Lack of proper deployment policy is apparent by comparing the number of stakeout operations and incidents of commercial robberies between first and second period. A 26.7 percent increase in stakeout operation resulted in an additional decrease of 23.6 percent in commercial robberies. Instead of keeping the stakeout operations, at least, at this level third period saw a 58.8 percent reduction in stakeout operation coupled with the introduction of TAC units. During this period commercial robberies increased by 17.4 percent. The fourth period's deployment pattern followed that the previous period causing commercial robberies to take over a 50 percent jump from the forecasted number.

The inability of TAC to be substituted for stakeout was indicated after the third period's evaluation. This information was, however, not used as feedback for subsequent deployment pattern. The primary objective of TAC operations was to effect on-site apprehension. But after 12 months and 8,234 operations, only 4 on-site apprehensions were made.

The evaluation of the project at this micro level shows that even though the original project goal came close to being achieved, the project, during the entire 24 months of operation, fell short of its potential. The evaluation further shows that the stakeout and decoy concepts in commercial and robbery reductions have high indications of being effective, yet lack

of proper planning and feedback from interim evaluations did not provide sufficient opportunities to establish this credibility.

## CONCLUSION

The experience gained during the course of the High Impact Evaluation shows that lack of proper quantitative assessment of the crime problem at the inception of the program has made the outcome very deceptive.

The very important dynamic nature of the impact crimes was ignored during the course of 5 years of program operation. Another less apparent inconsistency in the evaluation design stems from the fact that a macro level goal achievement was anticipated by concentrating on projects at very micro level, independently addressing certain categories or subcategories of crimes without any formalized priority structure.

The most important lessons learned from the impact evaluation are:

1. Any crime reduction program must at the onset assess quantitatively the trends and have some indications of what might be expected (forecasted values) had the project not been in operation. These trends and forecasts should be used as a yardstick for measuring the success or failure of the projects.
2. Sound and consistent planning of the concepts is imperative in order to correlate the project with the achievements.

METHADONE PROGRAMS AND CRIME REDUCTION: A COMPARISON BETWEEN  
NEW YORK AND CALIFORNIA PATIENTS\*

Dale K. Sechrest  
Deputy Director, Commission on Accreditation for Corrections

From 1969 to 1975 the Addiction Research and Treatment Corporation (ARTC) drug treatment program, located in the Bedford-Stuyvesant/Fort Greene area of Brooklyn, New York, was evaluated by the Columbia University School of Social Work, Yale Medical School, and the Harvard Law School Center for Criminal Justice. The findings and conclusions reported here are drawn from the Center for Criminal Justice study, which focused on reductions in criminal activity of methadone patients due to ARTC treatment. In addition, findings are reported and comparisons made with patients admitted to the Santa Clara County, California, Methadone Treatment and Rehabilitation Program.

While several questions were addressed in both studies, findings presented here are in response to the following questions:

1. Was patient criminal activity reduced on preaddiction levels, or lower, thus having an impact on community crime rates?
2. Did ARTC treatment produce decreased criminal activity only for specific groups of patients and not others?
3. What background and program treatment factors were most related to program success?

Finally, the future of methadone treatment programs of the ARTC type was addressed with recommendations regarding an overall approach to drug addiction treatment.

---

\*This paper draws on the final report prepared by the author on the methadone maintenance project carried out by the Center for Criminal Justice, Harvard Law School. The author assumed direction of the project during its final 2 years and is understandably indebted to those who preceded and assisted him: Professor James Vorenberg; Professor Lloyd E. Ohlin; Eleanor Halprin; Edward Dauber; Dr. Gila Hayim; Dr. Alden D. Miller; Paul Cirel; Marion Coates; and Professor Irving Lukoff, Columbia University School of Social Work. Opinions expressed are those of the author and do not imply endorsement of the Law Enforcement Assistance Administration which provided financial assistance under Grant No. 75-NI-99-0073.

ARTC and the California program were basically low-dose methadone maintenance drug treatment programs, with very broad admissions policies designed to make services available to virtually the entire heroin addict populations served. The supportive services available were similar to most drug treatment programs, including social services, individual and group counseling, assistance in job-finding, referrals to other community service agencies, and educational and legal assistance. Decisions on chemotherapy were made by medical personnel.

As with most such programs, both ARTC and the Santa Clara County program emphasized that the following benefits should accrue to patients and the community: (1) decreased drug use of all kinds; (2) decreased criminal activity; (3) improved patient performance in several areas of social functioning, such as employment, family relationships, relationships with friends and associates (including criminal justice agencies); and (4) improved use of time.

#### RESEARCH METHODS

Follow-up data on criminal activity were obtained from official New York City and California State records, with all of the inherent problems of comprehensiveness and accuracy. An additional problem was the impossibility of developing an appropriate control group against which to measure the performance of program patients. The random placement of addicts into experimental (treated) and control groups (untreated) could not be arranged. Furthermore, there were no strictly comparable alternate treatment modalities to which addicts could be referred for research purposes. Therefore, all measures of reduced criminal activity have been made against the patient's performance at the time of program admission (baseline), or in the preaddiction period.

Arrest rates and mean severity scores were computed using individual patient arrest data. Both measures were computed for three specific periods of patient activity: preaddiction, addiction, and postprogram entry. The arrest rate was computed based upon the number of arrests divided by the number of years spent in each period. The mean severity score was computed based upon the severity of all arrests summed and averaged for each period. The severity scale used was based on a modified version of the Sellin-Wolfgang scale (Sellin and Wolfgang, 1964, and New York City Criminal Justice Evaluation Project).

Offenses were grouped into seven major categories for analysis: drugs (possession and sale), property, forgery, robbery, assault (including rape and homicide), prostitution, threshold offenses (possession of weapons and burglary tools and criminal trespass), and violations (disorderly conduct, loitering, public intoxication, gambling, etc.). Eighty-two (82) demographic and background variables were used in a detailed analysis of the ARTC population.

Between October 8, 1969, and June 23, 1971, 990 addicts entered the ARTC program. Intensive analysis focused on a subgroup of 473 of these patients on whom official criminal activity was available for a 3-year period after program entry, all of whom had entered between October 8, 1969, and July of 1970. As will be shown, the 3-year period of followup was critical to an adequate analysis of the data. The third population studied were

methadone patients from the Santa Clara County, California, Methadone Treatment Program.\* That program began operation in February of 1970, and 271 patients who entered that program in 1970 and the first half of 1971 were used as a comparison group with the 473 ARTC patients. Although the Santa Clara County program accepted addicts at age 18 instead of 21, in almost all other respects the ARTC and California programs were similar (Sechrest and Dunckley, 1975). These populations were similar in age and length of addiction, differing primarily in ethnic/racial background.

Data Collection. Primary data collection was the responsibility of the evaluation team of the Columbia University School of Social Work, which was headed by Dr. Irving Lukoff. The Columbia Team had the primary responsibility since they were conducting studies on the broader social implications of the ARTC drug treatment program. Methadone dosage and urine test data were produced through the auspices of Creative Bio-Medics and the Yale University medical research team. Center staff coded arrest data into severity scores and provided the data on the California population, in addition to conducting studies on court and police practices relating to addicts in Brooklyn.

## FINDINGS

Patient Population Characteristics. Brooklyn patients were 13 percent Caucasian, 78 percent black, and 8 percent Spanish (Puerto Rican descent). California patients differed primarily in their cultural background with 50 percent Caucasian, 6 percent Black, and 44 percent Spanish (Mexican-American descent). About 80 percent of each population was male. The Californians were younger--29.5 years as opposed to 32.7 years (Brooklyn) average age at entry, and had used drugs a shorter period of time--9.4 years as opposed to 11.4 years (Brooklyn) at entry. This was partially due to the fact that California patients could enter the program at an earlier age. Just over half were married in each subgroup, and each had about the same proportion with a high school diploma.

It is not surprising to find a great deal of criminal justice system involvement in both populations; for the Brooklyn patients, 88.5 percent of the total population reported being arrested at some time in their lives, 61.5 percent ever convicted, and 65.5 percent having done time in a jail, prison, or penitentiary, with a mean stay of about three years.\*\* In the California population, 95.1 percent had been arrested at some time in their lives.

Overall Crime Reductions. Arrest rates rose very significantly (.001 level) from the preaddiction to the addiction period for both populations, males and females. In the second year after entry arrest rates began a

---

\*Santa Clara County is located 60 miles south of San Francisco and has a population of just over 1 million residents; the principal city is San Jose, which contains about half of those residents.

\*\*It was not clear whether this time represented detention time and sentenced time combined, or just the latter.

decline, and for the Brooklyn population, which was followed 3 years, rates were at the preaddiction level for males, but not for females, at the third year after program entry.

The importance of the 3-year followup is shown in the nondrug rates, which did not begin to show a decline until the third year after entry. Almost all of the initial decline in criminal activity at the second program year was due to a decline in drug crimes. The drug arrest decrease made treatment appear to have more impact on overall criminal activity than it actually produced. It can be hypothesized that much of the initial postentry decline might have occurred in any event, since these patients appeared to have reached a drug crime peak in the year prior to program entry. The point at which nonprogram influence declined, or "natural" decline, stops and program intervention becomes a factor cannot be determined without a control group population. One might assume an unassisted decline in criminal activity to at least the total addiction period level, particularly if there was no longer a serious habit to support. Therefore, a further decrease in criminal activity to the preaddiction level was seen as the desired goal. In this sense, ARTC showed some apparent success at the third year after program entry. A similar pattern was found for severity scores, with their marked decline beginning in the third year after program entry.

It was possible, of course, that the decline in criminal activity as measured by arrests reflected changes in police arrest policies rewarding drug addicts rather than the results of the treatment program. The declining arrest rates appeared to show the effects of either or both of these influences. For Brooklyn as a whole, overall narcotic arrests decreased 39 percent from 1971 to 1972 due to changes in police policy (Crime and Analysis Division, New York City Police Department, 1971). This policy stated that narcotic arrests would be the primary concern of the Narcotics Division which would concentrate its efforts on higher levels of drug activity, i.e., major dealers and distributors. This was coupled with an apparent feeling on the part of the patrolman that minimal levels of addiction would have to be tolerated and attempts made to refer known addicts to treatment programs in lieu of arrest (Coates and Miller, 1974). In order to determine the impact of these policy changes, the experience of the 1969 Brooklyn patients was compared to 1970 patients for total arrests and drug and nondrug arrests during each of the 3 years after program entry. Both groups of patients showed arrest rate reductions greater than what was expected due to the shift in police policies, based on overall narcotic arrest rates in the community. A similar comparison using crime severity scores led to the same conclusion.

Another complication with the comparisons across time was that more recent police records may not have been complete. Analysis of the arrest data by the Columbia University team suggested that the third-year rates might increase by as much as one-third had we been able to recheck the records in another year or so, since a lag in reporting arrests affected the data in each period. This would of course weaken the decline in arrests shown, even though rates for earlier periods would increase slightly also. Still, such increases would not eliminate completely the downward trend that began in the third year. The third year rates would fall somewhere between the preaddiction rates and the rates for the total addiction period.

This complication applied only to the arrest rates. Since the severity scores are averaged across recorded arrests, randomly missed data did not systematically affect mean severity scores.

Arrest rates and severity scores were also computed for the California population. The overall patterns of arrest rates and severity scores were similar to the Brooklyn population. However, arrest rates were much higher for the California population, with severity scores lower on the average. Increased arrests for violations, property crimes, and drug offenses accounted for these differences. In other words, California addicts were arrested more times per year but for less severe types of crimes on the whole. Brooklyn addicts, on the other hand, were arrested less often, but when arrested it was for a more severe type of crime.

The California population began its decline in criminal activity at the second year after program entry. No data were available for the third year after entry. While the nondrug crimes did not show the persistence found for Brooklyn patients, they did not decline as much as the total arrest rates, and the female population showed a consistent increase in the rate, but not the severity, of nondrug arrests.

With respect to patient age, the younger Brooklyn patients had higher arrest rates. Severity scores showed, however, that patients in the 22 to 29 age group were more likely to be arrested for more severe crimes. Arrest rates for the Californians were similar, but greater severity was clearly attributed to the younger patients.

What specific offenses showed significant decreases subsequent to program entry? Drug offenses showed the greatest fluctuation over time, but no group of patients showed a significant drug arrest rate reduction from the preaddiction period in the second or third year (Brooklyn) after program entry; however, the female patients in both populations showed earlier declines in drug arrests in comparison to the preaddiction period.

For the California population property arrests showed no significant decrease in relation to the preaddiction period, but the postentry decrease was significant for males and females. For Brooklyn patients, the total population had no significant decrease in property arrests, although male patients did decrease property arrests significantly (.05 level). Brooklyn female patients retained a significantly higher (.05 level) rate of property arrests at the third year after entry over the preaddiction period, although it was still a decrease from the second year after entry (.15 to .07 arrest rate).

For the Brooklyn population forgery arrests declined after program entry, never having been a serious category of criminal activity. For California patients, however, forgery increased significantly (.001 level) in the second year after program entry from the preaddiction period, the addiction period, and from the first year after entry, indicating a sharp increase in forgery arrests before and after program entry, for both males and females. There is no apparent explanation for this increase in arrests for forgery offenses; however, the increases were principally for younger patients who were more likely to be program dropouts reverting back to drugs, using forgery to support their habits as they had in the period of addiction.

Numbers of robbery arrests were greatest in the addiction period for both populations, primarily involving males (90 percent). The rates remained relatively low and stable, however, with no significant changes across any periods, although Brooklyn patients, male and female, accounted equally for a small increase in the postentry period.

Assault arrests accounted for the inability of Brooklyn female patients to return to their preaddiction level for nondrug arrests. California male and female patients significantly increased their assaultive behavior from the first to the second year after program entry. Brooklyn males showed an increase in assault arrests in the second year after entry, but declined thereafter, as did Brooklyn females.

Threshold arrests remained relatively stable over time for both populations, male and female, with patients overall returning to the preaddiction level of involvement.

Brooklyn males showed a significant decrease (.001 level) in minor violations from the preaddiction period to the third year after program entry, while females remained at a higher (but insignificant) rate than the preaddiction period, this being a decline from the first year after program entry. Neither California males nor females had declined to the preaddiction rate at the second year after entry, both having significantly greater rates (.05 level), although rates did decline from the first year after entry.

In conclusion, there was an overall decline in arrests for Brooklyn patients from program entry to the third year after program entry. The decline was not immediate, taking more than 2 years to return to the preaddiction rate of criminal activity. The greatest declines were for drug arrests, minor violations, and, to some extent, property crimes. At the second year after program entry California patients had an overall decline in arrest rates. However, forgery and assault rates showed persistent increases before and after program entry. Whether these differences from Brooklyn patients can be attributed primarily to program dropouts or to cultural differences is not clear, although Brooklyn patients did show the same pattern for assault from the first to second year after program entry, only to decline further in the third year after entry. Robbery rates also showed slight increases for Brooklyn patients in the postentry period, which persisted for male patients through the third year. It appears that an initial increase in assaultive behavior occurs in both populations, and then declines in subsequent years. The increases may be due, in part, to a return to alcohol use, although this is not a clear finding.

Precinct Criminal Activity. Though it seemed unlikely that reductions in the criminal activity of ARTC patients, as indicated by arrest rates, would be sufficient to produce lower overall community crime rates, criminal complaints obtained from the New York City Police Department were examined for the ARTC service area ("catchment area"). Ten precincts were studied for reductions in relation to the number of ARTC patients served and complaints to the police per 10,000 residents in each precinct. There were no marked reductions in criminal activity in the four (4) precincts served primarily by ARTC in comparison with six (6) adjacent Brooklyn precincts. The relative rankings of criminal complaints by precinct remained about the same from 1968 through 1973, the period 2 years before and 4 years

after the program began. The overall conclusion was that the crime rate decreased slightly for the entire area (10 precincts), but not necessarily in those areas where ARTC patients were located in greatest numbers.

What accounts for this apparent contradiction in results? If individual rates of crime by patients were decreasing, why were crime rates relatively stable in the community? The simplest explanation is that ARTC was not serving enough patients to generate a community-wide impact. No data were available on the number of addicts actually located in the ARTC catchment area, but figures from the New York City Narcotics Register show that Brooklyn, followed closely by Manhattan, had the largest number of first-reported new addicts in the City--close to 9,000 in the peak year of 1972, a very large number of new addicts compared to the number treated in the ARTC program. An alternative explanation is that the program did not reach the most crime-prone population of addicts, and that ARTC patients were older and more likely to be "maturing out" or "burning out" of criminal activity. The sample ARTC population (473 patients) was age 33, on the average, with an addiction history spanning just over 11 (11.4) years. New York City Narcotics Register figures show that almost 85 percent of the individuals reported as "New Cases" in 1970 were age 30 or less, although this proportion went down slightly in 1971-73 (New York City Narcotics Register, 1973). Data on the ARTC sample show consistently higher arrest rates for 22- to 29-year-old patients. The FBI Uniform Crime Reports regularly show that about half of all crimes reported are for people 25 years of age or younger. The implications of these conclusions for future planning will be presented after other findings are summarized.

Characteristics Related to Program Outcome. One of the primary goals of the study was to determine those characteristics related to decreased criminal activity. To determine those factors most likely to lead to such decreases, 82 background/demographic and program performance variables were used with 5 outcome variables in a stepwise multiple regression analysis.\* This is a more powerful variation of multiple regression, which allows for the choosing of independent variables which will provide the best prediction possible with the fewest independent variables. Using this technique with 5 outcome variables, a total of 25 independent variables were found in various combinations to be significantly related to outcome.

In general, there was no set of independent variables which gave a strong prediction of reduced criminal activity, for drug or nondrug arrest rates or severity scores, although females showed better results than males.

For males and females a better preprogram drug history and decreased drug use while on the program (fewer morphine positives) were related to decreased criminal activity at the third year after program entry. While the latter has no predictive value at program entry, it does emphasize the importance of a sound urine testing program and the importance of taking action where dirty urines are found--i.e., intensifying treatment effort. White males were particularly prone to continue drug crimes after entry. However, use of hallucinogens, high liquor intake, and morphine positives

---

\*Outcome variables were: Total Arrest Rate, Drug Arrest Rate, Nondrug Arrest Rate, Total Severity Score, and Nondrug Severity Score.

were significant components of the prediction equation. Apparently, white patients with a high self-reported drug use and dirty urines are a group requiring special attention, particularly if drug arrest rates are a concern.

It is of interest that a higher age of first daily heroin use tended to lead to lower nondrug arrest rates in male patients. This finding supports the thesis that patients who start drug-taking older may have skills which enable them to support themselves both before and after program entry.

For female patients several background factors were significantly related to decreased criminal activity. These included such variables as being better educated, not living alone at entry, and having low residential mobility. An additional predictive factor for males in relation to decreased criminal activity was family stability. Treatment variables were not found related to outcome for males or females.

It was found that age at program entry was significantly related to outcome. It appears that the ARTC population was older and may have completed a cycle of drug use which has enabled them to "burn out" or "mature out" of both drug-taking behavior and criminal activity. Programs which deal with the younger addict seem likely to have a much more difficult time achieving this level of success in reducing criminal behavior in patients.

#### SUMMARY AND CONCLUSIONS

Methadone progress in Santa Clara County (California) and Brooklyn, New York (Addiction Research and Treatment Corporation) were evaluated to determine the extent of reductions in individual and community-wide criminal activity. Findings showed an overall decline in official (recorded) arrests for Brooklyn patients in a 3-year follow-up period. California patients showed similar declines in a 2-year follow-up period. Assault rates for both populations showed less of a decline, however. It was found that Brooklyn patients were arrested less frequently but for more severe offenses across preaddiction, addiction, and postprogram entry periods. Criminal complaints to the police in the ARTC area were not reduced as compared to surrounding precincts. Individual and program performance characteristics were examined for Brooklyn patients in relation to reductions in criminal activity. A better preprogram drug history and decreased drug use while on the program were related significantly to decreased criminal activity for male and female patients, and older patients had slightly better outcomes.

For future drug treatment programs three major program requirements are recommended. First, that the epidemiologic outreach model described by Hughes (1972), Greene (1974), and others be a requirement of any such program. Second, that the administration of the program be sound and the program be committed to a continuous, long-term effort. One-year or 2-year efforts are simply not effective. And third, that the range of services provided be both administratively integrated and broadened in combined maintenance and drug-free approaches, even to include morphine maintenance for short periods of time after very careful screening procedures.

The treatment of heroin addiction takes time. It cannot be accomplished in 2 years from the majority of addicted individuals. Aggressive behavior, major and minor law violations, and drug violations continue, including alcohol abuse. Reductions in community criminal activity as a result of program activity will not be found until epidemiologic methods are applied to stop the production of new addicts in the community. Programs for treatment must be attractive, having something to offer in return for stopping drug abuse and criminal activity, and in order for programs to be effective they must be long term.

Finally, the range of services provided by these programs must be expanded, but only after all services for addicts have been integrated into one administrative structure. Nonmaintenance or "drug-free" programs should be operated alongside maintenance programs. A central screening mechanism should be developed in each community for the initial processing of identified addicts, whether they come from the criminal justice system or are volunteers. A decision should be made as to the proper treatment for them. This treatment need not be static or of one kind only. Too often methadone maintenance is the "only shop in town" for the addict. Perhaps after appropriate screening the addict could even be given morphine or methadone maintenance for a brief period, then moved to a therapeutic community, and then given close supervision and support in the community. The goal at all times would be on developing the patient's ability to manage his life and develop a lifestyle over time which would allow him to become self-sufficient, healthy, and law-abiding.

At this point in time we are faced with some choices. We can continue to work with those individual addicts who volunteer for our programs, whether by free choice or some form of law enforcement pressure. In doing so we will continue to find that they are the older, more motivated addicts who are ready to try something--anything. If we choose to continue in this way, we can forget about producing significant reductions in community criminal activity (i.e., the fabled "impact") and continue to work with whom we can. The money will probably be well spent in assisting these individuals in developing a lifestyle which will be at least nondamaging to society, even if they become (or remain) welfare recipients or find marginal employment. On the other hand, by using existing techniques, such as the epidemiologic field approach, and by utilizing existing administrative structures in the application of an integrated use of a broad range of services in a given community, it may be possible to cope with large-scale heroin epidemics now and in the future. This would reduce the frustration inherent in treating addicts who have lost a great part of their lives to addiction, and in the long run it may impact directly on the criminal justice system and society as a whole.

## BIBLIOGRAPHY

- Coates, R. C., and Miller, A. D. "Patrolmen and Addicts: A Study of Police Perception and Police-Citizen Interaction." J. of Police Science and Admin., Vol. 2, Fall 1974.
- Greene, M. H. "An Epidemiologic Assessment of Heroin Use." A.J.P.H. Suppl. 64: 1-10, 1974.
- Hayim, Gila J. "Changes in Criminal Behavior of Heroin Addicts: A Two Year Follow-up of Methadone Treatment," July 1973, p. 35 (unpublished mimeo).
- Hughes, P. H., Senay, Ed., and Parker, R. "The Medical Management of a Heroin Epidemic." Arch. Gen. Psychiatry 27: 585-591, 1972.
- Hughes, P. H., Sanders, C. R., and Schaps, E. "The Impact of Medical Intervention in Three Heroin Copping Areas." Proceedings, Fourth National Conference on Methadone Treatment, San Francisco, Calif., 1972.
- New York City Narcotics Register. "Analysis of Narcotics Addiction Trends Through June, 1973," September 1973.
- New York City Police Department, Crime Analysis Division. "Statistical Report, Narcotics, New York City, 1972." (Mimeo.)
- Sechrest, D. K., and Dunckley, T. E. "Criminal Activity, Wages Earned, and Drug Use After Two Years of Methadone Treatment." Addictive Diseases: An Int. J. 1: 491-512, 1975.
- Sellin, Thorsten, and Wolfgang, Marvin. The Measurement of Delinquency. New York: John Wiley and Sons, Inc., 1964.

## STREET LAYOUT AND RESIDENTIAL BURGLARY<sup>1</sup>

Julia Nutter, Douglas W. Frisbie, and Carol Bevis  
Crime Prevention Through Environmental Design Project  
Governor's Commission on Crime Prevention and Control  
St. Paul, Minnesota

This report summarizes a preliminary study of street layout as a possible deterrent of crime. Basis for the study is the work of Oscar Newman.

In his book, Defensible Space, Newman suggests that residents who perceive an area as their territory have greater concern for what happens there. Newman further states, "It is possible to subdivide the existing fabric of city streets in order to create territorially defined blocks and areas." As the territorial subdivision of streets in an area increases, the residents are more likely to increase surveillance and can better recognize who does and does not belong in the area. Thus, when territoriality increases, surveillance also increases, both of which contribute to a reduction of crime.<sup>2</sup>

Street layout can deter crime by strangers in ways other than increasing the territoriality and surveillance by residents. Interviews with burglars indicated they prefer to be familiar with the areas they victimize and they select targets which are convenient for both access and departure.<sup>3</sup> Burglars (like everyone else) are probably less familiar with, and find less convenience in, those areas which are somewhat isolated from the rest of the city due to inaccessible streets. These interviews additionally indicated that burglars avoid areas where they might be more easily identified as a stranger.

Finally, street layout may indirectly deter crime by local residents. Blocks less accessible to other streets will probably be traveled most often by local residents. If local residents are the primary users of their streets, those streets can more easily be adapted in a territorial way by the residents. Not only will it be easier to observe strangers on these blocks but greater concern for the area by residents should result in greater surveillance. Residents will be less hesitant to challenge anyone, stranger or known, who is engaged in suspicious or disruptive behavior.

There are examples in several cities of streets being changed for the purpose of reducing crime. Experiments in St. Louis and Brooklyn suggest that a relationship between street arrangements and crime has a basis in fact, not just in the theories and common wisdoms listed above. In St. Louis, several streets were closed at one end. Residents assumed responsibility for road and streetlight maintenance on their streets and, in

return, received a slight rebate on their city taxes. The rate of reported crime is lower on these closed streets and residents manifest proprietary feelings by surveilling the street more and questioning the intentions of strangers.<sup>4</sup>

St. Marks Avenue in Brooklyn was redesigned to slow traffic. Symbolic gateways were placed at each end of the street. A mid-block portion of the street was completely closed to traffic, turning it into a play and communal area. Residents reportedly defined the area as their own--illustrated by their cleaning the street every Saturday morning--and felt that crime had lessened significantly.<sup>5</sup>

In Berkeley, California, a complex set of traffic barriers and diverters was designed solely for traffic control. The impact of these barriers on crime was unclear due to concurrent introduction of new police patrol tactics.<sup>6</sup>

The experiments in St. Louis, Brooklyn, and Berkeley suggest that the purposeful redesign of streets for reducing crime is a promising strategy. Prior to redesigning streets in parts of Minneapolis, however, it is wise to determine the extent to which existing types of street designs in Minneapolis exhibit differing crime rates.

Police have long contended that there are fewer crimes on cul-de-sacs and dead ends than on other streets<sup>7</sup> but there has been little effort anywhere to statistically document this common wisdom. Newman's theory provides support for this contention.

This initial investigation looks at residential streets and their associated rates of residential burglary. If investigation finds that residences along certain kinds of streets are burglarized less, the experimental use of those street types is more likely to be a successful use of resources.

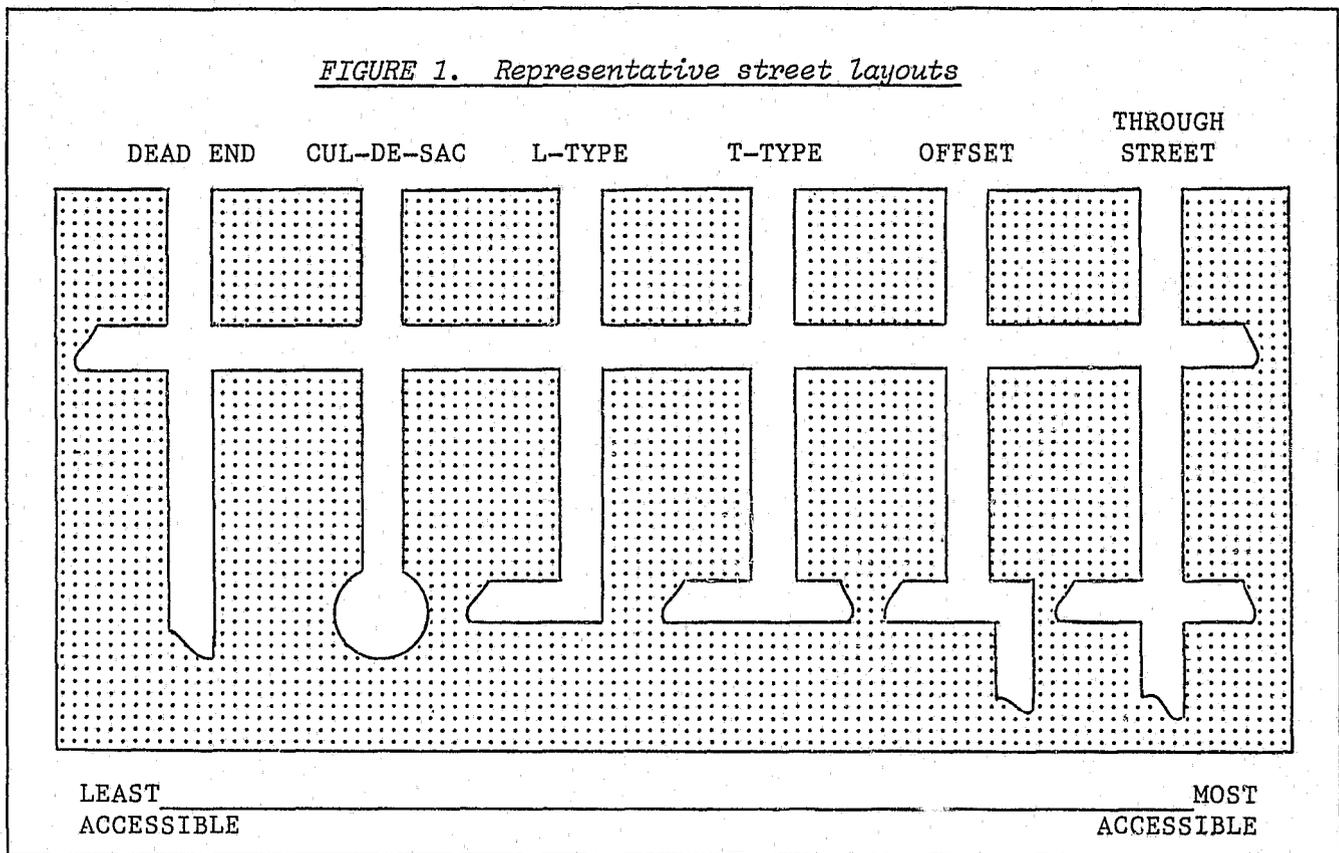
Of course, redesign of streets in an attempt to reduce crime rates could restrict police patrol and emergency vehicle access to these streets. Thus, any redesign that makes streets less accessible to criminals must also consider unwanted effects on the accessibility of needed city services. In addition, the redesign of streets should not be considered a complete cure, merely one of several crime preventive steps to be used with caution. As such it should be used only after careful analysis of the crime problem suggests it as a logical strategy.

#### METHODOLOGY

The analysis for Minneapolis is based on six basic types of streets and the residential burglary rates associated with them. The street layouts considered for this analysis are indicated in figure 1. The streets are ordered from those which are generally the least accessible to those which are most accessible.

To sample each street layout, a multistage random sampling method was used. The stages included randomly selecting 30 census tracts from the 127 tracts in Minneapolis, selecting street types within those tracts, and eliminating selected street blocks which had no residences.

FIGURE 1. Representative street layouts



A study sample and a control sample of blocks were selected. The 65 study blocks consisted of 11 dead ends, 13 cul-de-sacs, 13 L-types, 16 T-types, and 12 offsets.

For each study block a control block was designated. The control block was defined as the nearest through-street block feeding into the study block.

For both study and control blocks, "block" meant a section of street with an intersection at each end and no intersection between the ends.<sup>8</sup>

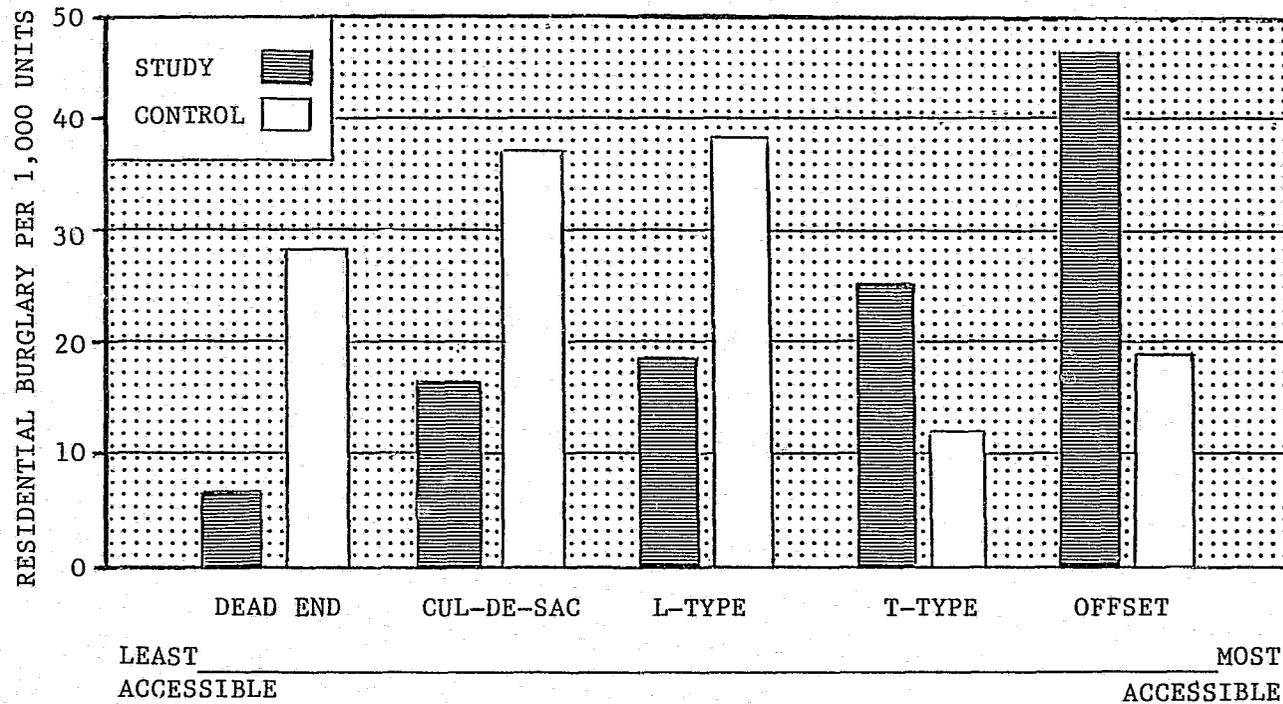
After locating study and control blocks, the number of housing units on each block was recorded as was the number of residential burglaries.<sup>9</sup> Both pieces of information were used in this preliminary study.

#### FINDINGS

As indicated in figure 2, the data show a noticeable pattern of lower residential burglary rates for housing on those study blocks with lower accessibility. There is an upward trend that relates increasing street accessibility with rising burglary rates.

Figure 2 also compares burglary rates for study blocks with those of the control blocks. Figure 2 shows the residential burglary rates were

FIGURE 2. Residential burglary rate for study and control blocks<sup>a</sup>



<sup>a</sup>N = 2,348 residential units upon which burglary rates were based.

Control blocks are the through blocks nearest to their corresponding study blocks.

lower on most study blocks than on their corresponding control blocks. This pattern did not hold for T-types and offsets.

It may be that these findings are a result of chance pairings of study blocks with control blocks. Statistical techniques used indicated that the results would have occurred by chance no more than once in 20 times. This is strong evidence that the results reflect a consistent, not a chance, pattern.<sup>10</sup>

In conclusion, the data show that dead end, cul-de-sac, and L-type blocks have lower residential burglary rates than do more accessible control streets. The same conclusion cannot be made, at least within the bounds of this study, for T-type and offset streets.

#### CONCLUSION

This initial study has demonstrated that street layout does affect residential burglary rates in Minneapolis. Our findings indicate that less accessible streets, such as cul-de-sacs, dead ends, and L-type streets, have lower residential burglary rates. The findings are consistent with theory, conventional wisdom, and experiments in other cities.

This study also indicates Oscar Newman may be too pessimistic when he says:

"The creation of 'community of interest' cannot be accomplished simply by the trend noted above if (1) at least one member of the pair had a burglarized residence and if the residential burglary rate was lower for the study block, or if (2) neither member of the pair had a burglarized residence but the study block had a greater number of residences (and hence a greater probability of being burglarized).

"Assuming there was no true trend, 50 percent of the pairs should coincidentally fit the trend and 50 percent should not. As it was, 65 percent of the pairs fit the trend:

	<u>Number of pairs of blocks fitting the trend</u>	<u>Number of pairs of blocks not fitting the trend</u>
Dead end	9	2
Cul-de-sac	7	6
L-type	<u>8</u>	<u>5</u>
Total	24	13

"Because we sampled blocks instead of looking at all blocks in Minneapolis, we may have, by chance, oversampled pairs fitting the trend. If we had looked at all block pairs in the city, maybe only 50 percent would have fit the trend. The probability (expressed as a proportion) of our drawing a random sample with 65 percent of the pairs fitting the trend when, overall, only

50 percent of the pairs fit the trend is less than one in twenty (that is, significant at the .05 level based on a binomial probability test).

"Setting up zones outside the dwelling for the collective use of proxemic dwellers, it involves as well the setting up of covenants . . . among groups of residents to somehow guarantee the nature of tenant occupancy and commitment to shared values."<sup>11</sup>

Physical layout of streets can affect crime without burdensome legal agreements and without any more resident homogeneity than already exists.

Present findings, coupled with information from other sources, suggest redesign of streets, properly used, is a promising strategy for crime reduction in urban areas. Application of this strategy should be limited to areas where the crime problems warrant it and should be planned in conjunction with safety and traffic needs of the area. In addition, until further analysis is complete, redesigning streets to control crime should be used on an experimental basis. This limited use will permit analysis of crime data to measure the impact of redesign on crime as well as to measure the impact of redesign on other aspects of neighborhood life.

Further study will concentrate upon street layout at the scale of census tracts. The study will use graph theory and network analysis to calculate indices of accessibility for each tract in Minneapolis. One-way streets, physical barriers to travel, variation in number of lanes and variation in volume of traffic will be some aspects of accessibility considered. Multiple regression will be the principal technique for determining the direction and strength of association between crime rates and accessibility indices. Further study may also expand the sample of individual streets so the interaction between street layout and variables like residential/commercial mix and single/multifamily mix can be estimated.

## NOTES

1. This research was supported by grant #76DF050005 awarded to the Governor's Commission on Crime Prevention and Control by the Law Enforcement Assistance Administration. Points of view and opinions stated in this report are those of the author(s) and do not necessarily represent the official position or policies of the Governor's Crime Commission or the Law Enforcement Assistance Administration.
2. Oscar Newman, *Defensible Space* (New York: Collier Books, 1973), pp. 60-62.
3. Based on interviews in 1976 with 45 burglars imprisoned in State of Minnesota institutions.
4. Oscar Newman, "Community of interest--design for community control," *Architecture, Planning and Urban Crime* (London: NACRO, 1974), pp. 26-35.
5. Newman, *Defensible Space*, pp. 60-62.
6. *Six Months Experience--Berkeley Traffic Management Plan* (City of Berkeley: Deleau, Cather and Company, 1976), pp. ii-v.
7. Based on discussions with Minneapolis police officers.
8. Blocks were classified into a design type according to whichever of the block ends was least accessible.
9. Residential burglaries recorded by the Minneapolis Police Department from July 1974 through June 1975.
10. The techniques compared members of 37 pairs of blocks. Each pair consisted of a study block and its corresponding control block. The relationship between the members of each pair was classified as supporting.
11. Newman, "Community of interest--design for community control," *Architecture, Planning and Urban Crime* (London: NACRO, 1974), p. 9.

SECTION V  
THE METHODS OF EVALUATION

Preceding page blank

RECIDIVISM AND RESEARCH DESIGN: LIMITATIONS OF  
EXPERIMENTAL-CONTROL RESEARCH

Robert Martinson and Judith Wilks

A critical issue in current discussions of evaluation research in criminal justice is how knowledge can be cumulated from existing information and brought to bear upon crucial questions of planning and policy. This area--sometimes called "evaluation of evaluation"--presents novel problems of methodology, analysis, and summarization of findings and results.

In October of last year, in a report titled "Knowledge in Criminal Justice Planning," we described in a preliminary way a general procedure which we thought would prove useful in combining the results of experimental evaluations of treatments with the results of the standard ("after-only") research collected from annual reports of criminal justice agencies. These two bodies of data have hitherto been developed side by side with little mutual impact and it was our belief that their proper combination might provide both a check on results and a substantial increase in analytic capacity.

Our preliminary report was based on 3,005 recidivism rates taken from 162 documents that had been coded at that time. The present paper is based on 7,341 recidivism rates taken from over 470 documents. We are still in the editing process for a good proportion of our data and thus we are limited in this brief paper to some preliminary comparisons of experimental and nonexperimental research. The tables we will present today raise a host of questions which we have not yet had the time to explore. We do think, however, that they are questions which it would be hazardous to ignore in a period when experimental research is proliferating in the field of criminal justice.

Judith Wilks and myself are especially concerned with this issue given the monumental inflation and widespread misinterpretation of our previous work--The Effectiveness of Correctional Treatment--which was based on only a few hundred experimental-type studies and in which we proceeded on the project-by-project level which, unfortunately, dominates the field to this day.

Indeed, the most important departure in our present research has been to establish a unit of analysis which cuts across projects or individual studies and at the same time permits making use of the differences among studies. In place of the project-by-project approach which uses the individual study as the unit of analysis, this procedure, we thought, could

lay the basis for accumulating and comparing study findings. This would permit study characteristics to play a role in explaining findings as well as leaving open the opportunity to sharpen and broaden the questions that can be answered with the data.

#### THE METHODOLOGY OF THE SURVEY

For a more complete description of the search procedure, the classification of documents received, and the variables which we have coded, it is necessary to read our preliminary report. Let me just briefly define a few terms which will help clarify what we have done.

The two key concepts are the "batch" and the "computable recidivism rate."

a. Batch. A "batch" is any number of persons at some specifiable location in the criminal justice system for whom a "proper" recidivism rate is computable. A proper recidivism rate must specify what proportion of a batch are recidivists. We use the term "parent batch" to conveniently refer to a universal set which contains two or more batches. For example, a universal set of, say, 1,000 male and female parolees may be broken into one batch of 800 male parolees and one batch of 200 female parolees. Each of these batches are coded as "exclusive" since they exhaust the parent batch and contain no members in common. Once the "exclusive" batches have been selected, the remaining batches constructed from a universal set are coded "nonexclusive."

All tables in this paper are based on exclusive batches with an N of 10 or more, located in the postadjudicatory area of criminal justice. That is, the batches begin at probation and include shock probation, group homes, standard imprisonment and aftercare, special forms of release from imprisonment including study, work, and halfway house release, special treatment on parole, and all the way through to "max out" or release from imprisonment without supervision.

b. Recidivism Rate. The primary unit of analysis in this survey is the computable recidivism rate. Each such rate specifies what proportion of any batch shall be identified as "recidivists" according to some operational definition of recidivism utilized by the researcher. Such an operational definition will normally specify the length of time which the batch was followed up, and, in general, the procedures used by the researcher to decide whether any given member of the batch does or does not fall into the category of "recidivist."

A batch may have more than one recidivism rate. For example, a batch of parolees may have a suspension rate, an arrest rate, a technical violation rate, and a return-to-prison-with-new-conviction rate. Additional batch rates may be computable beyond those actually computed by the authors of the document. When such is the case, the decision-rule is to compute as many rates as is possible, i.e., to exhaust the rate-generating capacity of the document.

Let me say, parenthetically, that documents vary widely in the degree to which they produce batches and therefore rates. The typical experimental-control study might produce as few as two batches--one experimental batch which received the treatment being evaluated, and perhaps one control batch which received standard processing. On the other hand, we recently received a computer printout from Marc Neithercutt who is Program Director of the Uniform Parole Reports. This document crossclassifies parolees by year of release, State, sex, and current offense. It contains 1,123 exclusive batches and will produce over 9,000 computable recidivism rates when it is entered into our system. Since each document has a unique identifier, we will be able to compare the rates produced by the Uniform Parole Reports with the rates produced by other sources of parole data.

This general procedure generates a large number of exclusive recidivism rates. For each of these rates, it is possible to code any number of additional items of information. The items and codes which we have used so far were derived empirically with the aim of maximizing the information-generating capacity of the body of documents produced by our search procedures. Our data analysis so far has concentrated on methodological variables such as research design, time in followup, whether the research used a sample or a population, the definition of recidivism used, and more detailed information on treatment. We have coded about two additional IBM cards of data including descriptive information on batches (age, sex, current offense, criminal history), the funding and publication source of the document, and so forth. We have not yet been able to run this data.

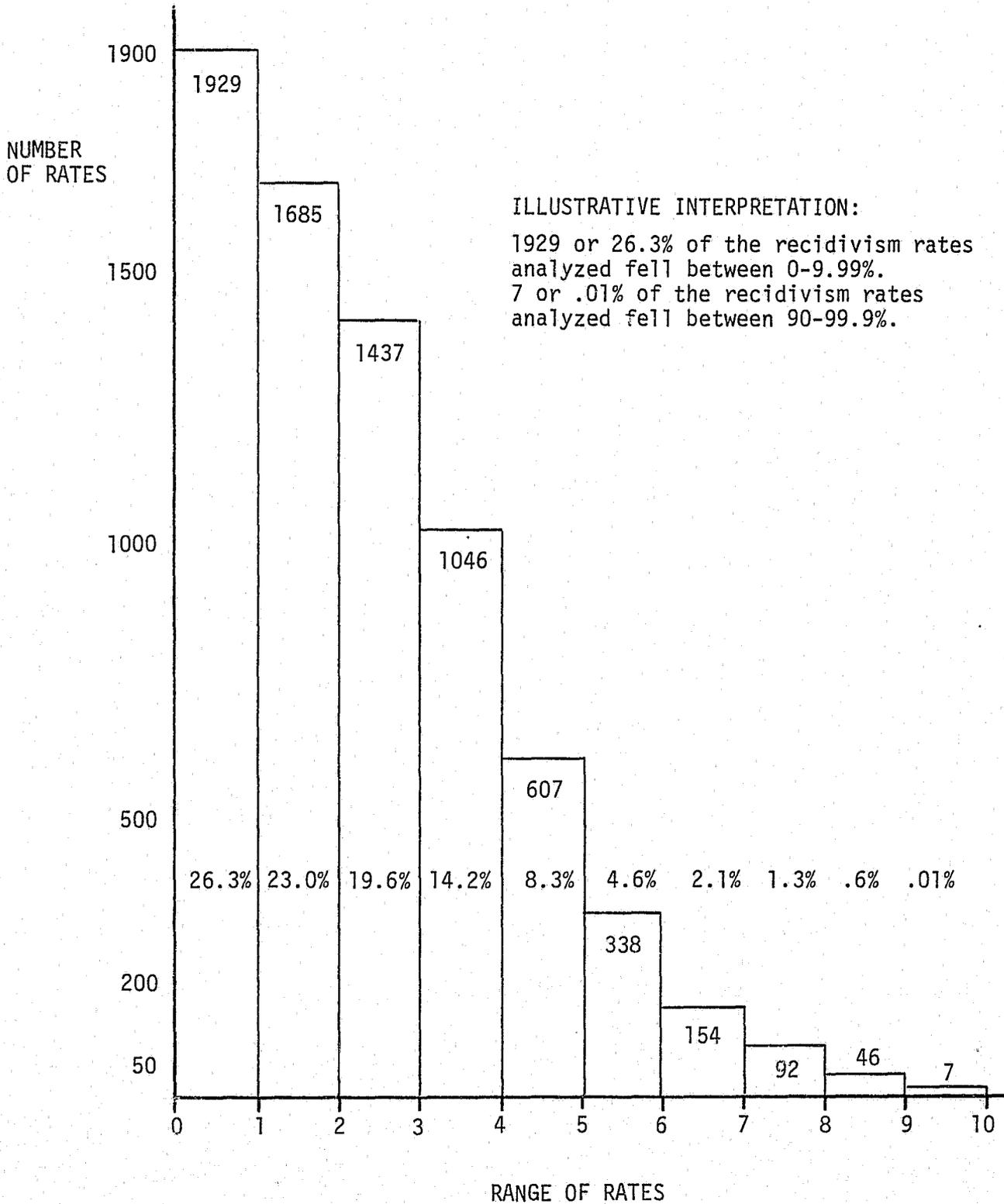
#### THE MEAN OF THE RECIDIVISM RATES

Figure 1 displays the distribution of the 7,341 rates which we have so far coded. To underline what we have done, let us call this the "fruit salad" distribution of recidivism rates. We have deliberately mixed together oranges, apples, pears, bananas, and even a little cumquat. We have permitted every methodological and substantive influence to modify the rates as much as the body of research examined will permit.

The mean of the rates of the "fruit salad" distribution is 23.3 averaging over the last 40 years or so. This mean has not changed by as much as 5 percentage points from the time we coded the first thousand or so rates even though we have coded additional sets according to convenience. This mean of the rates is below the general estimate of one-third made by Daniel Glaser a decade ago; and it is well below the estimate of about two-thirds that one frequently finds mentioned in textbooks, in the media, and even by criminal justice professionals. According to this distribution, a rate of 60 percent or above is not impossible to find, and is in fact found in the research we have examined about four times out of one hundred. Such a rate is real, but it is rare. The idea that the recidivism rate in the United States is high seems somewhat incompatible with this distribution.

For those who wish to argue that such rates would be higher if we caught more offenders, we would merely suggest that they substitute for the word recidivism throughout this paper, the term "system re-processing rate." The mean of the rates would then be the rate at which the system

FIGURE 1. Distribution of recidivism rates



of criminal justice as it is reflected in these studies tends to reprocess the offenders over which it has jurisdiction.

## RESEARCH DESIGNS

Figure 2 is based on the categories of research design used to describe studies in The Effectiveness of Correctional Treatment. We have added categories 19 through 22 to reflect the extension of our survey to include the after-only research typically produced by State divisions of research and by other governmental agencies. It is noteworthy that a total of 5,638 (or 76.8 percent) of the recidivism rates are found in the after-only category which might serve as a rough index of the available information which we did not utilize in our previous work. In all of the tables which we present today, we have collapsed the research designs into three basic categories. Category 1 includes matched or random allocation of subjects and is designated "random" in the tables. Category 2 includes the remainder of the experimental or quasi-experimental research we have coded and it is designated "nonrandom" in the tables. Category 3 is the after-only without control group type of research almost all of which is ex post facto.

## ISSUES IN EXPLANATION

Table 1 presents the distribution of recidivism rates by type of research design for varying periods of followup. The expected pattern is that rates based on longer periods of followup will be higher than those based on shorter periods with a tapering off of the rates as time increases. This pattern holds for the after-only category, but not for the random, nonrandom, or "total set" categories.

Table 2 partially supports a finding presented in the Preliminary Report, namely that samples generate higher recidivism rates than populations. There is an exception, however. In the nonrandom category of research design, populations generate higher rates than samples.

Looking now at the combined effects of research design category and population versus sample, the expected pattern for recidivism rates by months of followup holds for the after-only category of research which employs populations, but not for the random and nonrandom categories which employ populations (table 3). In the case of samples, this expected pattern does not hold absolutely for any of the categories of research designs.

Two other tables are included. Table 4 shows the distribution of recidivism rates of the various research design categories for various definitions of recidivism. Table 5 shows the distribution of rates by design and location in the Criminal Justice System. Neither of these tables, substance aside, illustrates the existence of a clear and systematically patterned relationship between the recidivism rates associated with the various categories of research design.

Our aim today is not to present you with answers, but to ask you to help resolve the conundrum which we have tried to briefly illustrate. Can we anticipate that a different method of controlling for research design would produce more systematic patterns? Can we anticipate that controlling

*FIGURE 2. Number of recidivism rates distributed by research designs classified according to degrees to which they meet criteria of internal validity*

Type of design				Method of obtaining subject pool	Method of allocation to experimental and control groups	Research designs
Pure	Ex post facto	Simulated				
1 (N:134)	5 (31)	9 (0)	random	Probability sample	Matched or random allocation	Classical design and after-only with control or comparison group
2 (666)	6 (274)	10 (19)		Nonprobability sample		
3 (6)	7 (127)	11 (2)	nonrandom	Probability sample	Nonmatched or nonrandom allocation	
4 (48)	8 (284)	12 (70)		Nonprobability sample		
13 (0)	15 (4)	17 (0)		Probability sample	Not applicable	Before-after without control
14 (4)	16 (34)	18 (0)		Nonprobability sample		
19 (12)	20 (4,725)		after-only	Probability sample	Not applicable	After-only
21 (12)	22 (889)			Nonprobability sample		
Total 882	Total 6,368	Total 91				

TABLE 1.--Mean recidivism rates by months in followup

Months in followup	Total set				Random				Nonrandom*				After-only			
	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%
1-6	15.3	1,367	.94	19	17.9	260	.88	23	20.1	125	.82	22	14.0	982	.96	17
7-12	21.2	2,522	.77	35	32.9	269	.71	23	27.1	208	.67	36	19.4	2,055	.76	36
13-18	32.0	245	.66	3	35.2	156	.67	15	33.1	36	.43	4	21.7	53	.56	1
19-24	25.7	1,192	.69	16	24.3	219	.83	19	24.6	75	.77	13	26.1	893	.64	16
25-36	30.2	520	.56	7	33.2	70	.71	6	41.3	57	.43	10	28.1	393	.53	7
37-60	28.9	830	.62	11	22.6	79	.83	7	44.8	73	.39	13	27.9	673	.61	12
Over 60	28.2	615	.61	8	31.4	71	.74	6	7.8	3	.74	1	27.9	541	.58	10
Unknown	27.6	49	.82	1	45.5	10	.71	1	29.5	2	.24	1	22.6	37	.77	1
Total	23.2	7,340	.75	100	26.9	1,124	.81	100	27.2	579	.67	100	22.0	5,637	.73	100

\*Includes before-after designs.

$\bar{X}$  = Mean recidivism rate.

CV = Coefficient of variability.

% = Percent of category.

TABLE 2.--Mean recidivism rates by population or sample

	Total set				Random				Nonrandom*				After-only			
	$\bar{X}$	N	$\frac{C}{V}$	%	$\bar{X}$	N	$\frac{C}{V}$	%	$\bar{X}$	N	$\frac{C}{V}$	%	$\bar{X}$	N	$\frac{C}{V}$	%
Population	20.9	4,927	.74	67	19.5	156	.91	14	34.2	72	.51	12	20.8	4,699	.74	83
Sample	28.2	2,414	.72	33	28.1	968	.79	86	28.5	507	.69	88	28.1	939	.65	17
Total	23.3	7,341	.75	100	26.9	1,124	.81	100	29.2	579	.67	100	22.0	5,638	.73	100

\*Includes before-after designs.

$\bar{X}$  = Mean recidivism rate.

CV = Coefficient of variability.

% = Percent of category.

TABLE 3a.--Recidivism rates for populations by months in followup

Months in followup	Total set				Random				Nonrandom*				After-only			
	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%
1-6	13.6	888	.96	18	29.5	2	.24	1	41.0	31	.40	43	12.6	855	.93	18
7-12	18.4	1,799	.75	36	34.5	7	.71	3	18.5	5	.73	7	18.4	1,787	.74	38
13-18	16.9	58	.70	1	6.7	18	.64	12	--	--	--	--	21.5	40	.53	1
19-24	23.3	599	.72	12	81.2	3	.14	2	19.5	14	.77	19	23.1	582	.70	12
25-36	27.9	346	.51	7	16.7	18	.67	12	39.9	13	.35	18	28.1	315	.50	7
37-60	27.0	636	.60	13	14.9	54	.76	35	34.5	9	.46	13	28.0	573	.58	12
Over 60	27.3	587	.59	12	23.6	54	.76	35	--	--	--	--	27.7	533	.58	11
Unknown	12.2	13	1.11	1	--	--	--	--	--	--	--	--	12.2	13	1.10	1
Total	20.9	4,926	.74	100	19.5	156	.91	100	34.2	72	.51	100	20.8	4,693	.74	100

\*Includes before-after designs.

$\bar{X}$  = Mean recidivism rate.

CV = Coefficient of variability.

% = Percent of category.

TABLE 3b.--Recidivism rates for samples by months in followup

Months in followup	Total set				Random				Nonrandom*				After-only			
	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%
1-6	18.5	479	.88	20	17.8	258	.89	27	13.2	94	.70	19	23.8	127	.82	14
7-12	28.0	723	.72	30	30.8	252	.71	26	27.5	203	.68	38	25.9	268	.74	28
13-18	36.6	187	.57	8	38.9	138	.58	14	33.1	36	.43	7	22.2	13	.67	1
19-24	28.0	593	.66	25	23.5	216	.87	22	25.8	61	.77	12	31.6	316	.50	34
25-36	34.8	174	.60	7	38.9	52	.62	5	41.8	44	.45	9	28.2	78	.62	8
37-60	35.1	194	.62	8	39.3	25	.53	3	46.2	64	.38	13	27.4	105	.77	11
Over 60	47.7	28	.50	1	56.3	17	.36	2	7.8	3	.73	1	44.5	8	.45	1
Unknown	33.1	36	.69	1	45.5	10	.71	1	29.5	2	.24	1	28.3	24	.60	3
Total	28.2	2,414	.72	100	28.1	968	.79	100	28.5	597	.69	100	28.1	939	.65	100

\*Includes before-after designs.

$\bar{X}$  = Mean recidivism rate.

CV = Coefficient of variability.

% = Percent of category.

TABLE 4.--Mean recidivism rates by definition of recidivism

	Total set				Random				Nonrandom*				After-only			
	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%
100% minus suc- cess rate	39.4	419	.53	6	55.1	47	.41	4	49.2	30	.37	5	36.5	342	.55	6
Short of arrest	19.8	967	.80	13	24.5	151	.91	14	25.0	65	.69	11	18.5	751	.76	13
Short of conviction	30.3	598	.73	8	30.8	380	.74	34	31.8	71	.58	12	28.1	147	.79	3
Short of prison sen- tence or return	22.5	391	.72	5	28.3	42	.56	4	17.8	96	.93	17	23.2	253	.68	4
Prison sen- tence or return-- old charge	17.8	1,035	.81	15	23.7	136	.88	12	24.5	85	.76	15	16.3	864	.75	15
Prison sen- tence-- new charge	11.7	639	.88	8	20.7	60	.70	5	7.8	12	.83	2	10.9	567	.86	10
Prison--old or new charge	25.2	3,237	.64	44	21.5	308	.84	27	34.8	220	.51	38	24.9	2,709	.63	48
Unknown	4.5	4	--	1	--	--	--	--	--	--	--	--	4.5	4	--	1
Total	23.3	7,340	.75	100	26.9	1,124	.81	100	29.2	579	.67	100	22.0	5,637	.73	100

\*Includes before-after designs.

$\bar{X}$  = Mean recidivism rate.

CV = Coefficient of variability.

% = Percent of category.

TABLE 5.--Mean recidivism rates by location in the criminal justice system

	Total set				Random				Nonrandom*				After-only			
	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%	$\bar{X}$	N	C V	%
Standard probation	21.2	2,116	.80	29	22.4	605	.83	54	26.6	101	.67	17	20.3	1,410	.79	25
Jail sentence and probation	18.7	142	.86	2	16.2	6	.72	1	49.5	2	.14	1	18.4	134	.87	2
Partial physical custody --preprison sentence	27.1	215	.78	3	32.4	34	.72	3	25.9	37	.81	6	26.1	144	.79	2
Imprisonment and standard aftercare	23.6	4,129	.71	55	35.4	195	.64	17	31.0	266	.65	46	22.5	3,668	.69	65
Work-study--furlough release	19.7	228	.97	3	21.7	156	.91	14	12.5	10	1.30	2	15.8	62	1.10	1
Partial physical custody --postprison sentence	25.9	209	.72	3	35.0	22	.57	2	26.2	86	.71	15	23.7	101	.75	2
Early release from prison	22.3	18	.77	1	--	--	--	--	26.5	5	1.01	1	21.4	13	.64	1
Parole with special treatment	30.1	213	.61	3	29.7	73	.70	6	32.9	71	.51	12	27.7	69	.60	1
"Max out"	49.4	70	.55	1	69.7	33	.26	3	54.5	1	--	--	30.6	36	.66	1
Total	23.3	7,340	.75	100	26.9	1,124	.81	100	29.2	579	.67	100	22.0	5,637	.73	100

\*Includes before-after designs.  $\bar{X}$  = Mean recidivism rate. CV = Coefficient of variability.

% = Percent of category.

for offender characteristics will establish a patterned relationship which will allow us to decide which design provides the most useful information given the current state of knowledge about criminal justice? What about controlling for risk? Source of funding? Time in treatment? And we could also look into more detailed characterization of treatment such as counseling or behavior modification or job training.

Would controlling for these or any number of other variables assist us in deciding which type research design category is going to be productive of the most reasonable standards for evaluation of criminal justice outcomes? We can control for all the variables listed above and then some. We are in fact seriously soliciting from you ideas about which variables and which research designs should be given priority consideration. There is a massive amount of information now available, generated by countless researchers at the cost of millions of dollars, and we want to use as much of it as possible. But, more importantly, we want to try to answer critical questions such as: How can the criminal justice dollar best be spent to protect the public? How can the research dollar best be spent to answer this question and others that are critical for policy? That is why we are asking for your help today.

## THE THREE FACES OF EVALUATION: WHAT CAN BE EXPECTED TO WORK

Herbert C. Quay  
University of Miami

While the adequacy of the research design is clearly related to the confidence with which the outcome of any intervention study can be interpreted, an adequate design is a necessary but not sufficient condition for the evaluation of any treatment. The dangers of overemphasis on design, analysis, and outcome to the exclusion of other criteria of adequacy is, unfortunately, manifest in the widely cited review of correctional treatment by Martinson (1974).

This review has had considerable impact on correctional administrators although its conclusions have been questioned in a recently published critique (Palmer, 1975). What has been obscured in the controversy over "What Works" are the dangers of over reliance on restricted notions of what constitutes "good evaluation." Both Martinson and his critics have failed to consider the studies reviewed on dimensions other than experimental design.

Before we can legitimately conclude that a method of correctional treatment doesn't work, there is a great deal more we need to know about it beyond the design used to evaluate it and its outcome. Most of this necessary additional information relates to what may be called the integrity of the treatment program.

As we hope to demonstrate, the integrity of any intervention is multi-dimensional matter and we do not pretend to have identified all of the possible elements. We do, however, suggest that there are four areas related to intervention integrity and that each area has identifiable subaspects.

### THE NATURE OF, AND THE EMPIRICAL BASIS FOR, THE TREATMENT

Specificity with Which the Intervention Can Be Conceptualized. Crucial to the integrity with which an intervention can be carried out is the clarity of its conceptual basis. Addressing this issue is basically asking: What, exactly, is the treatment? Put another way: How accurately can we describe (and perhaps measure) the independent variable? The more specific the conceptualization, the greater the ease, all else being equal, of the implementation, and, perhaps more importantly, the greater the ease with which others can replicate the treatment procedures.

The Empirical Basis of the Intervention. Assuming that the intervention can be reasonably well specified, how well it is grounded in empirical

research became the next point of concern. Are its operations and techniques based on research findings or is the empirical base for its procedures lacking? While the validation of procedures in the laboratory does not guarantee their successful translation into the natural setting, the firmer the empirical base the more likely is a successful transition to the "real world."

The Proven Utility of the Intervention in Less "Complex" Settings. The correctional setting is a complex one, with frequently conflicting demands on its staff (custody versus treatment) and its clients (institutional adjustment versus behavior change). Before a treatment is imported into such an inhospitable matrix of forces, it ought to have demonstrated its utility in more benign climates. An intervention which already carries with it evidence for its utility in the consulting room, clinic, or classroom would seem a better bet (but not a sure thing) for the correctional setting than one which did not.

#### THE SERVICE DELIVERED

A second group of factors involved in program integrity have to do with the service actually delivered to clients: Does what actually happens meet the specification of the treatment?

Monitoring Program Elements. Of crucial importance to program integrity is the answer to the question--what actually happened? One cannot take programs at face value and some accounting of what actually went on is clearly necessary. If counseling is the treatment one needs to know if counseling sessions were actually held, how well were they attended, and did what went on in the sessions constitute counseling as specified by the particular model. These are critical questions, because without generally affirmative answers there really is no intervention to evaluate.

Duration of the Service. Some intervention may specify duration, at least in general terms, a priori. Others make duration a function of continuously or periodically measured results. However duration is arrived at, the extent to which this criterion is met must be considered in assessing program integrity. Rome was not built in a day.

Intensity of the Service. The prescribed length of a counseling session, or of an academic class, or a Synanon game, may be in many ways arbitrary and more a function of opinion than of fact, but some limits must be respected. Ten minutes per day of remedial reading is not likely to increase reading skills very much, let alone reduce recidivism.

#### PERSONNEL

A third aspect of program integrity has to do with those who are delivering the service. Since most treatment techniques are dependent upon who is doing the treatment, questions related to personnel may well be the most important in evaluating program integrity. There are at least four identifiable subareas related to personnel.

Degree of Expertise. It is probably best if the personnel implementing the treatment program know what they are supposed to be doing before

beginning the intervention; this applies equally to planners, administrators, service deliverers, and researchers. While personnel qualifications are sometimes difficult to specify, one would not generally expect accountants to build bridges. Prior education, reflected usually by academic degrees, may not provide ironclad guarantees of expertise but such credentials are better than no evidence for competence at all.

Amount of Training Provided. Many correctional interventions simply cannot be implemented if all personnel are required to have a high degree of prior expertise. In cases where personnel have few skills present at the outset, the nature and amount of training become critical. Intelligent and motivated correctional officers clearly can become reasonably good at personal counseling but not as a result of three 1-hour lectures, only two of which they attended. Again, performance may be more important than training, but most people would rather, if absolutely necessary, be operated on by a nurse than an orderly.

Degree of Supervision. Almost all human service providers receive some degree of help and guidance in their efforts, if only informally through meetings and discussions with peers. When the service is to be provided by those with relatively little experience, training and supervision become most important. How often, how intense, how utilized--the more information we have about each of the above, the better.

#### THE MATCH OF TREATER, TREATMENT, AND TREATED

This is a relatively new area that is still under exploration which can contribute to program integrity. Whether a rigid, authoritarian person can ever be trained to become a nondirective counselor is problematical, so treatment agents with these personal proclivities might better work in another modality. On a more sophisticated level, there is some evidence to suggest that certain types of treaters do better with certain types of clients. At this juncture, one cannot expect staff-client matching in most instances, but where it occurs it likely contributes to program integrity.

#### PROGRAM INTEGRITY AND THE EVALUATION OF CORRECTIONAL TREATMENT

It is unfortunate, but nevertheless true, that many of the treatment studies in corrections have provided little information relative to program integrity. A notable exception, however, is the detailed report on a large scale study of the efficacy of group counseling by Kassebaum, Ward, and Wilner (1971).

This project is referred to twice in the Martinson (1974) article.<sup>1</sup> Under the heading of "Group Counseling" we find: "Two (studies) (Kassebaum, 1971; Harrison; 1964) report no long-lasting effects" (p. 31). Under "Transforming the institutional environment," Martinson states: "Another study by Kassebaum, Ward, and Wilner (1971) dealt with a program which had been able to effect an exceptionally extensive and experimentally rigorous transformation of the institutional environment. This sophisticated study had a follow-up period of 36 months, and it found that the program had no significant effect on parole failure or success rates" (p. 33).<sup>2</sup>

It is unquestionably true that no effects of the treatment were demonstrated and that the research design was adequate. Yet, a reading of the report reveals much about why, in this particular instance, group counseling did not work.

An examination of this excellent report from the point of view of what it says about many of the elements of program integrity proposed above is revealing. With respect to the conceptualization of the treatment, in discussing "what is group counseling," the authors (Kassebaum, Ward, and Wilner) report:

"Nonetheless, Fenton's description of the interactional processes of the sessions (what goes on between group members and the leader) is couched in very general terms, and the theoretical bases on which group counseling is built are not clearly spelled out. The aims of group counseling are not easily operationalized, nor is it described in terms that lend themselves to the precise analysis of group structure or process" (p. 59).

The utility of the process, as judged by those actually involved in it in the California Department of Corrections, is revealed in data provided in table 3.1 on page 64. Here it is reported that only 40 percent of group counselors agreed that "Group counseling induces personality change" and only 30 percent agreed that "Inmates from group counseling violate parole less." This latter finding means that less than one-third of practicing group counselors felt that the treatment would affect the major dependent variable of the study (recidivism) and the only dependent variable subsequently of interest to Martinson (1971).

Despite these problems in conceptualizing the treatment and the expressed doubts about its efficacy as practiced by the California Department of Corrections, it was nevertheless possible to study its effects. As stated by Martinson et al.: "The limitation in conceptual precision, however, does not prevent us from studying the effects of group counseling participation when an appropriate research design is employed" (p. 59).

What can be learned about the service actually delivered to the clients and the personnel involved in that delivery? According to the authors:

"Operationally, group counseling means that ten or twelve inmates meet one or two hours per week under the guidance of a lay group leader. Some leaders are administrative personnel, caseworkers, teachers, guards, or clerical and technical staff workers; others are therapeutic specialists (physicians, social workers, and psychologists). Nonprofessional personnel in group leader roles, to some extent, are trained and supervised by the group counseling supervisor in each prison. In most cases, these supervisors hold B.A. degrees and have received graduate training in social work" (p. 59).

We do not intend to convey the impression that nonprofessionals cannot do group therapy, but if such is the case, training, initial and continued, is crucial. Kassebaum, Ward, and Wilner (1971) address the issue of training in some detail. Initially they report that: "Although one community living unit and the mandatory group counseling programs had begun, the supplemental training for group leaders was not yet in operation" (p. 84).

This is a rather important point and the authors go on to discuss training in considerable detail in chapter IV. Some reasonable judgments can be made about the quality of training from the following quotations:

"Two training programs for group counselors were conducted at Men's Colony--East. During the first year and one half of the institution's operation the in-service training of group counselors was provided through a series of one-hour monthly lectures by the supervisor of group counseling. In addition, sessions were scheduled in which group leaders raised specific questions that arose from problems in conducting their groups. These meetings were poorly attended, and it was our impression that many of the men who did come seemed to be apathetic and disinterested" (p. 86).

"During the study period the counseling coordinator was seldom consulted by the new group leaders. Although a small library of books on counseling and therapy was housed in the coordinator's office, it was little used at the time of the study. No list of available titles had been distributed to group leaders" (p. 87).

"It was our view that at the time the study began the instruction given to correctional personnel in counseling techniques was limited and was generally not regarded as very helpful to the leaders. It should be kept in mind that CMCE had just opened, and many of its staff were men entirely new to corrections. In-service training time was in short supply and heavily committed to the more immediate tasks of operating a new prison" (p. 87).

Reacting to this situation, the research project itself decided to offer supplementary training. Measures were taken to combat the indifference and absenteeism characterizing the counselor's response to the regular in-service counselor training. This training focused more on the personalities of the counselors than on the technical aspects of counseling and the researchers seemed to feel that something was gained.

The project staff's attempts to monitor the services delivered by direct observation were limited but revealing. In discussing the small groups, the researchers observe:

"Our observations and discussions with staff and inmates led us to conclude that the small groups were frequently beset with the following problems:

1. A tendency for superficiality, a lack of emotional involvement, and evidence of insincerity.

2. A tendency for talkative members to monopolize the discussion to the exclusion and boredom of others.

3. A feeling of frustration and a lack of confidence in leaders' or members' ability to 'do the job' without professional supervision.

4. A tendency to focus on stories and personal accounts that were not further analyzed or used for discussion but were used to provide competition for another inmate's account of his preprison experiences or exploits.

5. A tendency for staff members to permit periods of silence up to the length of the entire session because of their misinterpretation of 'nondirective counseling' or their own inability to elicit discussion instead of personal narratives and storytelling. In some cases this may reflect inadequate training, in other cases it reflects inadequate counselors.

However, as in the case of Group No. 3, some groups exhibited behavior which was, in the opinion of the observers, similar to therapy groups in noncorrectional settings where the leader is unobtrusive (sic) but in command of the situation, and his manner suggested relaxed self-confidence. The members spoke critically and spontaneously and gave evidence of trusting the leader and one another. Based on feedback from inmate interviews (which we consider later in this chapter), conversations with staff members, and the reports of the on-site research staff, there were some, but not many, groups (like No. 3) in which the conduct of the sessions approximated the goals of counseling set forth in the departmental training manual" (p. 123).

Since group stability had previously been suggested to relate to success, the researchers undertook to locate stable groups. They reported:

"We believed that Harrison's findings about the superior performance of stable groups remained open to further empirical examination, since no control groups were used, but we found that we could not identify any groups at CMCE that met his criteria of stability. (The CMCE group counseling coordinator in regard to this issue glumly remarked, 'We have a stable group if there's less than one leader change during a month.')

Men moved in and out of groups frequently as their jobs or institutional activities required, as they stopped attending counseling sessions, as the leaders changed, and the like. No group had the same leader for an entire year because of changes in work shifts, job assignments, vacation and other absences, meetings, etc." (p. 247).<sup>3</sup>

On the basis of the information provided one is hard pressed to conclude other than that the service actually delivered by minimally trained and inexperienced personnel was inadequate to the task--a view already expressed by other counselors in the system.

Now we may turn to the target population; the inmates, their nature, the perceptions of the process, and the methods whereby they were selected.

Chapter V begins:

"At the time our study was set for full operation at Men's Colony--East, 76 counseling groups had been established in the three quads under study. Twenty-three of the groups involved mandatory participation in small groups and 53 involved groups made up of voluntary participants. In addition, there were three community living groups made up of 50 men each, in which participation was mandatory" (p. 118).

One can conclude that most participation was frankly involuntary. The reasons for attendance by inmates are reported:

"Of particular interest to us was the finding that was based on the interviews conducted with men in prison and on parole, that the most consistently expressed view of group counseling was that its value was chiefly in satisfying the Adult Authority at parole hearings. Like class attendance in some universities, inmates felt that a participation in group counseling might not be a major factor in getting paroled, but a lack of participation was likely to be regarded negatively by the Adult Authority" (pp. 31-32).

As to the nature of the participants, we find:

"The composition of the inmate groups also varied. A few included inmates with several years experience at doing time together at San Quentin or some other prison; other groups were composed entirely of men who were strangers to each other and who were serving their first term in prison. Most groups were mixtures of these extremes" (p. 118).

Clearly, there was no attempt by the program managers to select, or even identify, those most likely to have benefited from the treatment, nor was there an attempt to compose the counseling groups in any systematic way.

What of the clients' perception of the treatment? On the basis of interviews with inmates the authors reported:

"Interviews conveyed the strong impression that relatively few inmates entered group counseling with the conviction that they were participating in a meaningful treatment program. The usual advice new inmates received from others was to the effect that counseling was not adequately nor honestly run, but that participation looked good to the Adult Authority, and, in fact, counseling was one of the

measurable items of an inmate's experience in prison (like school attendance, trade training, and disciplinary reports) that could be considered. Although participation may not help inmates to make parole, its absence, generally noticed by the Adult Authority, is often interpreted as a lack of interest in helping oneself and getting involved in the treatment program. For the Adult Authority, the record of length of participation in group counseling is a useful index of prisoner experience because it joins that relatively small list of activities that can be quantified and used in plus-or-minus fashion in determining parole eligibility" (p. 131).

Survey data obtained revealed such findings as: "Two-thirds agreed that correctional officers were not competent to run groups" (p. 136). Further, "There was agreement that men did not talk frankly (four out of five), and one-half of the respondents agreed that if too much was revealed it would be used against them" (p. 136).

It is obvious that group therapy as a ". . . group setting necessary for clients to feel free to discuss with security their own and each other's feelings and attitudes toward the situation in which they find themselves" (p. 59) was never accomplished.

Given the overall quality of the intervention, made abundantly clear by Kassebaum, Ward, and Wilner (1971) but totally ignored by both Martinson (1974) and Lipton, Martinson, and Wilks (1975), what reasonable person could have expected recidivism, or anything else, to have been reduced? Certainly not, as we have seen, those with prior experience in California prison group therapy. Finally, one wonders how this treatment could possibly be described as one "which had been able to effect an exceptionally intensive and experimentally rigorous transformation (italics ours) of the institutional environment" (Martinson, 1974, p. 33).

It is unlikely that any other research report available for review contained the wealth of information on program integrity provided by Kassebaum, Ward, and Wilner (1971). However, such does not negate the fact that conclusions about the effects of group counseling and milieu intervention should have been tempered by a consideration of the extent to which program integrity was seriously lacking in this program effort.

We do not wish to beg the question of outcome, even when narrowly defined in terms of recidivism--itself a term not always meaning the same thing in different studies. But to continue to ignore all aspects of the integrity of the treatment in arriving at conclusions about what does and doesn't work will be a major error. An error, unfortunately, with serious policy and practical consequences for corrections beyond those already suffered.

#### ASSESSING PROGRAM INTEGRITY

The development of techniques for the assessment of program integrity must have as high a priority as the development and routinizing (see Glaser, 1974) of evaluation. For purposes of evaluating the impact of treatment

training on line correctional staff, Johnson (1975) has developed a schedule to assess the integrity of a positive reinforcement programming type of correctional intervention. This schedule provides a quantitative estimate, through ratings of various program dimensions, of the integrity of an intervention based on the particular model. Field tests have suggested that the scale can be used reliably and further research with it is underway.

We urge the development of technique for the assessment of the integrity of programs based on other models. Without attention to this third face of evaluation valid conclusions as to the efficacy of various rehabilitative strategies will continue to elude us.

## NOTES

1. In the entire article, one finds very little concern for the nature of the service delivered as possibly affecting the outcome. The problem in implementation and staff attitudes reported by Zivan (1966) is reported and represents the only reference, in any detail, to program integrity problems in the entire review. Some greater attention is given to the nature of the target population in terms of amenability of the entire target group or some subgroup thereof (e.g., Adams, 1961, discussed on p. 29; Goldberg and Adams, 1964, discussed on p. 33; the Warren Studies, 1966a, 1966b, 1967, discussed in pp. 43-44). Yet amenability effects are, on the whole, disregarded in arriving at his conclusions (see also, Palmer, 1975).
2. In the subsequently published and much less widely cited book (Lipton, Martinson, and Wilks, 1975) this research referred to under a number of headings [e.g., Group Counseling (p. 224, pp. 236-237); Milieu Therapy (pp. 242-243, pp. 252-253, pp. 259-261)] none of these citations contains details of the project beyond a brief description of its design and results.
3. Despite the failure to find groups meeting the criteria of stability the authors analyzed outcome of what could be called "more" and "less" stable groups. No differences emerged.

## BIBLIOGRAPHY

- Glaser, D. Routinizing Evaluation Getting Feedback on Effectiveness of Crime and Delinquency Programs. Rockville, Maryland: National Institute of Mental Health, Center for Studies of Crime and Delinquency, 1974.
- Johnson, V. S. The Positive Reinforcement Programming Congruence Scales, Unpublished Manuscript, 1975.
- Kassebaum, G., Ward, D., and Wilner, D. Prison Treatment and Parole Survival: An Empirical Assessment. New York, Wiley, 1971.
- Lipton, D., Martinson, R., and Wilks, J. The Effectiveness of Correctional Treatment. New York, Praeger, 1975.
- Martinson, R. What works? Questions and answers about prison reform. The Public Interest, 1974, No. 35 (Spring), pp. 22-54.
- Palmer, T. Martinson revisited. Journal of Research in Crime and Delinquency, 1965 (July), pp. 133-152.

## MEASUREMENT OF CRIME PATTERN MODIFICATION: AN EXAMPLE, GUN CONTROL\*

Stuart Jay Deutsch  
Associate Professor  
School of Industrial and Systems Engineering  
Georgia Institute of Technology

### INTRODUCTION

The goals of many criminal justice programs center around a reduction of criminal behavior, typically in terms of recidivism or crime reduction. An approach to the measurement of the effects of program activities on goals of crime reduction is the subject matter of this paper. As such, it is not intended to bring the reader to operational capability but rather to overview the major considerations while developing a conceptual framework for the analysis. For more complete description of the individual efforts that are overviewed here, see [1,2,3,4,5].

In the first section, a univariate time series approach to model past crime occurrence is briefly discussed. The application of these models in an intervention analysis framework is then described, both with respect to the conceptual approach as well as the statistical formulation of the problem. Viability of the models in post evaluation and perhaps more importantly in "on-line process" monitoring is looked at with a sensitivity analysis. Lastly, a summary example of the potential of the approach is illustrated via the evaluation of the effects of the Massachusetts' Gun Control Law on gun-related crimes.

### UNIVARIATE CRIME MODELING--A REVIEW

Viewing monthly historical crime data as a realization of an underlying stochastic process, models have been developed for each index crime in 10 major metropolitan areas [2]. The general model class utilized is the multiplicative autoregressive-moving average models proposed by Box and Jenkins [1].

$$\phi(B^S) \phi(B)(\nabla^S)^D \nabla^d Z_t = \theta(B^S) \theta(B) a_t$$

---

\*This work was performed under grant number 75NI-99-0091 from the National Institute of Law Enforcement and Criminal Justice. Points of view or opinions stated in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice.

where  $\phi(B^S) = (1 - \phi_{1.S}B^{1.S} - \phi_{2.S}B^{2.S} - \dots - \phi_{p.S}B^{p.S})$ ,

$$\phi(B) = (1 - \phi_1B^1 - \phi_2B^2 - \dots - \phi_pB^p),$$

$$\theta(B^S) = (1 - \theta_{1.S}B^{1.S} - \theta_{2.S}B^{2.S} - \dots - \theta_{q.S}B^{q.S}),$$

$$\theta(B) = (1 - \theta_1B^1 - \theta_2B^2 - \dots - \theta_qB^q),$$

$$\nabla = (1 - B)$$

and  $Z_t$  is the reported level of crime at time  $t$ ,

$\Phi, \phi$  are seasonal and nonseasonal autoregressive parameters respectively,

$\Theta, \theta$  are seasonal and nonseasonal moving average parameters,

$D, d$  are integer values typically 0, 1, or 2,

$s$ , is an integer value reflecting the seasonal lag,

and  $a_t$  are unexplainable deviates or residuals assumed to be normally and independently distributed with a zero mean and constant variance ( $a_t \sim \text{NID}(0, \sigma^2)$ ).

The building of an appropriate empirical-stochastic model is accomplished by a three-stage iterative procedure of identification, estimation, and diagnostic checking. Identification refers to the isolation of a tentative model form from the general model class based upon data characteristics. Estimation refers to the numerical evaluation of the parameters of the tentative model and diagnostic checking refers to the testing of the fitted model so as not only to discover inadequacies but also to give direction for the updating of the form of the fitted model to obtain a new tentative identification.

For each of the cities modeled [2], crime types of homicide and forcible rape were described by a random walk or (0,1,1) process,

$$Z_t = a_{t-1} - \theta_1 a_{t-1} + a_t.$$

The index crimes of robbery, burglary, larceny, vehicle theft, and assault were found to be adequately described by a seasonal-nonseasonal form,  $(0,1,1) \times (0,1,1)_{12}$ ,

$$Z_t = Z_{t-1} + Z_{t-12} - Z_{t-13} - \theta_1 a_{t-1} - \theta_{12} a_{t-12} + \theta_{12} \theta_1 a_{t-13} + a_t.$$

## INTERVENTION ANALYSIS

The models of crime occurrence previously described can be utilized to project past observed occurrences to develop future forecasts of crime activity. The forecasts from these models are best in a statistical sense since they possess the property of minimum mean square error. A complete description of the forecasting of these models and the properties of the forecasts are contained in [1,2,6]. It should be emphasized that the forecasts produced from the models will be accurate representations of future events if, in fact, the system continues to operate in the future as it had in the past.

A rudimentary assessment of alterations in system operation from historical tendencies can be made by comparing future forecasts made at time  $T$  to their corresponding observed values monitored at  $T + k$ ,  $k = 1, 2, \dots$ . Thus, probability estimates can be made of the likelihood of the current observation deviating from the forecast given the underlying process, as manifest by the historical data, is unchanged [3]. Direct measurement of change in the historical process is also possible [4,5,7]. To see whether a process has changed after time  $T$ , the form of the historical model prior to  $T$  can be augmented to allow for a shift parameter. Statistical estimation and hypothesis tests are applied in evaluating whether a real shift or change has been observed. The coupling of this type analysis with new activities in the system and their time frame of implementation is what we shall refer to as intervention analysis.

## REPARAMETERIZATION OF UNIVARIATE MODELS

In intervention analysis the total set of observations of crime levels ( $n$ ) is partitioned into two sets. A set of observations of size  $n_1$ , the observed performance of the system prior to a formal system intervention associated with new activities, and a set of observations of size  $n_2$  which are realized after the activities were first implemented. For example, the integrated moving average model or  $(0,1,1)$  process in which,

$$Z_t = Z_{t-1} - \theta a_{t-1} + a_t$$

can be written as:

$$\underline{\text{for } t \leq n_1,}$$

$$Z_1 = L + a_1, \quad t = 1$$

$$Z_t = L + (1 - \theta) \sum_{j=1}^{t-1} a_{t-j} + a_t, \quad t = 2, 3, \dots, n_1$$

for  $n_1 + 1 \leq t \leq n_1 + n_2$ ,

$$Z_t = L + \delta + (1 - \theta) \sum_{j=1}^{t-1} a_{t-j} + a_t, \quad t = n_1 + 1, \dots, n_1 + n_2.$$

It should be noted that  $L$  denotes the starting level of the historical series at  $T = 1$  and  $\delta$  is a shift parameter that occurs at or after  $t = n_1$ . Thus, if the  $Z_t$  are in units of the number of crimes of say type A per month, then  $\delta$  is in units of change in the number of crimes of type A per month after time  $t = n_1$ .

The model prior to  $n_1$  and after  $n_1$  can be coupled by transformation. Transformation by,

$$Y_1 = Z_1, \quad t = 1$$

and

$$Y_t = Z_t - (1 - \theta) \sum_{j=0}^{t-2} \theta^j Z_{t-1-j}, \quad t = 2, \dots, n_1 + n_2$$

yields

$$Y_t = f_1(\theta)L + a_t, \quad t \leq n_1$$

and

$$Y_t = f_1(\theta)L + f_2(\theta)\delta + a_t, \quad n_1 + 1 \leq t \leq n_1 + n_2.$$

Thus, for all  $t$ ,  $Y = X\beta + \varepsilon$  where  $\beta' = [L, \delta]$  results a general linear model form. Assuming  $\varepsilon$  to be normally and independently distributed with a zero mean and constant variance allows probability statements about  $L$  and more importantly  $\delta$ . For further detail of these estimators and the specific tests of hypotheses see [4,5,7].

#### ON-LINE PROCESS MONITORING

The value of the aforementioned approach to the measurement of the impact of system interventions rests with the ability to detect small but real changes and the ability to detect such a change with minimal time lapse. That is, the method should allow the untangling of small shifts confounded in the data variation, while the process time needed for detection should be small to allow for corrective action. Thus, the method should be sensitive with small values of  $\delta$  and small values of  $n_2$  respectively.

Sensitivity analyses were conducted varying (1)  $\theta$  (which determines the variability of the time series, increasingly positive values giving rise to increasing variation while increasingly negative values giving rise

to increasing smoothness or less variability), (2)  $\delta$  (the magnitude of the shifts from values as small as 2 percent of the base level of the series (L)), (3)  $n_1$  (the historical information), and (4)  $n_2$  (the new observations after a system intervention). The results of these analyses indicate that small shifts ( $\delta = 2$ ) can be statistically detected when confounded in a high variation series with as few as one or two new observations. Fuller details of the sensitivity analysis may be found in [7].

In discussions to this point, the (0,1,1) model has been solely used in illustrations. With respect to measuring impact upon crime occurrences it should be again noted that the seven index crimes were either described by this model form completely or this model form was an integral component of the overall model as in the occurrences of multiplicative seasonal model forms. Here, the within-season variation, or the variation in 12-month spans, is described by this form. In these latter cases to insure early detection of change, it is suggested that between component be filtered out and monitoring be done solely with the within (0,1,1) form. Details of this transformation are given in [5]. Thus the (0,1,1) form is integral to the evaluation of impact of program activities of crime occurrence.

#### AN EXAMPLE: GUN CONTROL

In this section, an example of the use of intervention analysis as described is presented. In particular we address the evaluation of the impact of a legislative policy change on a goal of the reduction of gun-related crimes. This example is not intended to reflect a comprehensive evaluation effort but rather an illustrative application in the "real world." The Massachusetts' Gun Control is briefly summarized to put in proper perspective the specific dimension that is evaluated with the intervention analysis approach. The summary statistical results of the analysis are then presented. A more complete description of this evaluation is contained in [5].

In April of 1975, the State of Massachusetts formally put into operation a gun control law which mandates a 1-year minimum sentence upon conviction of carrying a firearm without a special license. The consequences of this law merit serious study for at least two reasons. First, this State-level attempt to curb firearm violence represents a substantial variation from present and prior policy not only in Massachusetts but in the entire United States. Does this altered policy have a deterrent effect? Secondly, prior to the commencement of this law, there was virtually no limit on judicial discretion in providing minimum sentences. What is the effect of this increased pressure on the prosecuting and judicial elements of the criminal justice system?

Although Massachusetts' law on the carrying and ownership of firearms is multifaceted, it can be summarized as follows:

A. A Firearms Owner Identification (FOI) card is required in order to own or possess either a firearm or ammunition. This card can only be issued to nonaliens over 18 years old who have never been convicted of a felony or hospitalized for drug addiction, drunkenness, or mental illness. The unusual aspect of this facet of the law is that only about 40 percent of the States require prospective firearms purchasers to prove in advance

of acquiring a gun that they have not been excluded according to the above criteria.

B. In addition to satisfying the criteria mentioned in part (A), prospective handgun purchasers in Massachusetts must also satisfy police of their need to own the handgun, whereupon the police may issue a special license if they are satisfied that such a need exists. However, the police are not required to issue the license even if need of ownership has been established. Although these two facets of the law attempt to curb the availability of firearms, they do not prohibit the importation of firearms from contiguous States and their illegal possession.

C. While the first two facets of the law are directed toward curbing the availability of firearms (including handguns), the third facet is concerned with the carrying of firearms. Although carrying a firearm in most other States is also a criminal offense, the unique feature of Massachusetts' law is the mandatory 1-year minimum sentence upon conviction of carrying a handgun without a license to carry or purchase or carrying a rifle or shotgun without an FOI card. Prior to the enactment of the new law, there was virtually no limit on judicial discretion in providing minimum sentences. Under the new law, sentences cannot be suspended and parole cannot be granted until at least 1 year has been served in jail.

Although the mandatory jail sentencing does remove most judicial leeway in sentencing a defendant, the defendant can still escape the mandatory 1-year sentence via three avenues. First, if a person is apprehended with a firearm on his person, the police can file a charge of merely possessing an unlicensed gun in contrast to carrying an unlicensed gun. The possession violation does not carry a mandatory minimum penalty. Secondly, the prosecutor can also press for the lesser violation of possession, regardless of the initial police charge. Thus, the prosecutor still retains the plea bargain option and all its ramifications. Finally, the judge or jury can always find the defendant guilty of the lesser charge. It has been suggested that, "the one-year minimum will only invoke mandatory one-year jail terms for carrying firearms without a license to the extent that police, prosecutors, and judges want it to produce such results. If there is strong resistance from any single link in this chain, the mandatory minimum can be avoided" [8].

Although the impact of the new law on the prosecuting and judicial elements of the criminal justice system is uncertain, it has been hypothesized that, while the number of jury trials for carrying violations will increase, the number of prosecutions and convictions will decrease [8]. The type of defendant will also change in that he will have a prior criminal record involving violent crimes committed with a gun. Furthermore, the new law may lead to more jail sentences of duration less than 1 year since more defendants will be charged with the lesser possession violation. This aspect of the new law should also influence the crime reports by increasing the number of possession violations and simultaneously decreasing the number of carrying violations. Although these facets of the law's impact merit investigation, this research is specifically concerned with the deterrence properties, if any, that the new law may have on the commission of certain gun-related crimes.

To measure the effectiveness of the new law as a deterrent to carrying guns and the commission of gun-related crimes, the offenses of homicide, assault with a gun, and armed robbery will be examined for a change in their occurrence levels prior to and after the enactment of the law. Also, because of the localized nature of crime and the criminal justice system and the concentration of crime in bigger cities, the City of Boston will be used as the evaluation site.

Monthly Uniform Crime Report data for assault with a gun, homicide, and armed robbery were collected from January 1966 to October 1975. These time series were modeled using the multiplicative autoregressive-moving average model class previously described. The seasonal component of the resulting models for assault and robbery was filtered from these data series. Figure 1 exhibits the latter segment of the raw assault data, and its corresponding seasonally adjusted data. Figure 2 exhibits the latter segment of the raw armed robbery data and its corresponding seasonally adjusted data. The homicide model had no seasonal component, being simply described by a random walk process. Figure 3 exhibits the homicide data for Boston. It is to be noted that the two seasonally adjusted series of assault and robbery and the homicide series are all now described by the (0,1,1) process.

Each of these series was used in intervention analysis starting with January of 1974 which corresponds to  $n_1 = 85$  for the assault and robbery series and  $n_1 = 97$  for the homicide series. If the new observations  $n_1 + 1, n_1 + 2, \dots$  comprising the set of  $n_2$  observations exhibited an insignificant shift (e.g.,  $\delta = 0$ ), then the base point time index was updated and the analysis repeated. That is, in the first update we would move to  $n_1 = 85$  for assault and robbery which corresponds to standing at February 1974. This procedure was repeated sequentially until a significant shift was estimated. For assault and robbery the point at which the historical process was observed to have shift corresponded to  $n_1 = 98$  or February 1975. For the homicide series, no significant shift was detected. Table 1 summarizes the statistical information for assault and robbery and table 2 the corresponding analysis for homicide. For more complete details see [5].

With respect to assault and robbery we see a coincident decrease or shift in level a few months prior to the official enactment of the law. However, this time frame is consistent with the early publicity of law provided by the news media. However, no statistically significant change has been noted for homicide, at least through October of 1975. Perhaps due to the large proportion of residential homicides, any future impact of gun control on homicide in general may not be detected for several years, if ever.

## CONCLUSIONS

An approach to the measurement and evaluation of crime pattern modification by planned activities has been described. The procedures of building empirical-stochastic models for crime occurrence, their reparameterization for intervention analysis, and their versatility for on-line system monitoring discussed. The techniques were illustrated for the gun-related crime types with respect to the potential for general deterrence associated with the Massachusetts' Gun Control Law.

FIGURE 1. Boston's monthly occurrences of assault with a gun  
(series A) and the seasonally adjusted data (series B)

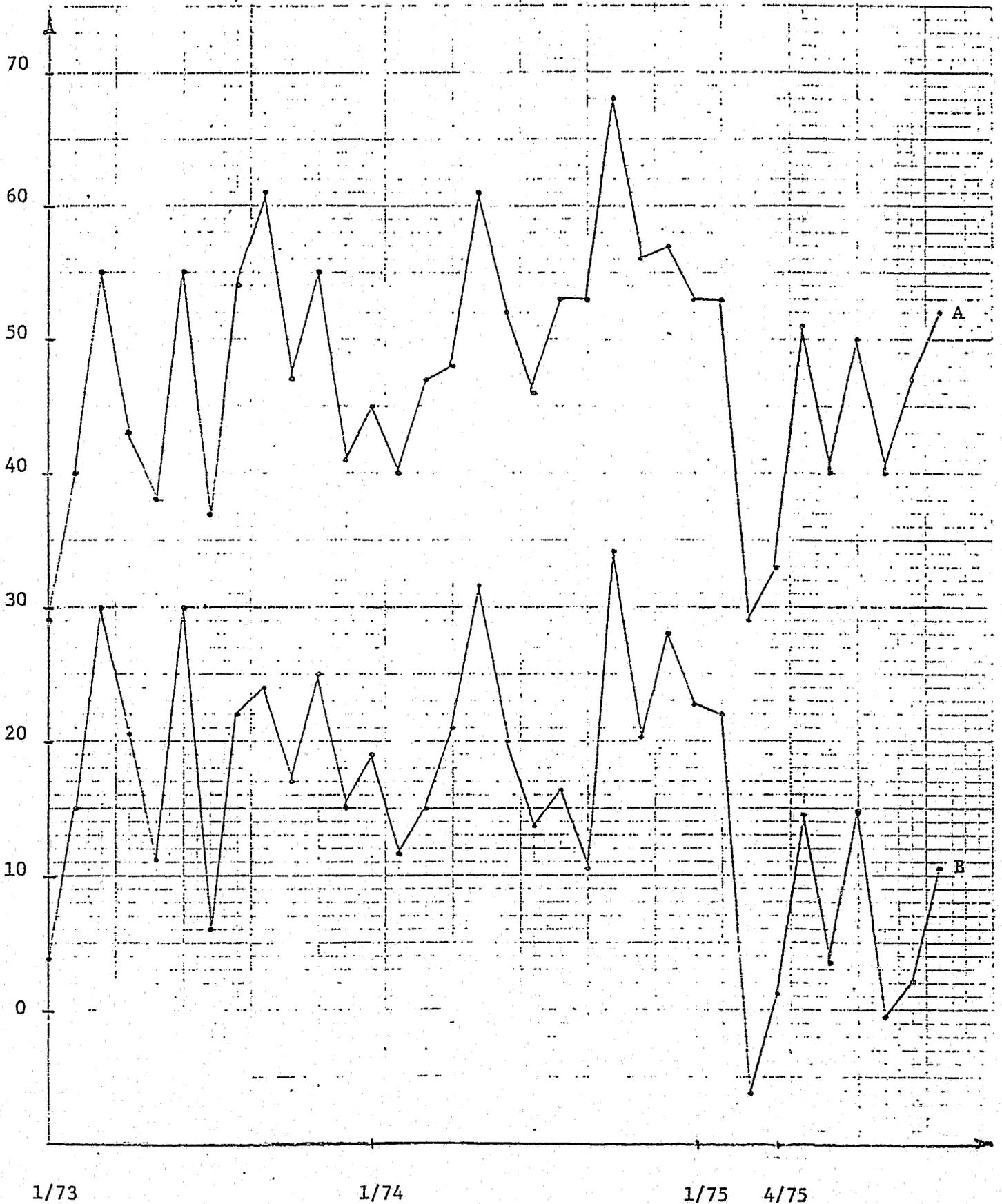


Figure 2. Boston's monthly occurrences of armed robberies (series A) and the seasonally adjusted data (series B)

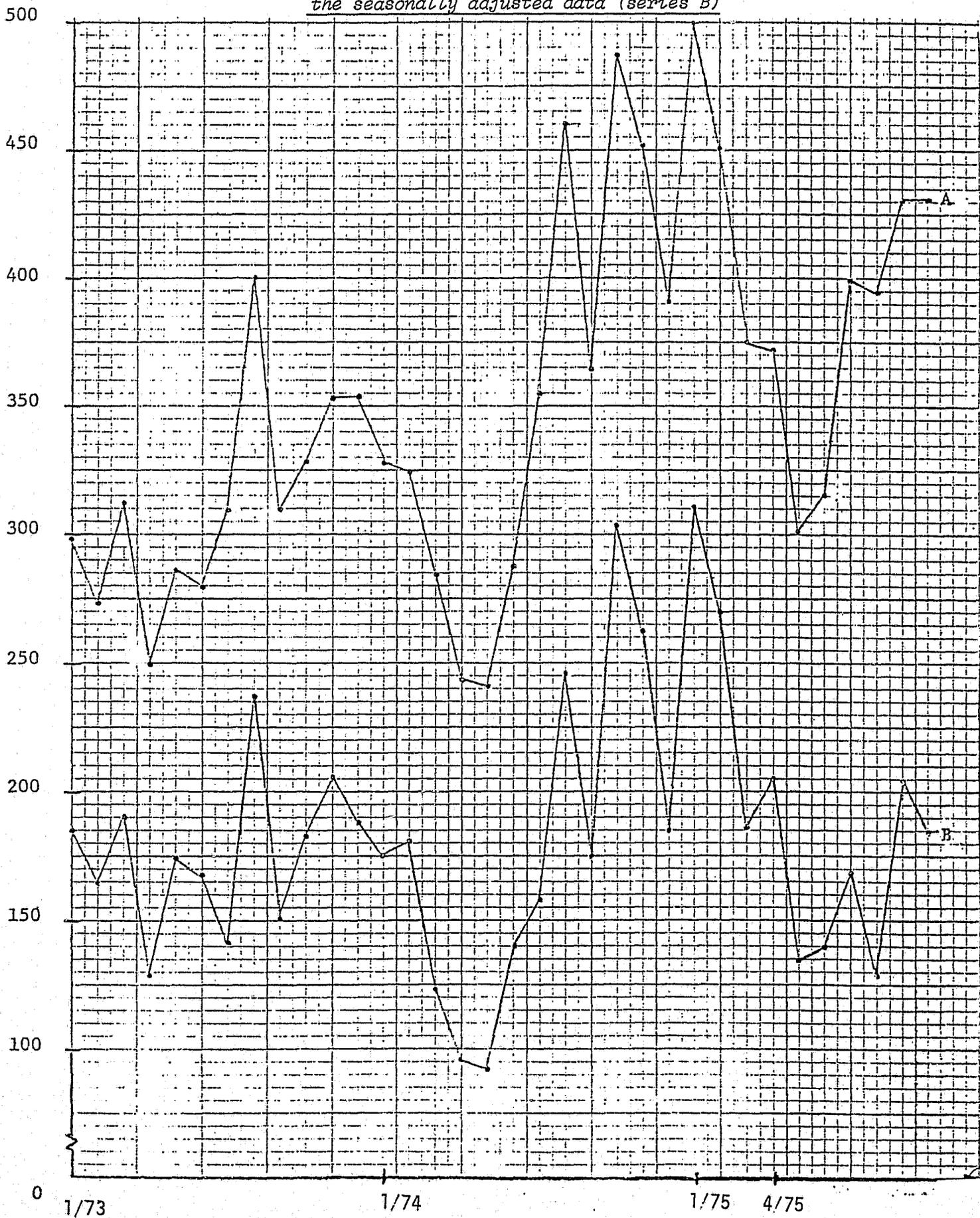


Figure 3. Boston's monthly occurrences of homicide  
(no seasonal component)

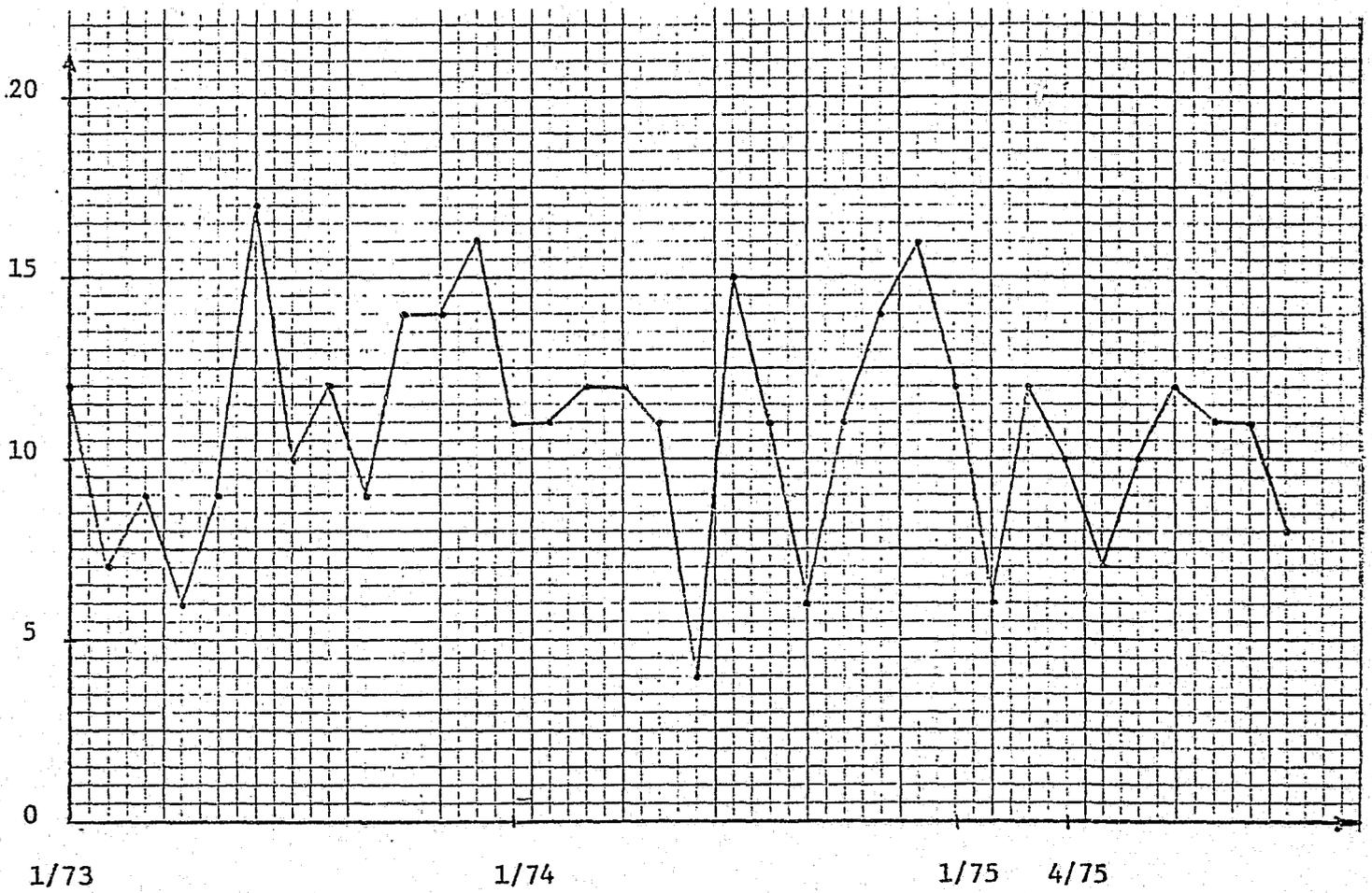


TABLE 1.--Shift detection results for assault with a gun and armed robbery

$n_1$	$n_2$	$\hat{\delta}$	T	Sig. level	95% conf. int.
Assault with a gun					
98	1	-28.21	-3.52	0.001	(-44.11, -12.31)
	2	-24.56	-3.96	0.000	(-36.88, -12.24)
	.	.	.	.	.
	.	.	.	.	.
	8	-18.01	-3.76	0.000	(-27.50, -8.52)
Armed robbery					
98	1	-80.36	-2.17	0.033	(-153.86, -6.86)
	2	-73.28	-2.22	0.029	(-138.74, -7.82)
	3	-85.35	-2.63	0.010	(-149.84, -20.92)
	4	-87.73	-2.72	0.008	(-151.62, -23.83)
	5	-86.96	-2.72	0.008	(-150.50, -23.42)
	6	-87.69	-2.74	0.007	(-151.10, -24.28)
	7	-86.98	-2.70	0.008	(-150.90, -23.06)
	8	-86.93	-2.71	0.008	(-150.55, -23.30)

TABLE 2.--Shift detection results for homicide

$n_1$	$n_2$	$\hat{\delta}$	T	Sig. level	95% conf. int.
108	1	1.50	0.40	0.690	(-5.93, 8.94)
109	1	-5.78	-1.55	0.125	(-13.19, 1.62)
110	1	1.32	0.35	0.727	(-6.14, 8.77)
111	1	-0.93	-0.25	0.804	(-8.36, 6.49)
112	1	-3.76	-1.01	0.316	(-11.15, 3.63)
113	1	0.96	0.26	0.798	(-6.43, 8.35)
114	1	2.78	0.75	0.456	(-4.58, 10.13)
115	1	0.25	0.07	0.947	(-7.09, 7.59)
116	1	0.20	0.05	0.957	(-7.11, 7.51)
117	1	-2.84	-0.77	0.442	(-10.12, 4.44)

## BIBLIOGRAPHY

1. Box, G. E. P., and G. M. Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, California, 1970.
2. Deutsch, S. J., "Stochastic Models of Crime Rates," Journal of Criminal Justice, 1976.
3. Deutsch, S. J., "Effectiveness Measurement for Crime Control," Criminology, 1976.
4. Box, G. E. P., and G. C. Tiao, "A Change in the Level of a Nonstationary Time Series," Biometrika, 52, pp. 181-192, 1965.
5. Deutsch, S. J., and Francis B. Alt, "The Effect of Massachusetts' Gun Control Law on Gun-Related Crimes in the City of Boston," Evaluation Quarterly, 1977.
6. Deutsch, S. J., and P. E. Pfeifer, "Forecasting Multiplicative Incidence Models," Technical Note #5, Department of Justice--LEAA, Grant #75NI-99-0091, 1975.
7. Deutsch, S. J., and Francis B. Alt, "Estimation of Shifts in Stochastic Models of Crime Occurrence," Modeling and Simulation, Vol. 7, No. 2, pp. 780-785, 1976.
8. Zimring, F. E., "Massachusetts' New Mandatory Minimum Sanctions for Gun Law Violations: A Preliminary Research Design," submitted to NILECJ, May 20, 1975.

EXPERIMENTAL DESIGN AND CAUSAL ANALYSIS  
IN CRIMINAL JUSTICE EVALUATION

John G. Heilman and Marshall T. Miller

The purpose of this paper is to outline a practical research design for causally oriented evaluations of criminal justice projects and programs. The design carries a shamelessly technical title, the "ex post facto staggered time series design." Nevertheless it can be easily understood, and responds to the situation of many practicing criminal justice evaluators. In particular, ex post facto data collection is especially appropriate for many programs the Law Enforcement Assistance Administration (LEAA) funds at the State level, even though ex post facto designs are normally discouraged for causal analysis.

Our argument begins with a brief elaboration of the above points. First, we describe our own experience as what we call "practicing evaluators," the group to whose needs this paper is ultimately addressed. For nearly 3 years our six-man research staff at Auburn University has served as the outside evaluator for LEAA-funded programs in Alabama. In this time we have evaluated, in one way or another, over 100 individual grants. Working with a limited staff, we have not attempted to study causation. The emphasis in our work has been to help project managers establish measurable objectives, data collection procedures, and internal evaluation criteria. Our goal has been to help the State planning agency make informed management and planning decisions. Similar efforts in other States appear to have comprised the majority of criminal justice evaluation work.

So, we treat as practicing evaluators those persons who are paid by a governmental agency to evaluate a range of projects and programs on an on-going basis, with relatively limited funding. They must often study projects already underway. Thus, the problem of practicing evaluators is to find suitable data and a suitable analytic design within a restricted period of time, on limited resources, and after project activities have commenced.

Another problem now confronting many evaluators is the transition from simple monitoring, which LEAA has so far accepted as a kind of "evaluation," to causal analysis, which LEAA now accepts as the only "true" evaluation. The difficulty is that research designs permitting causal analysis often require experimental controls obtainable only in laboratory settings. To explain this difficulty we now review Donald Campbell's discussion of research designs and possible threats to their validity.

Validity threats are factors complicating the researcher's decision of whether a particular variable or stimulus caused an observed effect or outcome. To make this decision the researcher must be able to discount as

many alternative causes, or rival hypotheses, as possible. Of course, the researcher needs to concern himself only with those rival hypotheses that are plausible. The greater the number of plausible rival hypotheses that a research design can refute, the better the design. The researcher never actually proves his own hypothesis, but by eliminating rivals he lends credibility to it.

Many of the rival hypotheses are actually artifacts of the design itself and are therefore considered threats to the validity of the design. Among the validity threats Campbell includes are these:

1. History: Events, other than the experimental treatment, occurring between pretest and posttest and thus possibly causing measured changes.
2. Maturation: Processes within the respondents or observed social units producing changes as a function of the passage of time itself, such as growth, fatigue, and secular trends.
3. Instability: Unreliability of measures, fluctuations in sampling persons or components, autonomous instability of repeated or "equivalent" measures.
4. Testing: The effect of taking a test upon the scores of a second testing. Put another way, the effect of publication of a social indicator upon subsequent readings of that indicator.
5. Instrumentation: Changes in the calibration of a measuring instrument, or changes in the observers or scoring procedures which produce unreliable measurements.
6. Regression artifacts: Relationships which appear only because persons or treatment units have been selected upon the basis of their extreme scores.
7. Selection: Variation in the techniques used to select experimental groups, producing different mean levels on the measure of effects.
8. Experimental mortality: The nonsystematic loss or dropping out of respondents from comparison groups during the experiment.
9. Selection-maturation interaction: Selection biases resulting in differential rates of "maturation" or autonomous change.

These nine factors threaten only the validity of conclusions reached about the groups under study and are therefore termed "internal" threats to validity. Campbell also identifies "external" validity threats. These problems challenge the researcher's attempt to generalize from his specific results to a broader universe of cases. We consider only the internal validity threats because Campbell's external validity threats are mainly problems of reactivity (interaction between the experimenter and experimental subject), which do not seriously threaten the staggered time series design we

shall propose. In fact, our design effectively controls external validity threats, although the proof is too long for inclusion here.

Campbell considers the extent to which several types of research designs control for various validity threats. He distinguishes between pre-experimental, quasi-experimental, and true experimental designs. To illustrate these designs Campbell uses the following system. The letter O is used to indicate an observation or measurement of a variable that is under consideration. The letter X represents the introduction of a treatment or the starting of a funded program. A horizontal row represents the experimental activities affecting a single group or agency being studied. Thus the following configuration shows that one group or agency was studied, and this entity was measured once before the treatment started and once afterward.

O   X   O

A second row shows that two groups are being compared. The first row records the experimental activities on the first group and the second row the experimental activities on the other group.

O   X   O  
O        O

The preceding configuration shows two groups being studied. The first group was given the treatment and is termed the experimental group; the second group was not given the treatment and is termed the control group.

The use of dashes to separate rows indicates that the groups are not matched and thus are not equivalent. If groups are matched then no dashes are used, and if matching was achieved by randomly assigning individuals to the groups, then an R precedes the row. The following two illustrations show respectively (1) an experimental group and a control group that were not matched and (2) an experimental group and a control group that were matched by random assignment.

(1)    O   X   O  
      -----  
      O        O

(2)    R O   X   O  
      R O        O

The procedures that practicing evaluators have used to conduct monitoring evaluation fall under what Campbell calls pre-experimental designs. These designs include the following.

One-shot case study

X   O

One-group pretest-posttest design

0 X 0

Static group comparison

X 0  
-----  
0

These designs score very poorly on controlling for internal validity threats. The evaluation work done employing these designs is variously termed formative, enumerative, or categorial, as distinguished from causal studies, which are usually termed analytical or summative. The analytical studies have for their ideal designs the true experimental designs. Thus it would appear that the practicing evaluator only needs to switch from pre-experimental to true experimental designs. Unfortunately, to switch is not so easy. A depiction of true experimental designs will show why.

Pretest-posttest control group design

R 0 X 0

R 0 0

Solomon four-group design

R 0 X 0

R 0 0

R X 0

R 0

Posttest-only control group design

R X 0

R 0

These designs call for random assignment of treatment, a practice beyond the realm of possibility for practicing evaluators. As mentioned before these evaluators are often thrust into research situations over which they can exert little or no control. There may even be a question of legality concerning random selection in criminal justice programs. Also, many programs to be evaluated have been in operation several years before the evaluator is asked to study the effect. For these reasons true experimental designs are not feasible for practicing evaluators.

Fortunately, there is a group of designs that Campbell terms quasi-experimental. He proposes them as useful alternatives for researchers who do not have the flexibility and control required by true experimental designs. The quasi-experimental designs attempt to control for validity

threats by elaborating the pre-experimental designs. For our purposes, the elaboration can take the form of extending the number of observations (time series) or the number of groups being observed (experimental and control groups). Three quasi-experimental designs are illustrated below:

Time series design

0 0 0 0 X 0 0 0 0

Nonequivalent control group design

0 X 0

-----  
0 0

Multiple time series

0 0 0 0 X 0 0 0 0

-----  
0 0 0 0 0 0 0 0 0

These designs also appear to present serious drawbacks to the practicing evaluator. First, extending the number of observations requires an expanded time period for evaluation. Such an expansion both requires more funds and delays the results of evaluation. Few practicing evaluators can tell a governmental agency that it must wait several years for results. Also, elaborating an evaluation design by observing more groups requires the cooperation and availability of comparative groups. It would seem now that even quasi-experimental designs are useless to practicing evaluators.

We shall argue, however, that certain characteristics of LEAA-funded criminal justice programs actually encourage the use of quasi-experimental designs for assessing causal relations in criminal justice programs. The key characteristics of LEAA-funded programs are these:

- (1) Many programs are continuation efforts and have been funded for several years.
- (2) Often the same or similar programs have been introduced at different times and in various locales.
- (3) Much of the data needed for evaluation is available independently of the evaluator's research efforts.

We shall now suggest how each of these advantages can encourage the creation of a quasi-experimental design for the practicing evaluator.

1. Continuation Funding. The funding of many programs for a period of years permits the use of time series analysis. There are two ways to do a time series study: either one can analyze programs that have operated for several years, or one can start a study and hope that it will be supported over a long period. Since the governmental agencies supporting practicing evaluators want immediate results, the latter alternative is inappropriate. The former alternative, analyzing programs that have operated for several

years, requires ex post facto analysis, a frequently criticized practice. Later we shall respond to the criticisms; for the moment we simply suggest that the ex post facto study complements the existence of continuation programs.

The time series design is well suited to causal analysis. This design answers most of the internal threats to validity, in particular the mention of instability. The essence of this problem is that some variables, such as traffic fatalities, naturally fluctuate from year to year. Thus a study which has just one pretest and one posttest, say, a year apart, may be vulnerable to instability as a rival explanation of change. However, if the period over which observations are taken can be increased, then instability can be controlled for.

Statistical tests also control for instability, but these tests depend on random assignment. For the practicing evaluator the time series design is a more appropriate control. Thus, the continuation of programs for a period of years invites the use of time series design.

2. The Funding of the Same or Similar Program at Various Locations and Times. State planning agencies often fund similar projects in different locations but start the funding at various times. This staggered sequence also encourages the use of time series design, in two ways. First, the existence of many comparable programs starting at different times permits the researcher to control for the effects of external, historical events. The effects of historical events would show up at roughly the same calendar time in all functioning programs and thus be recognizable. Second, the addition of a control group or case which does not receive the treatment would permit the researcher to control for historical trends. These would show up in the control group or case over time, just as they would in the experimental groups or cases. Of course, for the practicing evaluator this control group probably will not be equivalent to the experimental group. Campbell allows for nonequivalence by developing a Nonequivalent Control Group design.

In sum, although the effect of external historical trends normally remains a rival hypothesis when time series design is used, the use of a control group, even a nonequivalent group, controls the problem. And the staggered sequence of projects being studied permits control of historical events as validity threats. So, the use of a control group (receiving no treatment) and other comparative groups (receiving staggered treatments) effectively combats most validity threats. The staggered time series design is depicted below.

0	0	0	0	X	0	0	0	0
0	0	0	X	0	0	0	0	0
0	0	0	0	0	X	0	0	0
0	0	0	0	0	0	0	0	0

3. The Existence of Data Covering Periods of Years. It is helpful to recall that Campbell's main interest is psychological and educational evaluation. An important difference between this field and criminal justice evaluation is the former's reliance upon tests to obtain data. In psychology and education many of the phenomena being observed, such as I.Q. score, psychological type, or self-concept rating, are artifacts of testing. Campbell argues in effect that data should be collected as the events they reflect occur, rather than ex post facto, because in his field any other method of obtaining data is seldom used. This lack of enthusiasm for ex post facto studies is understandable, because such studies in education or psychology can seldom be done: after-the-fact collection of data is usually by definition impossible.

By contrast, criminal justice events do not always require experimenters to record them as data. Much of the data required by evaluation are being routinely recorded by public agencies and have been recorded for years. These data are normally artifacts of routine institutional behavior--case-load statistics, arrests, indictments, etc. They predate any evaluation effort. Although not always easily accessible, they can be obtained by the diligent researcher. And it is precisely their availability that makes the staggered, time series design both possible and appealing.

The control of validity threats and flexibility offered by this design justifies the effort of data collection. The elaborated design enables the evaluator to consider not only "what is the impact of LEAA funding?" but also "what was the impact of LEAA funding?"

Two possible arguments against using ex post facto designs may thus be refuted as follows. First, the wealth of existing criminal justice data dissolves the contention that the experimenter needs to be present to record data. Second, the feeling that true social research occurs only when the theory precedes data collection is unfounded. Observation of naturally occurring phenomena and generalization from these observations is the original scientific method.

To be useful, a research design must be applicable, and so the last step in this presentation is a sample application of the design we have outlined thus far.

Since 1971 the State planning agency in Alabama responsible for distributing LEAA funds has supported a program to improve the prosecutorial efficiency of district attorneys. This program consists of funding local district attorneys for purposes of increasing their staffs, upgrading their staffs, and providing inservice training. We at Auburn have been asked to evaluate the impact of this program.

The first consideration in terms of applying the staggered time series design to this evaluation is that many of the district attorneys being funded now have been funded for several years. Thus, provided we can obtain data generated prior to the program's inception, a time series study is impossible. The next consideration is that not all district attorneys in the State have received funding; thus nonequivalent control groups exist. Finally, the program was begun at different times for different

district attorneys. This method of funding enables the evaluator to select staggered comparative groups.

The independent variable or treatment in this evaluation is the LEAA-funded program. The dependent variables or effects to be studied are those factors that would indicate improved prosecutorial efficiency: case backlogs, instances of nolle prosequi, prosecution success rates (convictions divided by indictments), and elapsed prosecution time (time from arrest to final disposition). Even though no formal evaluation effort has been in effect, data for these dependent variables are on record with the circuit courts.

We now consider how our proposed design handles internal validity threats in this case. The time series feature controls for maturation, regression, and instability. The nature of the data and method of observation, naturally occurring phenomena that are institutionally recorded, discount the plausibility of testing and instrumentation as rival hypotheses. Of course the evaluator must investigate whether recording procedures were changed or improved; the use of staggered comparative groups, however, strengthens the design against this threat. When measuring entire social units rather than individuals within a group, mortality is not a significant factor either. History, selection, and selection-maturation interaction are all controlled for by the presence of a control group, other comparative groups that are staggered, and again by the time series observations.

To summarize, the authors have proposed a quasi-experimental research design for practicing evaluators. The design is most completely termed an ex post facto, staggered, time series design. This design takes advantage of three aspects of LEAA-funded programs: (1) it employs a time series design to take advantage of the numerous programs that have received continuation funding; (2) it uses staggered, nonequivalent comparative groups and a nonequivalent control group to take advantage of the existence of similar programs starting at different times; and (3) it advocates ex post facto analysis to take advantage of the large amount of criminal justice data already in existence. When subjected to a validity checklist, the above design scores as high as the true experimental designs discussed by Campbell. Thus, the ex post facto, staggered, time series design appears to provide practicing evaluators with a feasible research design for conducting the causal analysis LEAA requires.

## BIBLIOGRAPHY

- Campbell, Donald T., "Reforms as Experiments," collected in Handbook of Evaluation Research, Struening, Elmer L. and Guttentag, Marcia, eds. Beverly Hills: Sage Publications, 1975. Appeared originally in a different form in American Psychologist, vol. 24, no. 4 (April 1969).
- Campbell, Donald T., and Stanley, Julian C., Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally and Company, 1963.
- Hatry, Harry P., Winnie, Richard C., and Fisk, Donald M., Practical Program Evaluation for State and Local Government Officials. Washington, D.C.: The Urban Institute, 1973.
- Isaac, Stephen, and Michael, William B., Handbook in Research and Evaluation. San Diego: EdITS publishers, 1971.

DETERMINANTS OF THE SERIOUSNESS OF CRIMINAL ACTIVITY:  
THE MISDEMEANOR-FELONY DISTINCTION\*

Ann Dryden Witte  
Assistant Professor

Peter Schmidt  
Associate Professor

Department of Economics  
University of North Carolina

1. INTRODUCTION

The ideal program evaluation is an evaluation which randomly selects individuals from a population to participate in a program. The postparticipation performance of individuals selected for the program is then compared on the basis of the program objectives with that of the individuals who were in the population but did not participate. Significant differences in performance of the two groups in such an evaluation can be attributed to program participation since the two groups should not differ systematically in any other characteristic due to the random selection. However, evaluations of criminal justice programs are rarely able to use this ideal technique because of legal or moral restrictions on choosing program participants. Even where such restrictions do not exist this technique has not been used extensively in criminal justice evaluation because these evaluations are usually only begun after a program has been in operation for a period (often substantial) of time. Selection criteria for ongoing programs are often informal and change over time and are rarely, if ever, random.

Due to these difficulties, criminal justice evaluators have developed a number of techniques for evaluating programs under less than ideal conditions. One of the most interesting of these techniques has been performance prediction.<sup>1</sup> This technique develops a statistical model to predict the performance of individuals in a population of interest, e.g., prison inmates. It then compares the performance of a group of individuals from this population who participate in a program with the performance predicted by the statistical model. If the model is accurate, a significant difference between the actual and predicted performance of this group should be the result of program participation.

---

\*The analysis contained in this paper was supported by a grant from the North Carolina Department of Correction. However, the statements and conclusions in the paper are solely those of the authors, and do not necessarily represent the opinion of the Department of Correction.

The most common criterion of success in criminal justice evaluations which have used the above technique, or indeed in most other criminal justice evaluations as well, is a reduction in the rate of recidivism, defined as the percentage of individuals returning to criminal activity. While most evaluators have realized the crudeness of this measure, they have been constrained in their ability to use more sophisticated measures of criminal activity (e.g., the frequency and seriousness of criminal activity) because of the lack of appropriate multivariate statistical techniques to analyze such variables. Measures of the frequency or seriousness of criminal activity are often categorical, generally polytomous rather than binary, and even when continuous are usually skewed and truncated at 0, since some individuals do not return to criminal activity.<sup>2</sup>

It is inappropriate to use readily available multivariate statistical techniques such as multiple regression analysis which are based upon the normal distribution in analyzing such variables which are clearly not normally distributed.<sup>3</sup> In recent years statisticians and econometricians have developed multivariate techniques to analyze dependent variables such as the above<sup>4</sup> and this paper illustrates the use of one such technique to develop predictions of one measure of the seriousness of criminal activity.

Specifically we investigate what determines into which of three categories an individual's most serious conviction during a follow-up period will fall. The categories are in order of decreasing seriousness: (1) conviction for a felony, (2) conviction for a misdemeanor, (3) no conviction. This variable is analyzed using the logit model.

The next section of the paper gives a brief description of the data set on which the analysis is based. Section 3 describes the model and gives the main results of our analysis. Section 4 gives predicted probabilities of the various categories for certain types of individuals. The final section summarizes the paper and suggests ways in which this model can be used in criminal justice evaluation.

## 2. THE DATA

The data used in this paper consist of information on the postrelease activities of a random sample of 641 men who were in prison in North Carolina in 1969 or 1971. The activities of these men were followed for an average period of 37 months after their release from incarceration.<sup>5</sup> Although the sample was drawn from a single administrative area, the population from which it was drawn is quite representative of the population in medium and minimum custody prison units throughout North Carolina.<sup>6</sup> The data are unusually complete in three senses. First, they contain detailed information on postrelease activities, since 71 percent of the sample of inmates was interviewed an average of 37 months after release. Second, criminal records are unusually complete. They were obtained by writing all areas where there was any indication that an individual lived after release and all areas where self-report, FBI check, or North Carolina Department of Correction's records indicated that an individual had contact with criminal justice authorities. Finally, it should be emphasized that this data set is essentially a random sample of all releases.<sup>7</sup> Unlike many other data sets, it is not limited to individuals on parole.

### 3. MODEL AND RESULTS

As mentioned in section 1, the dependent variable to be analyzed is the type of the most serious crime<sup>8</sup> for which an individual is convicted during a follow-up period after release from prison. Three possible outcomes are considered:

- (1) NONE, indicating that the individual was not convicted of any crime during the follow-up period;
- (2) MISD, indicating that the most serious conviction was for a misdemeanor;
- (3) FEL, indicating that the most serious conviction was for a felony.<sup>9</sup>

Due to the fact that this variable is categorical rather than continuous, the logit model<sup>10</sup> is an appropriate model to use. The logit model specifies that the logarithm of the ratio of the probabilities of these various outcomes for the *i*th individual, denoted  $P_{1i}$ ,  $P_{2i}$ , and  $P_{3i}$ , respectively, are a linear function of the explanatory variables. Formally

$$\ln(P_{ri}/P_{si}) = \sum_{j=1}^k \beta_{r,s,i} X_{ij}$$

where *r* and *s* go from 1 (NONE) to 3 (FEL), *k* is the number of explanatory variables, and  $X_{ij}$  represents the value of the *j*th explanatory variable for the *i*th individual. The  $\beta$ 's are coefficients to be estimated. The method of estimation used is maximum likelihood.<sup>11</sup>

Since little theoretical or empirical work<sup>12</sup> has been done on the determinants of the seriousness of criminal activity, the explanatory variables that we use are essentially those variables that have been found to be associated with individual criminal activity in previous studies.<sup>13</sup> Specifically the explanatory variables used are as follows: a constant term, denoted CNST; a dummy variable equal to one for whites and zero for nonwhites, denoted RACE; the number of convictions prior to the one leading to the sample sentence,<sup>14</sup> denoted CONVBS; a dummy variable equal to one if the sample sentence was for a felony, and zero if for a misdemeanor, denoted MF; a dummy variable equal to one if the individual's record indicates a serious problem with alcohol, or use of hard drugs, denoted ALKY; age (in hundreds of months) at the time of the individual's first arrest, denoted AFA; a dummy variable equal to one if the individual's release from the sample sentence was supervised (e.g., parole), and zero otherwise, denoted SUPER; age (in hundreds of months) at the time of release from the sample sentence, denoted AAR; the length (in tens of months) of the individual's follow-up period, denoted LENG; a dummy variable equal to one if the sample sentence was for a crime against a person, and zero otherwise, denoted PERS; a dummy variable equal to one if the sample sentence was for a crime against property, and zero otherwise, denoted PROP; the number of prison rule violations during the sample sentence, denoted RULE; a dummy variable equal to one if the individual was married when released from the sample

sentence, and equal to zero otherwise, denoted MS; number of years of schooling completed, denoted SG; and a dummy variable equal to one if the individual participated in the North Carolina prisoner work release program, and zero if he did not, denoted WR.<sup>15</sup> The sample contained 596 individuals (out of the original 641) on whom this information was available.

The maximum likelihood estimates of the parameters of the logit model as specified above are given in table 1. For example, the numbers in the first column, under the heading "MIS/NONE," are the coefficients of the explanatory variables in the equation whose dependent variable is the logarithm of the ratio of the probability that the most serious conviction is a misdemeanor (MIS) to the probability of no conviction (NONE). As one example, the positive coefficient of RACE indicates that if an individual is white, the probability of MIS increases relative to NONE. The numbers in parentheses under the coefficients are "asymptotic t ratios," as obtained from the information matrix. They are asymptotically distributed as  $N(0,1)$  under the null hypothesis that the associated coefficient equals zero. Therefore a "t ratio" in excess of 1.96 in absolute value indicates statistical significance at the .05 level, for the usual two-tailed test.<sup>16</sup>

Two variables, ALKY and LENG, have very similar effects, in that they affect the probability of a conviction, but do not seem to affect the probabilities of a felony relative to a misdemeanor. These results indicate that the probability of a conviction for either a misdemeanor or a felony (MIS, FEL) is increased, relative to the probability of no reconviction (NONE), if an individual is alcoholic or a user of hard drugs (ALKY = 1) and the longer an individual's activities are followed (the larger LENG).

Three variables, RACE, CONVBS, and AAR, affect the probability of conviction for a misdemeanor relative to none, but do not affect the probability of a felony relative to none or a felony relative to a misdemeanor. These findings indicate that the probability of conviction for a misdemeanor is increased relative to no conviction, if an individual is white (RACE = 1), the more previous convictions he has (the higher CONVBS), and the younger he is at release (the lower AAR).

The coefficients on which all of the above statements are based are significantly different from zero at the 5 percent level (two-tailed test). The five variables discussed above have similar effects in that they affect the probability of a conviction relative to no conviction, but do not affect the probability of a felony relative to a misdemeanor. (The coefficients of none of these variables in the last column are significant at even the 10 percent level.)

Three variables, RULE, AFA, and MF, affect the probability of either felony (AFA) or misdemeanor (RULE and MF) relative to none and also affect the probability of conviction for a felony relative to conviction for a misdemeanor. Thus, having more rule violations when incarcerated (a large value of RULE) or being incarcerated for a misdemeanor (MF = 0) increases the probability of conviction for a misdemeanor relative to no conviction and the probability of conviction for a misdemeanor relative to a felony. The younger an individual is when first arrested the greater the probability of conviction for a felony relative to no conviction or relative to a misdemeanor. These results which are all significant at the 5 percent level

indicate that these three variables affect the seriousness of the crime for which an individual is convicted as well as whether or not he will be convicted. This crime will be likely to be more serious (more likely to be a felony than a misdemeanor) the fewer rule violations an individual had when incarcerated, if he was incarcerated for a felony and the younger he was when first arrested.

Two variables, WR and SUPER, affect only the probability of a felony relative to a misdemeanor but not the probability of conviction relative to no conviction. Thus an individual who is on work release and who is supervised when released is less likely to be convicted for a felony relative to a misdemeanor, i.e., is more likely to be convicted of a less serious offense. The above result for WR is significant at the 5 percent level and that for SUPER at the 10 percent level. This result is particularly interesting as participation in work release and supervision on release are variables over which the criminal justice authorities have substantial control. It is interesting that, assuming that our results are due primarily to program participation,<sup>17</sup> an evaluation of either of these programs which used the traditional criterion for success, significant differences in the percent returning to criminal activity, would have judged both these programs as failures. However, an evaluation which investigated our measure of the seriousness of conviction as well would have judged them successful in the sense of decreasing the seriousness of postrelease criminal activity.<sup>18</sup>

Four variables, MS, SG, PERS, and PROP, have no significant effects, at normal levels of statistical significance, on any of the dependent variables and hence need not be discussed. In order to make our analysis more concise we redid it dropping these four variables.<sup>19</sup> The results for this reduced specification are given in table 2. The results for the variables still in the specification are very similar to those in table 1, except that the level of significance of three coefficients is altered. The coefficient for SUPER in the FEL/MIS equation increases in significance and is now significant at the 5 percent level (two-tailed test). The coefficient of RULE in this equation decreases in significance and is significant only at 10 percent level. The coefficient of AAR in the FEL/NONE equation increases in significance from the 10 percent to the 5 percent level.

#### 4. SOME FURTHER RESULTS

In order to further illustrate the above results, we have used the fitted logit model to predict the probabilities that an individual's most serious conviction will be for a felony or a misdemeanor or that he will have no reconviction. This has been done for three hypothetical individuals with disparate personal characteristics, for various periods of time after release. For illustrative purposes we have given these hypothetical individuals very different characteristics in order to illustrate the importance of personal characteristics in determining the seriousness of postrelease criminal activity. It is exactly these types of predictions that would be used to evaluate a criminal justice program.

Table 3 gives predicted probabilities for the "average individual," defined as an individual whose values for the explanatory variables equal the sample mean.<sup>20</sup> For various lengths of time (LENG), table 3 gives

TABLE 1.--Results of logit analysis of the seriousness of criminal activity: all variables

Variable	MIS/NONE	FEL/NONE	FEL/MIS
CNST	.59607 (0.75)	.58473 (0.71)	-.011342 (-0.02)
RACE	.51513 (2.39)	.42313 (1.21)	-.091998 (-0.29)
WR	.31417 (1.49)	-.43965 (-1.20)	-.75382 (-2.21)
ALKY	.87052 (3.86)	.75775 (2.12)	-.11277 (-0.35)
SUPER	.15621 (0.60)	-.59185 (-1.41)	-.74806 (-1.90)
CONVBS	.09104 (2.28)	.07366 (1.44)	-.017388 (-0.48)
RULE	.28800 (3.27)	.08003 (0.66)	-.20797 (-2.03)
AAR	-.50911 (-3.62)	-.39151 (-1.69)	.11840 (0.55)
AFA	.21556 (1.41)	-.75452 (-2.15)	-.97001 (-2.88)
LENG	.35557 (4.50)	.28229 (2.37)	-.073282 (-0.68)
MF	-.89607 (-2.78)	.29429 (0.60)	1.1903 (2.66)
MS	.13685 (0.61)	-.11897 (-0.31)	-.25581 (-0.70)
SG	.061812 (-1.46)	.025620 (0.33)	.087431 (1.21)
PERS	-.36852 (-1.22)	-.61722 (-1.02)	-.24870 (-0.43)
PROP	-.057209 (-0.19)	.09092 (0.19)	.14813 (0.34)

TABLE 2.--Results of logit analysis of the seriousness of criminal activity: reduced specification

Variable	MIS/NONE	FEL/NONE	FEL/MIS
CNST	-.10143 (-0.25)	.83607 (0.92)	.93745 (1.06)
RACE	.55595 (2.63)	.45439 (1.33)	-.10156 (-0.33)
WR	.30503 (1.46)	-.48086 (-1.32)	-.78589 (-2.32)
ALKY	.88674 (3.95)	.76646 (2.15)	-.12027 (-0.37)
SUPER	.098358 (0.39)	-.67029 (-1.62)	-.76864 (-1.98)
CONVBS	.094241 (2.35)	.07950 (1.56)	-.014736 (-0.42)
RULE	.27512 (3.19)	.083882 (0.70)	-.19124 (-1.88)
AAR	-.46633 (-3.53)	-.46539 (-2.16)	.000942 (0.00)
AFA	.21158 (1.39)	-.69666 (-1.99)	-.90824 (-2.72)
LENG	.34561 (4.43)	.27653 (2.31)	-.069081 (-0.64)
MF	-.90199 (-3.27)	.30702 (0.90)	1.2722 (3.34)

TABLE 3.--Predicted probabilities of various outcomes: average case

Length of time (in months)	P(NONE)	P(MISD)	P(FEL)
1	.522	.426	.051
2	.514	.434	.052
3	.506	.442	.052
4	.497	.450	.053
5	.488	.458	.053
6	.480	.466	.054
7	.472	.474	.054
8	.463	.482	.055
9	.455	.489	.055
10	.447	.497	.056
11	.438	.505	.056
12	.430	.513	.057
14	.413	.529	.058
16	.397	.544	.059
18	.381	.559	.060
20	.365	.575	.060
22	.350	.589	.061
24	.334	.604	.062
26	.319	.618	.062
28	.305	.632	.063
30	.291	.646	.063
32	.277	.659	.064
34	.263	.672	.064
36	.250	.685	.064
42	.214	.721	.065
48	.182	.753	.065
60	.129	.807	.064
72	.089	.845	.062

P(NONE), P(MIS), and P(FEL). If comparing the prediction for such an individual with his actual performance, one would choose the prediction for the length of time corresponding to the follow-up period.

As would be expected, the probability of no conviction falls as the length of the follow-up period increases. It is approximately 43 percent after 1 year, 33 percent after 2 years, 25 percent after 3 years, and 9 percent after 6 years. The probability of being convicted for a misdemeanor rises with length of time, which is also expected. After 1 year this probability is 51 percent; after 2, 60 percent; after 3, 69 percent; and after 6 years, 85 percent. Somewhat surprisingly, the probability of being convicted for a felony remains relatively stable as length of time changes. This probability hovers around 6 percent for all follow-up periods. For periods less than 8 months it is around 5 percent; it rises to a peak of 6.5 percent after approximately 4 years; and then again declines to approximately 6 percent for long follow-up periods.

These predicted probabilities can be used to get at least a rough idea of the fit of the model. The usual measures of goodness of fit that are available in, say, a linear regression model, are not available here. Furthermore, it is somewhat difficult to compare predictions to actual frequencies since each observation (individual) has a different set of explanatory variables. However, we can hope for at least a rough correspondence between the sample proportions in our three categories, and the predicted probabilities based on average characteristics and the average length of the follow-up (which was 37 months). The predicted probabilities of NONE, MIS, and FEL, for our "average individual," for a follow-up period of 37 months are 24.4 percent, 69.1 percent, and 6.5 percent. These compare fairly closely to the actual sample proportions of 27.0 percent, 63.9 percent, and 9.1 percent. This is encouraging evidence for the adequacy of the model.

Table 4 gives the predicted probabilities for a very favorable case. This would be an individual who is old (AAR = 6) and nonwhite, who is not an alcoholic or hard drug user, whose release from imprisonment was supervised, who had no previous convictions (prior to his sample crime), who committed no rule violations while in prison, who was first arrested at an advanced age (AFA = 4.8), whose sample sentence was a misdemeanor, and who participated in the prisoner work release program when serving this sentence.

Such an individual is much more likely to receive no conviction and less likely to receive a misdemeanor than is our "average individual." However, the most dramatic feature of table 4 is the decreased probability of a felony for the individual with favorable characteristics compared to the "average" individual.

Finally, table 5 gives the predicted probabilities for a very unfavorable case. This would be an individual who is young (AAR = 2.6), white, who is an alcoholic or hard drug user, whose release from imprisonment was unsupervised, who had four previous convictions, and four rule violations during his sample sentence, who was very young when first arrested (AFA = 1.56), whose sample sentence was for a felony, and who did not participate in the prisoner work release program when serving that sentence.

TABLE 4.--Predicted probabilities of various outcomes: "best" case

Length of time (in months)	P(NONE)	P(MISD)	P(FEL)
1	.808	.190	.001
2	.803	.196	.001
3	.797	.201	.001
4	.792	.207	.001
5	.786	.213	.001
6	.780	.218	.001
7	.774	.224	.001
8	.768	.230	.002
9	.762	.237	.002
10	.756	.243	.002
11	.749	.249	.002
12	.743	.256	.002
14	.729	.269	.002
16	.715	.283	.002
18	.701	.297	.002
20	.686	.312	.002
22	.671	.327	.002
24	.656	.342	.002
26	.640	.358	.002
28	.624	.374	.002
30	.608	.390	.002
32	.591	.407	.002
34	.574	.423	.002
36	.557	.440	.002
42	.506	.492	.003
48	.454	.543	.003
60	.355	.642	.003
72	.267	.730	.003

TABLE 5.--Predicted probability of various outcomes: "worst" case

Length of time (in months)	P(NONE)	P(MISD)	P(FEL)
1	.139	.499	.362
2	.136	.502	.362
3	.132	.505	.362
4	.128	.509	.362
5	.125	.513	.362
6	.122	.516	.362
7	.118	.520	.362
8	.115	.523	.362
9	.112	.527	.362
10	.109	.530	.361
11	.106	.533	.361
12	.103	.536	.361
14	.096	.543	.360
16	.091	.549	.359
18	.086	.555	.358
20	.081	.561	.357
22	.077	.567	.356
24	.072	.573	.355
26	.068	.579	.353
28	.064	.584	.352
30	.061	.589	.350
32	.057	.595	.348
34	.054	.600	.346
36	.050	.605	.345
42	.042	.619	.339
48	.035	.633	.332
60	.024	.658	.318
72	.016	.684	.303

Comparing these probabilities with those for our "average" individual the most dramatic change is the large increase in the probability of conviction for a felony, particularly in the period immediately after release from incarceration. The probability that such an individual will be convicted of a misdemeanor is not terribly different from that of the "average" individual and the probability of no conviction is much lower than that of this "average" individual.

Since there were no individuals in our sample all of whose characteristics were actually as favorable or unfavorable as those in which tables 4 and 5 are based, these results should be considered illustrative only. They are meant to demonstrate the type of predictions given by our model and the strong influence of personal characteristics on the probabilities of various types of crimes.

## 5. SUMMARY AND CONCLUSIONS

In this paper we have investigated the determinants of one measure of the seriousness of the criminal activity an individual will pursue after release from prison, the misdemeanor/felony distinction. Specifically, we investigate what determines into which of three categories an individual's most serious (as measured by sentence length) conviction during a follow-up period will fall. The categories are no conviction, conviction for a misdemeanor, and conviction for a felony. This dependent variable was analyzed using the logit model, based on a sample of men released from the North Carolina prison system.

Several of our independent variables were found to affect the probability of a conviction of some crime, but not to affect the relative probability of conviction for a misdemeanor or felony. We found that an individual is more likely to be convicted of some crime (misdemeanor or felony) if he is an alcoholic or a user of hard drugs, if he is white, the longer his activities are followed, the more convictions he had before entering prison, and the younger he is on release.

Three variables affect both the probability of conviction for some crime and the relative probability of conviction for a felony or misdemeanor. Having more rule violations when incarcerated or being incarcerated for a misdemeanor increases the probability of conviction for a misdemeanor relative to either no conviction or to conviction for a felony. The younger an individual is when first arrested the greater the probability of conviction for a felony relative to no conviction or relative to a misdemeanor.

Two variables affect only the probability of a felony relative to a misdemeanor, but not the probability of conviction to no conviction. An individual who participates in the prisoner work release program when incarcerated or is supervised when released is less likely to be convicted for a felony compared to a misdemeanor.

As a way of illustrating our findings, we also predicted probabilities of various types of convictions, based on the estimated logit model, for individuals with various values for the explanatory variables. These

predicted probabilities illustrate the strong influence of personal characteristics on the seriousness of postrelease criminal activity as well as the type of predictions possible with this model.

The authors suggest that this model can enrich criminal justice evaluations by allowing them to investigate, in a multivariate statistical context, postrelease criminal activity in its full complexity (frequency and seriousness) rather than merely considering whether or not an individual returns to crime. While we have only used the logit model to investigate one measure of the seriousness of postrelease criminal activity, the model can be used to analyze most other categorical measures of the seriousness or frequency of criminal activity. Judging by past evaluations it seems likely that evaluations which analyze recidivism more thoroughly, by using measures such as the above, will show more programs to be successful than evaluations which concentrate solely on whether an individual again commits a criminal act.

## NOTES

1. A number of examples are contained in Savitz and Wolfgang (1970).
2. Frequency is usually measured as the length of time until return to criminal activity or the number of returns during same time period. Both of these variables would be truncated at zero and skewed, and the number of returns is often handled in such a way as to make it a multiple category (polytomous) variable rather than a continuous variable. (See Wolfgang, Figlio, and Sellin [1972] for examples.) Seriousness is usually measured by time sentenced, or by type of conviction (felony versus misdemeanor) or by one of the formal seriousness indices developed by criminologists. Seriousness indices such as Rossi's (1974) and the felony/misdemeanor distinction are polytomous while continuous seriousness indices such as Sellin and Wolfgang's (1964) are both skewed and truncated at zero.
3. See Theil (1971) for a discussion of the difficulties involved.
4. For example, see Theil (1969), Amemiya (1973), and Amemiya and Boskin (1974).
5. A follow-up period of 3 years is commonly recommended for studies of ex-offenders. See U.S. President's Commission (1967) or Mulvihill and Tumin (1969).
6. See Witte (1975) for a detailed description of sampling and data collection methods. Due to the original purpose for which the data were collected (an evaluation of the North Carolina prisoner work release program), the population from which it was drawn excluded men convicted of sex offenses, serious drug offenses, or as public drunks.
7. With the exceptions noted in footnote 6.
8. The most serious crime is defined as the one for which the longest prison sentence was received. If no prison sentence was imposed, seriousness is measured by the amount of the fine imposed. A suspended or probationary sentence is considered less serious than an active sentence, but more serious than a fine.
9. The felony/misdemeanor distinction in North Carolina is quite similar to most other States. A misdemeanor generally carries a prison sentence of 2 years or less. See Institute of Government (1970) for the details of this distinction.
10. For a detailed description of the logit model, see McFadden (1974), Nerlove and Press (1973), or Theil (1969).
11. The appendix to Schmidt and Strauss (1975) contains an explicit statement of the expression for the various probabilities, the likelihood function, and the information matrix (which is used to calculate standard errors).

12. Wolfgang, Figlio, and Sellin (1972) analyze the determinants of the seriousness of criminal activities for juveniles using multiple regression. Their analysis is only partially comparable to the present analysis for three reasons. First, they are studying juveniles rather than adults. Second, they use a different measure of seriousness. Third, they use an inappropriate statistical technique, multiple regression. However, we have taken direction from their findings where this seemed reasonable.
13. See Service (1972) for a review of these studies.
14. The sample sentence is the sentence that the individual was serving when sampled; the follow-up period begins with the release from the sample sentence.
15. Work release, a program which allows an individual to work at a regular job during the day and return to prison at night, is the major rehabilitative program in North Carolina.
16. All tests are two-tailed tests unless otherwise indicated.
17. It is possible that our results derive from biases in program selection rather than program participation. For example, if particularly reliable individuals are selected for work release and supervision, they may do better after release not because they participated in the program but because of other favorable characteristics for which we have not controlled. This seems more of a potential problem for work release than supervision since more difficult individuals are usually supervised when released in North Carolina. In a previous study (Witte, 1975), all quantifiable characteristics were controlled for and work release still reduced the seriousness of postrelease criminal activity.
18. It is interesting that a number of studies that have gone beyond simple recidivism rates in evaluating criminal justice programs have shown decreases in either the frequency or seriousness of criminal activity although no differences in simple recidivism rates. See Witte (1975), Taylor (1967), and Jesness (1965) for examples.
19. To confirm that these variables were jointly as well as individually insignificant, we carried out a likelihood ratio test of their joint significance. This test statistic is  $-2 \log \lambda$ , where  $\lambda$  is the ratio of the maximized values of the likelihood function under the two alternative specifications and is asymptotically distributed as  $\chi^2_4$  under the null hypotheses that these 4 variables are jointly insignificant. The value of the statistic is 6.569 which is clearly insignificant at normal levels of statistical significance.
20. The sample means are as follows: RACE = .503, WR = .462, ALKY = .481, SUPER = .316, CONVBS = 2.735, RULE = .672, AAR = 3.801, AFA = 2.803, MF = .308. (The variable LENG takes on values from 1 to 72.)

## BIBLIOGRAPHY

- Amemiya, T. (1973), "Regression Analysis When the Variable Is Truncated Normal," Econometrica 41, 997-1025.
- Amemiya, T., and M. Boskin (1974), "Regression Analysis When the Dependent Variable Is Truncated Lognormal, with Applications to the Determinants of the Duration of Welfare Dependency," International Economic Review 15, 485-496.
- Institute of Government (1970), Punishment Chart for Crimes of General Interest in the Superior Courts of North Carolina, Chapel Hill: University of North Carolina.
- Jesness, C. F. (1965), "The Fricot Ranch Study: Outcomes with Small Versus Large Living Groups in the Rehabilitation of Delinquents," Research Report No. 27, California Youth Authority.
- McFadden, D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, ed., Frontiers in Econometrics, New York: Academic Press.
- Mulvihill, D. J., and M. M. Tumin (1969), Crimes of Violence Vol. 12, Washington, D.C.: U.S. Government Printing Office.
- Nerlove, M., and S. J. Press (1973), Univariate and Multivariate Log-linear and Logistic Models, Santa Monica: Rand-R1306-EDA/NIH.
- Rossi, P., et al. (1974), "The Seriousness of Crimes: Normative Structure and Individual Differences," American Sociological Review 39, 224-237.
- Savitz, L., and M. E. Wolfgang (1970), The Sociology of Punishment and Correction, New York: John Wiley & Sons.
- Schmidt, P., and R. P. Strauss (1975), "The Prediction of Occupation Using Multiple Logit Models," International Economic Review 16, 471-486.
- Sellin, T., and M. E. Wolfgang (1964), The Measurement of Delinquency, New York: John Wiley & Sons.
- Service, P. (1972), The Recidivism of Persons Released from Facilities of the North Carolina Department of Correction During January-July, 1968, Raleigh, N.C.: North Carolina Department of Correction.
- Taylor, A. J. W. (1967), "An Evaluation of Group Therapy in a Girls' Borstal," International Journal of Psychotherapy 17, 168-177.
- Theil, H. (1969), "A Multinomial Extension of the Linear Logit Model," International Economic Review 10, 251-259.
- Theil, H. (1971), Principles of Econometrics, New York: John Wiley & Sons.

U.S. President's Commission on Law Enforcement and the Administration of Justice, Task Force Report (1967), Corrections, Washington, D.C.: U.S. Government Printing Office.

Witte, A. D. (1975). Work Release in North Carolina: An Evaluation of Its Post Release Effects, Chapel Hill, N.C.: Institute for Research in Social Science.

Wolfgang, M. E., R. M. Figlio, and T. Sellin (1972), Delinquency in a Birth Cohort, Chicago: University of Chicago Press.

## MANAGEMENT BY EVALUATION: PUTTING RESEARCH TO WORK

Dale Shears  
Evaluation Program Coordinator  
Michigan Department of Social Services

How many times have we watched a research report die?

After great labor, the report is finally finished. Another forest goes down, in order to supply the copies for its distribution. But the busy manager, with a desk full of urgently pressing items, gives the report only a quick glance--perhaps reads only the Summary--and then files it away. There is no impact. He can't use all those numbers.

I believe that a saddening amount of research effort is wasted every year in juvenile corrections because it fails in application what it has dared in design. It fails to move the manager.

For too long, management and evaluation have been a house divided in the rehabilitative sciences. Most research has been carried out by external monitoring agencies, often on a one-shot basis connected with grant monies. The price of this kind of organization is a high degree of irrelevance to the decisionmaker. As a result, research in juvenile corrections has had dismayingly little impact on services delivery.

As a manager, if you want to know whether you're getting where you want to go, you have simply got to take stock along the way. Management and evaluation, in short, are not really separate functions. They are aspects of the same function: serving our clients as intelligently and thoroughly as we are able.

Today I would like to share with you what we've been doing in Michigan's juvenile institutions over the past 5 years to try to bring about a marriage between evaluation and the decision process. Along the way, I'd like to discuss some of the things we've discovered that help juvenile offenders "make it." We call this research organization Management by Evaluation.

As a person who has found himself in the trenches of both program management and research, I believe that before an evaluation program is ever going to achieve management relevance, it must stick to four principles:

The first principle is that evaluation must be inseparable from the management process. It must not be a series of one-shot affairs designed to pacify various funding sources.

Instead, research must be organized as a management function--a regular and continuous feedback process internal to the agency itself. Key evaluation personnel should share management tasks and functions. At the very least, they must be regular participants in the issues and discussions that surround the group decision-making sessions of the agency. Without continuous contact and participation with management circles, research gets hopelessly out of touch.

The second principle is that evaluation efforts must achieve and then maintain continuity of procedure and instrumentation--across programs, and across time. Comparability of data is critical. If one program uses instrument A, another uses instrument B, and a third uses instrument Z, and then they all change over to "new and improved" instruments every year, what have we got? We have succeeded in resurrecting the myth of the Tower of Babel. We stand surrounded by all the data in the world--and no information at all.

The third principle is that evaluation efforts must focus on client objectives and outcomes as first priority. In a service agency, the client is the cornerstone of all endeavor. Substantial efforts here will be rewarded by better agency and worker support. "People listen when outcomes talk." Desirable and undesirable outcomes must be agreed upon, defined, made measurable, and assessed.

The fourth principle is that objectives and measurements must be renegotiated regularly to sharpen efforts and maintain commitment.

In the next few minutes, I would like to explore these four principles in greater detail, utilizing examples from our development in Michigan.

## HISTORY

We began in 1972 with a Steering Committee that developed an initial research design for Project STEADY, a Federal grant. From the beginning, however, the evaluation component was designed to outlive the grant and to achieve an in-house evaluation capacity. And from the beginning, the evaluation program was shaped by program managers--people who were experienced in working with institutionalized delinquent kids, and who were therefore in touch with some of the practical goals and problems.

Let me illustrate with an isolated example: We might have used the traditional paper/pencil personality tests, but the Steering Committee knew how poorly our kids read--and so we developed a true/false format that enabled us to test kids orally. We now use audiotapes for standardized testing, and the instrument has worked out very well.

An agency will probably be better off to develop and staff its evaluation program by people who have some knowledge of both the system and the kids. From a management viewpoint, outside evaluators seldom seem to raise

the necessary questions. As a consequence, when they speak they are too seldom heard. Good research is not done to an agency, it is done by and agency.

In the Management by Evaluation system, I want to describe, the priority research questions to ask are the ones that instruct managers in their decisions. Managers help negotiate both objectives and measurements-- and they are also accountable to those objectives and measurements.

## INSTRUMENTS

In Michigan's juvenile institutions, we have been trying to achieve Management by Evaluation for 5 years. We are using a variety of instruments. There are pre/post measures, which include an attitude/value scale called the Youth Opinion Poll, as well as educational achievement testing on the Stanford Achievement Test. There are process measures, which include a weekly Behavior Checklist as well as various management compilations such as length of stay and truancy rates. And there are outcome measures which assess arrest outcomes, job/school outcomes, and placement outcomes at 3 and 12 months following release.

At this point, the basic instrumentation has been in use long enough to permit us to establish baselines of achievement for each of our treatment centers, and to go on from there to establish objectives relevant to those baselines.

The Institutional Services Division in Michigan operates nine treatment centers, which turn over about 700 youth per year. Starr Commonwealth, a private agency, is now providing an additional source of comparative data. Data flows from all Division programs to a centralized Data Center, which stores, interprets, and feeds back information into a regular 6-month management cycle.

In order to compare programs of different size and length of stay, we make heavy use of ratio measures. For example, we divide academic gain by length of stay to produce a standard Education Index, and then compare educational effectiveness at different treatment centers. Programs that aren't doing the job quickly surface, and corrective action is begun.

Similarly, in studying truancy rates, we take the number of successful truancies at each Center, and then divide it by the average daily number of youth in care for the period studied. The resulting Truancy Index caused us to discover recently that some of our minimum security programs had 38 times more truancy per capita than others, and that our so-called "maximum security" programs were hardly that secure.

The booklet entitled The Institutional Centers, which you have before you, details 1976 management objectives and rates of achievement. Appendix A lists current goals and means of measurement.

## MANAGEMENT PROCESS

At this point, we have settled into a regular management process with the program managers of the residential Centers. The Data Center feeds back information continuously, but produces summary reports every 6 months. Feedback is then discussed and processed in individual management conferences at each treatment center. Out of these conferences come ideas for deleting, adding, or modifying goals or measurements to make them more useful to the managers. They also generate action plans to solve the problems that emerge.

In a final meeting of all the Division managers each spring, the goals and measurement procedures for the year ahead are negotiated, and the management process comes full circle. The new goals and objectives become the focus for the next go-around of evaluation activity.

Let me illustrate the two-way nature of this process with an example: Back in 1975, we developed our first goal in the area of length-of-stay. This goal had two parts: One was to increase the percentage of releases in the 4-7 month period; the other was to decrease the percentage of releases occurring after 11 months. Baselines were established and programs monitored. Quick turnaround was the objective.

In May of 1976, Center Directors and Division administrative staff decided to modify their 4-7 month release goal to read "maintain (not increase) the percentage of youth released within 4-7 months." One of the reasons for the shift in focus was research evidence that educational achievement level at release was related to arrest outcome success. In light of this finding, we didn't want to keep pushing for ever-shorter lengths of stay when it meant we couldn't accomplish as much academic gain.

Since last year's meeting, however, the evaluation program has unearthed evidence indicating that (1) length of stay is unrelated to outcome success, but (2) age at release is directly related to successful outcomes, possibly because older kids have greater opportunities in the area of employment, etc. What we now know is that if a youth is released under the age of 17, his re-arrest probability within 3 months is 1 in 3, but if he is released at age 17 or older, this probability drops to 1 in 5.

This finding has potential impact on our future determinations of appropriate length of stay. Programs are already using this information in making individual case decisions, but in May this year when we meet again, we are going to have to take a hard look at our length of stay policy, particularly in our shorter programs, in light of what seems to help kids.

What our management objective for length of stay is going to be after May I can't even guess. But I have detailed the evaluation of this objective as one example of the interaction we are accomplishing between management decisions and evaluation feedback.

## OUTCOMES

As you may already have gathered, the bottom line in the whole process is client outcome. Outcomes are crucial because, in the final analysis,

they are the yardstick against which every aspect of program, and of program evaluation, is measured. We don't put millions of dollars into special educational programs in our institutions out of a unique love for the art, but out of a profound hope that the additional educational skills will help kids make it. At bottom, everything we do comes back to (or should come back to) the kids.

In the rehabilitative sciences, we are very fortunate in one regard. Our overall objective--the rehabilitation of people who have been doing criminal things--is relatively clear. Consider, for example, the problem faced by a mayor. What is the guiding objective of a city? Lowering taxes? Raising employment? Building roads?

We have selected an outcome index based upon police/court dispositions; that is, the arrest process. We follow every youth into the community by telephone, using a structured interview with the youth's Community Care Worker. (Because our youth are State wards, when they leave us they will in almost every case return to the care of a Community Worker.)

With this procedure we have been able to achieve 96 percent successful tracking through 3 months (that's based on almost 1,200 kids), and 90 percent successful tracking through 12 months. In effect, we now have a tracking vehicle in place that gives us an ongoing capacity for longitudinal outcome study.

The principal outcome scale is a four-point Arrest Outcome Scale. We ask the Community Care Worker whether his youth (a) has had no police or court contact; (b) has had some contact short of arrest; (c) has been arrested or charged with an offense; or (d) not enough knowledge to say. If the youth has been arrested, we also determine whether the charge against him was for a felony, a misdemeanor, or a juvenile status offense.

In addition to arrest outcome information, we also collect information on job/school outcomes, and on placements subsequent to release.

Many other outcome measures were considered and rejected as measures of program effectiveness. To name several, we might have used recidivism (meaning return to one of our institutions); or survey instruments; or rating scales; or court records; or convictions.

Of all these alternatives, court records are a good source regarding the arrest event, but the courts normally don't have good case information in the employment, school, and placement areas.

Recidivism figures can be very misleading. Most youth, we have discovered, do not return to State institutions if they encounter future difficulties. They go instead to other juvenile settings or to adult corrections facilities. Indeed, a large number of youth arrested on felony charges remain right in the community on some form of court probation or observation. In short, while recidivism figures can make a program look very good, they are hardly an accurate index of program effectiveness.

Questionnaires and survey instruments are almost always only partially returned, and therefore their validity is distorted.

Rating scales of youth success tend to have serious reliability problems because they are not behaviorally specific.

The conviction record is similarly inadequate. Pretrial periods are too long, and processes like plea-bargaining are too erratic to measure well. Counting only convictions puts the focus on a legalistic definition of outcome that has little utility for people working in the rehabilitation process.

We finally selected the arrest outcome index. It is behaviorally specific, and highly reliable when checked against police and court records; it distinguishes between harassment and formal arrest action; it may readily be completed over the telephone; it permits collection of other valuable outcome data on employment, school, etc.: and it most directly measures what institutional delinquency services workers strive to prevent: further formal action by the juvenile justice system against youth that have been released.

#### NEWSLETTER

Information on the outcomes of each youth is written into a monthly Evaluation Newsletter, which is fed back to institutional workers. It lets them know how their kids have turned out, and at the same time provides valuable feedback with which they can sharpen their case-management skills. In fact, at this point several treatment centers have begun using the Newsletter to systematically explore their own placement failures.

In the past few months, we have finally arrived at our first relationship studies between program variables and felony arrest outcomes.

In the remaining minutes we have, I would like to sketch out some of the relationships we have been discovering. We are finding that:

- 30 percent of our male youth are arrested within 3 months of release; and 50 percent are arrested within a year. (Girls = 11 percent and 20 percent.)
- For males, almost 9 out of 10 arrests are on felony charges. (For females, only about 1 in 3.)
- The first 3 months following release seem to be a critical period. The arrest rate in the first 3 months is almost five times what it is in the next 9.
- Length of stay is not related to outcome success but age at release is directly related to arrest outcome success. The older the youth released, the better his chances of success.
- Educational achievement at release is related to outcome success.
- Job/school productivity is strongly related to success at avoiding arrest, but fully 59 percent of all our youth are doing nothing whatever productive (no job and no school) 3 months after they leave us.

- Although both jobs and school seem helpful, jobs seem more effective than schools in helping youth avoid arrest.
- Brighter youth are more often found in jobs; academically handicapped youth are more often found in schools.
- In general, the social and academic level that kids reach by the point of release (rather than how much they've gained while we've had them) seems to predict success.

You will find detailed reports of these studies in Appendix B of the Institutional Centers report.

#### SUMMARY

In the past 20 minutes, I have tried to describe for you the practical value of putting research to work through a management process we call Management by Evaluation. It is my firm conviction that you do not have to have a Hollywood Budget to do useful research. Nor is computer technology necessary.

We have put together this research program with three people and nothing more complicated than a hand calculator. I believe that similar programs could be initiated at low cost wherever management is sufficiently committed to exploring the results of what it is trying to do.

## DEVELOPMENTAL HISTORY AS A METHOD OF ORGANIZATIONAL EVALUATION

Jackwell Susman  
Institute for Advanced Studies in Justice  
American University

The ability of an organization to achieve its goals, to change its goals, and to fulfill the needs of its employees and its clientele is largely a function of its leadership (Etzioni, 1964). Yet leadership is a relatively neglected dimension of the evaluation of organizations. Furthermore much evaluation is cross-sectional rather than longitudinal despite the fact that organizations change over time. The early years are as formative for an organization as for a human being. Sometimes the first leader is strong enough to leave his mark on successive generations of an organization. In most cases, however, the first leader is replaced by one of almost opposite qualities, and this process is repeated thereafter. Pareto referred to this as the "circulation of elites": innovative leaders (he called them Foxes) are replaced by Lions who are concerned with survival (Parsons, 1968). Selznick (1957) suggests that Pareto's cyclical theory of social change could be salvaged and reformulated for use in the study of specific organizations rather than of whole societies and historical epochs.

This Yin-Yang conceptualization of organizational leadership needs a methodology that can capture its essence. Evaluation of an organization, if it utilizes a cross-sectional analysis, will almost completely misunderstand, misjudge, and misstate the social processes that influence organizational effectiveness. These social processes, and particularly the leadership process, insofar as they are dependent on changing circumstances both within the organization and between the organization and its environment (Evan, 1966). We need to move from reliance on cross-sectional studies to longitudinal studies utilizing time-series data. Furthermore, we need to move from case studies to comparative studies, that is, the systematic collection of information on relevant variables related to leadership and organizational effectiveness across a number of similar and disparate organizations (Scott, 1975). Only in this way can we isolate the variable and invariant aspects of organizational effectiveness.

This paper reports on an effort to apply Pareto's conception to a longitudinal case study of a public defender office in the District of Columbia. Data on the structure and functioning of the organization were gathered in two ways: organizational documents and interviews with expert informants. This follows the methodology suggested by Pennings (1973).

A particularly important issue, and one that goes to the heart of a study of a public defender office, requires that we put some relevant

content into Pareto's formal structural concepts. What does it mean to be an innovative leader in a public defender office? On the other hand, how do we characterize the Lions? To begin with, one of the major problems for the criminal justice system is that of the maintenance of conflict, or the control of tendencies toward cooperation (Skolnick, 1967). Administrative requirements produce strains both within organizations and between organizations toward "deviant behavior," i.e., cooperative relations. This tendency toward accommodation extends to the private bar, especially that strata of the bar that practices criminal law and generally is not part of a large law firm (Carlin, 1962). As far as defense attorneys are concerned these accommodationist pressures directly affect client control. Client control, how the lawyer gets the client to do what the lawyer wants, is yet another major problem of the criminal justice system that is partly bound up with but is also independent of pressures toward cooperation. The meaning of leadership and the effectiveness of the public defender system can best be assessed in terms of balancing these two problems.

How the original leaders of the D.C. Public Defender Service (known as the Legal Aid Agency until 1970) conceived of these problems and the steps that they took to resolve them had its roots in the Public Defender social movement of the early 1960's. This social movement was sparked by the United States Supreme Court's reassertion of its control of the criminal justice system in the late 1950's (Criminal Law Reports, 1973; McCloskey, 1972) and fueled by academic and public dissatisfaction with injustice and the traditional legal representation of the criminally accused by the bar. The ideology of the public defender movement combined, as do most ideologies, both elitist and democratic principles in its conception of reality and strategy for change (see, for example, Journal of the State Bar of California, 1966). At the risk of considerable oversimplification the two major principles of the ideology may be stated in terms of the role and function of public defenders. The sense of injustice was crystallized into the notion that the mass of the criminally accused are denied equality of legal opportunity (Packer, 1968); furthermore the system is not honest, and the complicity of the bar permits this dishonesty. Criticism of the prevailing norms of the criminal justice system and the efforts at reform which it engendered, were manifest in the creation of the Legal Aid Agency (later the Public Defender Service) of the District of Columbia in 1960. The ideology also implied that there were conduct norms that should be conformed to by public defenders.

The earliest leadership of the D.C. agency had the responsibility of translating the ideology into practice. The formal leader of the organization had a laissez faire attitude toward the management of the staff lawyers and a feeling that each of them were independent professionals. There were also differences in education, age, and background between him and the professional staff that made peer communication difficult. The informal leader, whose position of influence was later formalized, was able to exercise enormous power in defining organizational goals and recruiting and training the staff lawyers. He perceived himself as a member of the public defender social movement and attempted to operationalize its ideology. He believed that the role of the public defender was to keep the system honest and their function was to serve the poor. He tried to move the organization toward these sometime mutually exclusive goals of reform and service. The goal of reform of the system had two themes. The major innovative theme was an

attempt to abolish plea bargaining; and the minor, and more cautious, theme, was to develop test cases (see, for example, Brill, 1973). Client service also took on two faces; first, naturally enough, giving every man his day in court and second, providing nonlegal services to the agency's clientele. The latter was, despite its obviously innovative nature, much less likely to "rock the boat" than the former. Thus over and above the legislatively stated purpose of the agency--"make attorneys available to represent indigents"--the first leader of the organization established the contrapuntal policies that were variously adopted by the subsequent Foxes and Lions.

His concern for service led to the creation of a unit within the Agency which provided nonlegal services to indigents accused of crimes, principally outpatient psychiatric referral, job training, and employment (Institute of Criminal Law and Procedure, 1970). Community Services, as this unit was originally known, had a full-time director and part-time social worker; it also prepared presentence reports on each convicted client which could be used by defense counsel to recommend an alternative disposition and/or to challenge the court's presentence report. At its inception, this organizational activity met with strong resistance from staff attorneys who were not service oriented; their relationship and interest in their client was based solely upon their client's legal problem (which in any event ended with a finding of innocent or conviction) and the experience they would gather as they solved it. The leader finally had to issue an order requiring that the staff submit the names of defendants who pleaded guilty or were found guilty after trial to the Director of Community Services so that a presentence report could be prepared. This function was also made available to non-Agency attorneys representing indigents but it was not widely utilized by them probably for similar reasons. This lack of interest reflected the private practicing lawyers' conception of their role and responsibility to their clients which traditionally encompassed only the clients' legal problems and terminated when the trial was concluded, except when an appeal was entered.

The principal effort in terms of client service was but the other side of the coin of system reform. That every person should have his day in court was concomitant with the desire to abolish plea bargaining. Although justified by the leader and some staff on the grounds of equality of legal opportunity, reducing the number of defendants who plead guilty could not in itself be legitimated in terms of client services because most trials ended in conviction and defendants convicted after trial were generally more severely punished (sentenced) than those who plead guilty. Not only was pleading part of the system being changed, its latent dimension reflected on professional behavior. The public defender movement entertained a rather common stereotype of most criminal lawyers appointed to represent indigents; that the latter had low status in the profession (Watson, 1965) and their professional behavior as regards the existing legal order was uninspired and unchallenging. Or, as more colorfully described, defense attorneys function to "cool the mark out" and are Faustian if not immoral (Blumberg, 1967). In particular, these lawyers were thought to plead their clients guilty while "good" lawyers were believed to fight it out in court. A recent article by Sudnow (1965) made the same point and was criticized for doing so by Skolnick (1967). Earlier studies (1935), more quantitative but less analytical than Sudnow, led to similar conclusions and probably contributed to the stereotype. As Skolnick suggested and the present study

makes explicit, there are problems of motivation and leadership to be considered. PDS reformers, casting themselves in opposition to the stereotype, were attempting to institutionalize new behavior patterns, at the same time they could legitimate their spending more time and energy in the traditional (and romantic) workplace of the lawyer, the courtroom. The PDS attorneys, arguing with the Board of Trustees of the PDS for more trials, implied that they were merely seeking to emulate the best practicing attorneys retained by some accused criminals. Some lawyers, including some on the staff, expressed privately doubts about the stereotype but both to protect this myth and keep the goodwill of the Bar, the PDS only twice in its history published statistics on the disposition of cases represented by private retained counsel.

The reformist tendencies of the leader were manifest, then, in the desire to reduce plea bargaining to zero if possible. This would in part resolve the tension produced by administrative requirements toward cooperation in favor of maintaining the purity of the adversary system. But this effort on the part of the public defenders was likely to work against the interests of the client. On the one hand, conviction rates by trial are high; on the other hand, defendants who are convicted after trial are usually sentenced more severely than defendants who plead guilty. As a result, client control became a much more difficult and exaggerated proposition. The public defender movement and the agency's leadership were disposed to view client control as meaning that both the strategy and tactics of the case should be under the client's control, one of many interpretations of client control (Skolnick, 1967). The attorney offers the client his technical knowledge of various branches of law to implement the client's strategy. The public defender movement allowed for (in fact almost required) the possibility that a trusting relationship might develop between the attorney and client and the latter would seek and act on the advice of the former. What the public defenders did not expect and found hard to cope with was the hostility of their clients toward them. The racial dimension of this hostility grew year by year. But beyond that there was hostility growing out of the feeling by the accused that the public defenders, paid by the State, were part of the system and served the interests of the system (Casper, 1972). So it would seem that the public defender's clients were not always inclined to play the role the movement ideology assigned to them.

The implications of these observations for assessing organizational effectiveness are twofold. First, they remind us that the effectiveness of the public defender service may mean different things if we view it from the standpoint of the clients, the staff attorneys, or the leadership. Furthermore, there are constraints that influence the agency's goals and their attainment regardless of how these matters are resolved. Pressures toward cooperation and client control must be considered and dealt with whether the leadership is a Fox or a Lion.

We have explicitly chosen to view the effectiveness of the public defender service in terms of the perspective of its leadership. We are now in a position to clarify somewhat what we mean by an innovative leader. Our study of the D.C. Legal Aid Agency and its successor, the D.C. Public Defender Service, makes it clear that its first leader thought of his

organization's success in terms of two major goals: a low plea bargaining rate and a low trial conviction rate. No two measures, of course, can describe a complex structure such as an organization. Nevertheless, these indices which single out certain quantitative aspects are helpful in reducing the complexity of reality to a form which our inadequate minds can grasp.

In achieving these goals he was not altogether successful. As table 1 indicates, during the first leadership period, the years of the Fox (1963-1965), the agency's success rate in felony trials was rather low, about 22 percent. The plea bargaining rate in felony cases was about 47 percent; during the same period, the felony trial conviction rate for the District of Columbia was 55 percent (President's Commission on Crime in the District of Columbia, 1966).

The years of the Lion that followed (1966-1968) saw an increase of the felony plea bargaining rate to about 52 percent of the agency's felony cases. But the agency was also winning more of its felony trials; the rate increased to about 35 percent.

The cycle was repeated. The Lion was followed by another Fox (1969-1971). During the years of this leader the felony plea bargaining rate fell to about 23 percent, very low. Furthermore the agency was winning many more of its felony trials; the rate went to about 47 percent.

The successor was another Lion. The felony plea bargaining rate increased to about 39 percent. In addition, the agency was winning still more of its felony cases, about 52 percent.

In conclusion, it appears that there was a circulation of elites in leadership positions in the Public Defender Service of the District of Columbia. During the 11 years there were four leaders: a Fox, followed by a Lion, followed by a Fox, followed by a Lion. The basic policies, both major and minor, were established by the first leader, a Fox. He pursued the major innovative policies but laid the groundwork for the less innovative policies. The innovative policies were to limit plea bargaining and, of course, go to trial more frequently. The policies with more conservative implications were the creation of a community services unit in the agency and the effort to develop test cases. His organization was successful in accomplishing his goal of keeping the felony plea rate low but they had a low level of winning felony cases. This was not in the interests of their clients. However, the first leader was, as predicted by Pareto's theory, followed by a Lion. The pattern he established, a high felony plea bargaining rate was perhaps balanced by winning more of the felony trials. The next leader was also a Fox, and the plea bargaining rate declined; however, the proportion of felony cases won continued to increase. The last leader considered was also a Lion; the plea bargaining rate increased again, although the win ratio of felony trials continued to increase. One conclusion of this study is obvious; the outcome of trials is not a good measure of effective leadership. It is important to note that while the Lions were de-emphasizing low plea bargaining as a goal they were emphasizing the minor policies of the first leader. There was an enormous growth in the Community Services unit, which was in the interests of the agency's clientele. There was also increased emphasis upon test

TABLE 1.--Periods of leadership

Leadership periods (fiscal years)	Average number of attorneys including Director and Deputy	Average felony caseload	Average number of felony cases assigned or closed	Average number of guilty pleas	Rate of guilty pleas	Average number of cases found not guilty or dismissed	Rate of not guilty and dismissed
I (1963-65)	9	22	199	93	.467	44	.221
II (1966-68)	14	11	141	74	.524	50	.354
III (1969-71)	34	15	522	121	.227	187	.442
IV (1972-74)	46	24	1,094	425	.388	566	.517

Source: Annual Reports of the Legal Aid Agency of the District of Columbia, 1963-1969; Annual Reports of the Public Defender Service for the District of Columbia, 1971-1974.

cases; in whose interests these may be depends upon the issues and, importantly, how court decisions are implemented by the leadership of the criminal justice system. This is yet another parameter beyond the control of public defender leadership.

## BIBLIOGRAPHY

- Beattie, Ronald H. The Public Defender and Private Defense Attorneys. Berkeley: University of California, 1935.
- Blumberg, Abraham S. "The Practice of Law as a Confidence Game," Law & Society Review, 1:27-39, 1967.
- Brill, Harry. "The Uses and Abuses of Legal Assistance," The Public Interest, No. 31, Spring 38-55, 1973.
- Carlin, Jerome. Lawyers on Their Own. New York: Russell Sage Foundation, 1962.
- Casper, Jonathan D. Criminal Justice: The Consumer's Perspective. Washington: U.S. Government Printing Office, 23-29, 1972.
- Editors of the Criminal Law Reporter. The Criminal Law Revolution and Its Aftermath. Washington: Bureau of National Affairs, 1973.
- Etzioni, Amitai. Modern Organizations. New Jersey: Prentice-Hall, 61-67, 1964.
- Evan, William. "The Organization-Set: Toward a Theory of Interorganizational Relations" in Thompson (ed.), Approaches to Organizational Design. Pittsburgh: University of Pittsburgh Press, 175-191, 1966.
- Institute of Criminal Law and Procedure. Rehabilitative Planning Services for the Criminal Defense. Washington: Georgetown University Law Center, 1970.
- Journal of the State Bar of California. Symposium on Legal Services to the Poor, 24:214-311, 1966.
- McCloskey, Robert G. The Modern Supreme Court. Cambridge: Harvard University Press, 1972.
- Packer, Herbert L. The Limits of the Criminal Sanction. Stanford: Stanford University Press, 1968.
- Parsons, Talcott. The Structure of Social Action. New York: Free Press, 1:278-293, 1968.
- Pennings, J. "Measures of Organizational Structure: A Methodological Note," American Journal of Sociology, 79:696-704, 1973.
- Scott, W. Richard. "Organizational Structure" in A. Inkeles, J. Coleman, N. Smelser (eds.), Annual Review of Sociology, 1:1-20, 1975.
- Selznick, Phillip. Leadership in Administration. Evanston: Row, Peterson, 112, 1957.

Skolnick, Jerome H. "Social Control in the Adversary System," Journal of Conflict Resolution, 11:52-70, 1967.

Sudnow, David. "Normal Crimes: Sociological Features of the Penal Code in a Public Defender Office," Social Problems, 12:255-276, 1965.

Watson, Andrew S. "On the Low Status of the Criminal Bar," Texas Law Review, 43:289-311, 1965.

CRIMINAL JUSTICE INFORMATION NEEDS OF ILLINOIS  
LOCAL GOVERNMENT OFFICIALS\*

S. Burkett Milner, Ph.D.  
Director, Social Justice Research  
Associate Professor  
Social Justice Professions  
Sangamon State University  
Springfield, Illinois

Barbara E. Ferrara  
Research Specialist  
Center for the Study of  
Middle-size Cities  
Sangamon State University  
Springfield, Illinois

INTRODUCTION

The criminal justice literature has generally ignored the role of local government officials (LGOs) as criminal justice decisionmakers. According to the Statutes, local government officials do have a criminal justice role. While that role is not always easily or clearly recognized, their responsibilities for planning, budgeting, and the development of policy are fundamental to the administration of justice on the local level. Indeed, it has been asserted that how LGOs fulfill their role directly affects the quality of criminal justice services provided.

This paper highlights the methodology and some of the major findings of a project entitled, "The Criminal Justice Awareness Project." Funds to support this project were obtained from the Crime Prevention Commission, Jacksonville, Illinois, and the Illinois Law Enforcement Commission. Determining the nature and extent of the criminal justice information needs of local government officials was one of the research activities.

RESEARCH QUESTIONS

In Illinois, criminal justice services are organized, administered, and financed largely at the local government level. This, together with the fact local governmental decisionmakers are, for the most part, laymen with respect to criminal justice, poses the questions: Do local government officials have sufficient information to make sound decisions about criminal justice matters? If not, what are their informational needs? Are LGOs interested in the results of criminal justice research and evaluation? and What is the likelihood they will utilize any technical information offered them, including the results of criminal justice evaluation?

---

\*Support for this study was obtained through Grant #2056 provided by the Illinois Law Enforcement Commission.

## OPERATIONAL DEFINITIONS

Prior to outlining the methodology, note should be made of how some key terms used in the study were operationally defined:

Local government refers to political units known as counties, municipalities, cities, and villages.

Local government officials (LGOs) refers to members of a body of elected representatives (boards, councils, commissions) and appointed chief administrators (city managers, county managers, or administrators) who together constitute the governing authority of a local government unit.

Criminal justice refers to the functions and activities of police, courts, probation, parole, and corrections.

Knowledge refers to awareness of a fact, concept, technique, or process that may be gained through experience or instruction. Respondents who said they were familiar with a concept, had observed a process, and cited appropriate sources of information were considered more knowledgeable than those who said they were unfamiliar, had not observed, and were unable to cite appropriate information sources.

## RESEARCH PROCEDURES

This was not an evaluation study per se but it could be considered a formative evaluation, an exploratory first step, since certainly a great deal about the predilections of local government officials was learned.

The major objective of the study was to identify the informational needs of local government officials in Illinois with respect to criminal justice matters. To assess these needs meant in part asking the officials themselves what their needs were and determining what was being done already to reduce these needs. But it also required determining what knowledge base were the officials bringing to criminal justice decisionmaking. This meant, in turn, that to measure their knowledge levels, a baseline or standard first had to be specified.

The methodological activities for the study were therefore divided into two phases: Phase One asked, What should they know? and What can they be tested on? Phase Two asked, What do they know?

In attempting to bridge the gap between social science method and practical realities, the political sensitivities involved in doing this kind of research presented some difficult problems. With this in mind, clearly the officials could not be tested on every facet of the field of criminal justice. And the practical limits on specialized knowledge expected of the more often than not part-time local governmental decision-maker had to be considered. A body of knowledge had to be narrowed to that conceivably related to their statutory responsibilities for the administration of justice.

The method chosen for delimiting the field was judgment sampling. In a survey of practitioners, the universe of members of three professional

criminal justice associations in Illinois were asked, in a mailed questionnaire, to rate selected concepts, standards, techniques, and legislation-- in terms of their importance for local government officials to know.

A second judgmental sample--a panel of experts--was used in tandem with the survey of practitioners to help derive a standard. Eight known experts, familiar with administration of justice in the State of Illinois and with Illinois local government, served on the panel. Together they were broadly representative of the professions associated with the criminal justice system and, also, all with a background in local government. They were chosen according to their reputational base and inclusion on national commissions.

The panel was asked to consider the question: What do local government officials need to know in order to make sound decisions about criminal justice matters? An interesting feature of the research was the use at this point of a small group interaction technique known as Nominal Group Process. The technique was used to structure the panel's agenda. Its advantage was it ensured that each expert had an equal opportunity to propose and clarify items in response to the question.

The product of the nominal group process was a plurality of opinion as to what local government officials need to know about criminal justice. Involving experts in the identification of items supplied a legitimating base for the instrument to be taken into the field.

The benchmark or base line that came out of the two judgment samples encompassed four broad areas: concepts, processes, techniques, and laws. Among the top concepts were:

- (1) A system understanding of criminal justice.
- (2) A broader definition of crime climate.
- (3) The concept of due process.
- (4) Mythological assumptions about the system.
- (5) The broad spectrum of antisocial behavior.

The processes identified as important for the officials to know included system-offender transactions and steps in the general planning process. The legal aspects included standards for operating agencies and for facilities and the statutory definition of the criminal justice role of the local government official. The techniques included in the base line related to planning, budgeting, and financing, and to utilizing results of criminal justice research and evaluation.

Moving now from getting a baseline to doing the field work: A two-stage sampling technique was used in the study. The first stage involved drawing a stratified random sample--stratified on the basis of population--of Illinois local government units. The universe consisted of all 101 counties in the State (Cook County was excluded at the suggestion of the

funding agency) and all municipalities having a population of 5,000 or more (excluding those within the bounds of Cook County). The sample, which included 50 municipalities and 26 counties, was found to be representative of the State in terms of population of the government jurisdiction and geographic location within the State.

In the second stage, individual members of the government units were chosen. The process for selecting respondents was:

- (a) The chief administrator of the jurisdiction was always chosen.
- (b) Then by random process a member of a criminal justice committee was chosen.
- (c) Last, a member-at-large was chosen, again by random process.

Using the results of the judgment samples, an interview guide was prepared, pretested, and revised. The instrument contained a variety of items, including:

- (1) Demographic characteristics of the officials.
- (2) Criminal justice information already offered.
- (3) Kinds of criminal justice decisions made.
- (4) Persons and agencies consulted in making criminal justice decisions.
- (5) Familiarity with regional planning units.
- (6) Observations of facilities and processes of the criminal justice system.
- (7) Knowledge of legal constraints and requirements.
- (8) Perception of LGOs role in relation to criminal justice.
- (9) Perception of local criminal justice problems.
- (10) Satisfaction with available criminal justice services.
- (11) Attitudes toward intergovernmental cooperation.
- (12) Informational needs and methods for reducing these needs.

Two hundred and four (204) personal interviews were sought. A total of 186 interviews, averaging 1 hour in length, were obtained with the help of the State and regional law enforcement commissions. This meant a 91 percent response rate. The respondents were found to be representative of the universe in terms of population of their jurisdiction.

## FINDINGS

A. Criminal Justice Role and Activities. The criminal justice role of local government officials is neither foremost in the minds of the officials nor perceived to be important by many of them. When asked, "As a local government official, what kinds of criminal justice decisions do you make?," 43 percent of the officials responded they do not make criminal justice decisions. Other officials reported that they approve recommendations, but added that they regarded this as insignificant. One official commented, "My role amounts to so little, it doesn't really matter."

A number of officials characterized their role in relation to criminal justice as "minimal" or "indirect." Some of these responses reflected a certain concept of agency management. Some local governments assign the responsibility for overseeing specific agencies to committees. In these cases, committee members have a great deal of control over agency policy, practices, and resources, while those officials not on the committee have little, if any, say about the agency. The role of the latter may be merely to review the committee's recommendations. In commission forms of government, the responsibility for a given operating agency may be assigned to a single commissioner, with the commissioner rarely modifying his or her recommendations.

The comment of one official with regard to his role was echoed by several: "I try to be as little involved as possible. I don't believe in infringing on department heads." Since agency-head positions within counties tend to be elective offices, this orientation seems more typical of county than city officials.

It was anticipated that most LGOs would feel they had little to do with the administration of justice. A checklist of activities that LGOs could perform in relation to criminal justice operations was therefore included in the instrument. The activities were budget preparation, problem identification, goal setting, priority setting, program development, long-range planning, short-range planning, program evaluation, personnel decisions, seeking grants, and monitoring agencies and programs.

Significantly, the majority of officials reported they did engage in each activity with the exception of monitoring agencies and programs. Budget preparation was the activity in which the greatest number (85 percent) of LGOs were involved. Nearly three in four officials engaged in short-range planning and seeking grants. Thus it was found that local government officials are more involved in criminal justice planning than most of them realize.

B. Criminal Justice Decisionmaking. A series of questions were posed to determine from what agencies and persons local government officials seek information when making criminal justice decisions. The first question asked the officials what specific agencies at any level they consult when making decisions about criminal justice matters. Twenty-eight percent of the officials stated they do not consult any agency when making decisions. Of those who said they did, the majority (62 percent) mentioned only local agencies. Only a small number of officials said they consult a State or

Federal agency and, of those, a greater proportion were LGOs of cities, not counties.

The second question asked the officials which persons presently employed by the criminal justice system they consult when making decisions about criminal justice matters. One in 11 LGOs said they do not consult any such person. Over 80 percent of the rest mentioned only local system personnel. Fewer than 14 percent specified State in addition to local personnel.

The third question in the series asked the officials to identify which persons not employed by the criminal justice system they consult when making criminal justice decisions. Two of three LGOs indicate they did not consult any person outside the system. A higher proportion of county officials (80 percent) than city officials (59 percent) so indicated. Of those reporting they do consult persons outside the system, half consult other local government officials. This suggests that, in many cases, input is sought from merely a "circle of intimates." Only one in 10 LGOs identified a former employee of the criminal justice system as someone he or she consults and only one in five indicated they consult citizens. Only eight respondents mentioned consulting persons in the community who have expertise in criminal justice, such as college faculty.

Thus it was found that local government officials tend to seek information for criminal justice decisionmaking from only local agencies and persons.

C. Criminal Justice Planning. Each region in the State of Illinois is served by a regional criminal justice planning unit. Although the regional support staffs are not the sole source of information, they are a basic source available to all local government officials. They distribute information on funding sources, cooperative agreements, innovative approaches to crime problems, and so forth.

To determine whether LGOs are familiar with this source of planning information, the officials were asked to identify the name or number of their regional planning commission and the name of the regional planner. Only 27 percent of the respondents correctly identified their region and only 37 percent correctly identified their planner.

The emergence of criminal justice as a field of study and as a focus of increased attention by the Federal government has resulted in the availability of numerous publications related to the administration of justice. Although some of these publications are expensive, some are available without charge. Many contain useful information on planning and evaluating criminal justice services.

With this in mind, the officials were asked what specific publications, including government documents, professional magazines, newsletters, or journals, they now consult in making decisions about criminal justice matters. Over half (56 percent) of the officials reported they did not consult any publication. The publications most commonly mentioned by those who did were:

- (a) The Illinois Municipal Review published by the Illinois Municipal League;
- (b) Target published by the International City Management Association;
- (c) Police Chief published by the International Association of Chiefs of Police; and
- (d) The Illinois County and Township Official, published by the Illinois Association of County Officials, Township Officials of Illinois.

Thus it was found that local government officials are relatively unfamiliar with national criminal justice information sources.

According to the responses, the majority of officials were aware of the National Institute of Law Enforcement and Criminal Justice. The majority of city officials had heard of the National Advisory Commission on Criminal Justice Standards and Goals but the majority of county officials had not. Only one in nine officials knew of the National Criminal Justice Reference Service.

D. Information Needs. Since access to needed information has much to do with sound decisionmaking, the officials were asked two questions related to their perception of the availability of criminal justice information. The first asked, "When you are making criminal justice decisions, do you feel you always, usually, seldom, or never have the information you need to make sound decisions?" Over 60 percent of the officials responded that they "usually" have the information they need. However, only one in seven felt they "always" have the information. City and county officials did not differ on this item. Members of criminal justice committees were slightly more inclined than either chief administrators or members-at-large to feel they need more information.

The second question asked, "In relation to criminal justice matters at the local government level, what would you say are the principal information needs of local government officials?" The item mentioned most frequently was information on local criminal justice operations, i.e., expenditures, caseloads, crimes solved, etc. Other items included information on: crime problems in their locale; special criminal justice topics, e.g., juvenile delinquency; new legislation that affects local operations, e.g., the 1976 Illinois Alcoholism and Intoxication Treatment Act; how the criminal justice system is organized and supposed to operate; funding sources; and how similar communities are handling problems and with what results.

#### IMPLICATIONS FOR CRIMINAL JUSTICE EVALUATION

Local government officials are interested in obtaining more information on criminal justice planning. The results of criminal justice evaluation are one type of information they consider useful.

The fact that many have a full-time occupation apart from serving as a member of local government plus the fact most are laymen in relation to criminal justice, implies that local government officials have neither the time nor the technical knowledge to digest the quantitative evaluations produced by many evaluators. LGOs want information that is concise and useable.

Since local government officials seek information from other LGOs and from local system personnel, criminal justice evaluators might endeavor to get someone in the "circle of intimates" to present the results of evaluation to them.

In summary, local government officials are interested in the results of criminal justice evaluations but evaluators must develop a different strategy for providing the officials with the information if it is to be used.

## BIBLIOGRAPHY

### BOOKS

- Babbie, Earl. Survey Research Methods. Belmont, Calif.: Wadsworth Publishing Co., 1973.
- Cole, George F. The American System of Criminal Justice. North Scituate, Mass.: Duxbury Press, 1975.
- Cole, George F. Criminal Justice: Law and Politics. North Scituate, Mass.: Duxbury Press, 1972.
- Criminological Research and Decision Making. Publication No. 10. Rome: United Nations Social Defense Research Institute, 1974.
- Gorden, Raymond L. Interviewing: strategy, techniques, and tactics. Rev. ed. Homewood, Ill.: The Dorsey Press, 1975.
- Illinois Revised Statutes 1975. State Bar Association Edition. St. Paul, Minn.: West Publishing Co., 1976.
- Kenney, David. Basic Illinois Government. Carbondale: Southern Illinois University Press, 1974.
- Leonard, Edwin C., Jr. Assessment of Training Needs. Fort Wayne, Ind.: Midwest Intergovernmental Training Committee, 1974.
- National Advisory Commission on Criminal Justice Standards and Goals. A National Strategy to Reduce Crime; Criminal Justice System; Police; Courts; Corrections; Community Crime Prevention. Washington, D.C.: U.S. Government Printing Office, 1973.
- Research and Policy Committee of CED. Improving Productivity in State and Local Government. New York: Committee for Economic Development, 1976.
- Survey Research Center, Institute for Social Research. Interviewer's Manual. Rev. ed. Ann Arbor: The University of Michigan, 1976.
- U.S. Department of Justice and U.S. Department of Commerce, Expenditure and Employment Data for the Criminal Justice System 1974. Washington, D.C.: U.S. Government Printing Office, 1976.

### MONOGRAPHS

- Glaser, Daniel. Strategic Criminal Justice Planning. Rockville, Md.: National Institute of Mental Health, Center for Studies of Crime and Delinquency, 1975.
- Gottfredson, Don M., ed. Decision-making in the Criminal Justice System: Reviews and Essays. Rockville, Md.: National Institute of Mental Health, Center for Studies of Crime and Delinquency, 1975.

## HANDBOOKS

Ancel, Louis and Stewart H. Diamond. Illinois Municipal Handbook, 1973-1974 ed.; 1975 Supplement. Springfield, Ill.: Illinois Municipal League.

Howlett, Michael J. Handbook of Illinois Government. Springfield, Ill.: State of Illinois, 1974.

Kelley, Michael J. Police Chief Selection A Handbook for Local Government. Police Foundation and International City Management Association, 1975.

Local Governmental Affairs, Department of, and Northeastern Illinois Planning Commission. Intergovernmental Cooperation in Illinois. State of Illinois, 1976.

Miller, Delbert C. Handbook of Research Design and Social Measurement 2nd ed. New York: David McKay Co., Inc., 1970.

Rolewick, David F. A Short History of the Illinois Judicial Systems. Springfield, Ill.: Administrative Office of the Illinois Courts, 1971.

## VIDEOTAPES

"Interviewing." Princeton, N.J.: Mathematica, Inc.

APPENDIX

PANEL: THE HISTORY AND PROMISE OF CRIMINAL JUSTICE EVALUATION

Criminal Justice Evaluation: A NILECJ Perspective--Richard L. Linster, Office of Evaluation, NILECJ/LEAA

Criminology Evaluates Itself--Marvin E. Wolfgang, Ph.D., Director, Center for Studies in Criminology and Criminal Law, University of Pennsylvania, Philadelphia

The Promise of Program Evaluation in the New Policy Environment--Joseph S. Wholey, The Urban Institute, Washington, D.C.

Future Directions--Marcia Guttentag, Ph.D., School of Education, Harvard University, Cambridge, Massachusetts

PANEL 1: EVALUATION: ITS ROLE, ACTIVITIES, AND RELATIONSHIPS

Federal Program Evaluation and Policy Analysis--James B. Bell and Thomas W. White, The Urban Institute, Washington, D.C.

Evaluation for Decision-Making--Chris Webster, Law and Justice Planning Office, Olympia, Washington

Local Elected Officials Look at Criminal Justice Evaluation--Nancy Loving, U.S. Conference of Mayors, Washington, D.C.

Predetermining the Extent and Sources of Support/Resistance Surrounding Proposed Programs and Policies--William R. Brown, Charles M. Unkovic, James T. Clark, and Katherine D. Hardin, Department of Sociology, Florida Technological University, Orlando

Environmental Effects on Project Implementation--Ralph G. Lewis and Jack R. Green, Michigan State University, Criminal Justice Systems Center, East Lansing

Criminal Justice Information Needs of Illinois Local Government Officials--Robert A. Shay, Project Director, Crime Prevention Commission, Region 14, Illinois Law Enforcement Commission, Jacksonville and Burkett Milner, Sangamon State University, Springfield, Illinois

The Role of the Legislature in Program Evaluation: The Federal Experience in Criminal Justice--Joel Garner, Office of Evaluation, NILECJ/LEAA

PANEL 2: PRETRIAL DIVERSION/INTERVENTION

Voluntary Pretrial Diversion and the Question of Compliance: A Preliminary Evaluation--Diane L. Gottheil, Champaign County Adult Diversion Program, Urbana, Illinois

Assessing the Impact of Treatment and Other Factors on Successful Completion of a Pre-Trial Intervention Program--Jeffrey A. Fagan and Wilfred W. Recker, State University of New York at Buffalo

The Use of Psycho-Socio-Educational Instruments to Identify and Measure Deficiencies within Two Samples of Pre-Trial Intervention Offenders--David M. Weis, with Ronald J. Klein and Isadore Newman, University of Akron, Ohio

Evaluation of Adult Diversion Programs: The California Experience--Michael Agopian, California Department of Corrections, Los Angeles

Pre-Trial Diversion/Intervention Evaluation Report--Jeffrey Zlonis, Governor's Commission on Crime Prevention and Control, St. Paul, Minnesota

Improving the Decision-Making System in Arrest, Charging and Pre-Trial Screening and Diversion--Donald F. Blumberg, Decision Sciences Corp., Jenkintown, and Special Consultant to the District Attorney, Philadelphia, Pennsylvania

PANEL 3: EVALUATING ALTERNATIVES TO INCARCERATION

Intensive Special Probation Programs--Jerry Banks, Georgia Institute of Technology, Atlanta

Alternatives to Incarceration: Humane Corrections or Low Cost Social Control?--David Duffee and Peter B. Meyer, Division of Community Development, The Pennsylvania State University, University Park

The Effectiveness of a State Probation System--George W. Farmer, Probation Development Section, Columbus, Ohio

PORT Projects: Alternatives to Incarceration?--Kay A. Knapp, Minnesota Department of Corrections, St. Paul

An Evaluation of the LEAA Replication of the Des Moines Community-Based Corrections Program--William Rhodes, Florida State University, Tallahassee

Evaluating Community Service Delivery to Offenders--Anthony M. Scillia, Robin Ford, and Kenneth J. Klimusko, Community Corrections Services for the 16th Judicial Circuit, Geneva, Illinois

Shock Probation: A Natural Experiment on the Effect of a Short Period of Incarceration--Joseph A. Waldron, The Ohio Youth Commission, and Henry R. Angelino, The Ohio State University, Columbus

PANEL 4: COMMUNITY RESIDENTIAL TREATMENT FOR JUVENILES

Problems in Juvenile Justice Evaluation Design: Shelter Care Programs--  
Lawrence V. Grant, Governor's Commission on Crime Prevention and  
Control, St. Paul, Minnesota

Alternative Approaches to Juvenile Corrections: A Case Study--Marcia Slot-  
nick, Mayor's Criminal Justice Coordinating Council, New Orleans,  
Louisiana

Group Home Study in Washington State--Jack O'Connell, Law and Justice Plan-  
ning Office, Olympia

A Consumer Satisfaction Format for Evaluating Community-Based Treatment  
Programs for Juvenile Offenders--Robert J. Jones, Saraveen Fields, and  
Gary D. Timber, Western Carolina Center, Morganton, North Carolina

A Consumer and Outcome Evaluation of Community Group Homes for Juvenile Of-  
fenders: A Comparison of the Effectiveness of Trained and Untrained  
Child-Care Staff--Kathryn A. Kirigin, Achievement Place Research Project,  
University of Kansas, Lawrence

An Example of How Cost-Benefit Analysis Can Be Used for Management and Evalu-  
ation of a Community Program for Juvenile Corrections--Robert F. Minne-  
han, College of Urban Affairs and Public Policy, University of Delaware,  
Newark

PANEL 5: EVALUATING DRUG TREATMENT PROGRAMS

A Follow up Study of Former Clients of Drug Treatment Programs in Washington,  
D.C. and New York City--Marvin R. Burt, Institute for Human Resources  
Research, Bethesda, Maryland

Evaluation of an Institution-Based Drug Treatment Program--Cameron R. Dight-  
man, Office of Research, Department of Social and Health Services,  
State of Washington, East Olympia

Three Years Later: A Follow-up of Decertified ODAS Clients--Paul J. Gold-  
stein and Douglas S. Lipton, New York State Office of Drug Abuse Serv-  
ices, New York City

Self Reports and FBI Reports: Further NARA Analyses--Howard L. Kitchener,  
U.S. Bureau of Prisons, and Douglas Stewart, CONSAD Research Corporation,  
Pittsburgh, Pennsylvania

Drug Offender Rehabilitation Program: Recovery Rates, Personality Variables  
and Maintenance Factors--Gregory L. Little and Robert W. Wood, Shelby  
County Penal Farm; Richard S. Sweet, Memphis Correctional Center; and  
Haskel D. Harrison, Correctional Research and Evaluation Center, Memphis,  
Tennessee

Methadone Programs and Crime Reduction: A Comparison of New York and California Addicts--Dale K. Sechrest, Commission on Accreditation for Corrections, Rockville, Maryland

PANEL 6: EVALUATING THE PERFORMANCE OF POLICE

Using Citizen Surveys to Evaluate Policing--Gordon P. Whitaker, University of North Carolina, Chapel Hill

Consumer Evaluation of Police Performance--Helena M. Carlson and Markley S. Sutton, Department of Psychology, Lewis and Clark College, Portland, Oregon

Police Performance Measures--Jerry Needle, American Justice Institute, Sacramento, California

Why Do Arrests Fail in Court?--Brian E. Forst, Institute for Law and Social Research, Washington, D.C.

Personnel Performance Evaluation--Forced Choice--A Municipal Experiment--David M. Hanley, Indianapolis Police Department, Indiana

The Challenge of Patrol Performance Evaluation in the Study of Policewomen--Joyce L. Sichel, Vera Institute of Justice, New York City

PANEL 7: USING THE PHYSICAL ENVIRONMENT TO REDUCE CRIME AND FEAR OF CRIME

Methods for Evaluating Crime Prevention Through Environmental Design Programs--Leonard Bickman, Loyola University of Chicago and Westinghouse Electric Corp., Evanston, Illinois; and Michael D. Maltz, University of Illinois--Chicago Circle and Westinghouse Electric Corp., Evanston, Illinois

Data Requirements in CPTED Evaluations--Richard M. Titus, NILECJ/LEAA/CCP

Evaluation of Programs to Prevent Crime Through Environmental Design--William Brill, William Brill Associates, Annapolis, Maryland

Street Layout and the Occurrence of Residential Burglary--Douglas Frisbie, Governor's Commission on Crime Prevention and Control, St. Paul, Minnesota

The Private Streets of St. Louis--Oscar Newman, Institute for Community Design Analysis, New York City

Is 'Defensible Space' a Defensible Theory? A Field Experiment--James P. Levine, Brooklyn College of the City University of New York; Michael C. Musheno, University of Minnesota, Minneapolis; Denis J. Palumbo, Indiana University, Indianapolis

PANEL 8: USING MODELS AND SIMULATION IN EVALUATIONS, I

A Program Evaluation Model for Analyzing Deterrence and Displacement--  
Stuart T. Deutsch, School of Industrial and Systems Engineering,  
Georgia Institute of Technology, Atlanta

Modeling to Measure Deterrence and Incapacitation Effects--Alfred Blumstein,  
Carnegie Mellon University, Pittsburgh, Pennsylvania

Determinants of the Seriousness of Criminal Activity: The Misdemeanor-  
Felony Distinction--Ann Witte and Peter Schmidt, University of North  
Carolina at Chapel Hill

Applications of Modeling and Simulation in the Assessment of Program Al-  
ternatives to Incarceration--Daniel D. Smith, National Center for  
Juvenile Justice, Pittsburgh, Pennsylvania

Evaluation of a Police Implemented AVM System: The St. Louis Experience--  
Richard C. Larson, Kent W. Colton, and James M. Tien, Public Systems  
Evaluation, Inc., Cambridge, Massachusetts

PANEL 9: MANAGING THE EVALUATION RESOURCE: SOME SPA-RELATED ISSUES

Managing the Evaluation Resource: Some SPA-Related Issues--Peter G. Hart-  
jens, Director of Evaluation, Governor's Commission on Crime Preven-  
tion and Control, St. Paul, Minnesota

The Drip-Drip Flow of Evaluation Data: A Case Study of New York State's  
Evaluation Program--Robert L. Fisher, Senior Research Analyst (Crime  
Control), New York State Division of Criminal Justice Services, New  
York City

Using Grant Evaluation Funds Effectively--William P. Gloege, Administrative  
Analyst III, San Jose Police Department, California

Evaluation for Decision Makers: Using Results--Steve Vojtecky, Utah Council  
on Criminal Justice Administration, Salt Lake City

Evaluation for Decision Makers: The Santa Clara Experience--Leonard Zeitz,  
Director, RCJPB Evaluation Unit, and George M. Shannon, Jr., Director,  
Regional Criminal Justice Planning Board, Santa Clara County, San  
Jose, California

PANEL 10: CITIZEN ACTION AGAINST CRIME--EVALUATIONS

Neighbors Against Crime Together: A Project Evaluation--David I. Sheppard,  
Institute for Research, International Center, Reston, Virginia and  
David Q. Brodie, University of Denver, Colorado

Operation Identification--John D. Carr, Systems Analyst, Denver Anti-Crime  
Council, Colorado

Community Crime Prevention Program, July 1, 1973 to August 31, 1976--Kenneth E. Mathews, Jr., Senior Researcher/Evaluator, Seattle Law and Justice Planning Office, Washington

An Evaluation of the Open Garage Door Burglary Program--Ron Pennington, Research Associate, St. Louis County Police Department, Clayton, Missouri

Eden Prairie Prioritized Premise Survey Evaluation--John Richardson, Research Analyst, Governor's Commission on Crime Prevention and Control, St. Paul, Minnesota

#### PANEL 11: INSTITUTIONAL TREATMENT OF OFFENDERS

An Evaluation of a Prison Treatment Program--Barbara Cadow, Doctoral Student, and Edwin I. Megargee, Professor of Psychology, Florida State University, Tallahassee

Outcomes of Prison Vocational and Educational Programs on Recidivism and Employment Success--Mary Ann Evan, Coordinator of Corrections Education, Oregon State Department of Education, Salem; and Robert Mason, Professor of Sociology, Survey Research Center, Oregon State University, Corvallis, Oregon

Evaluation of a Correctional Education Treatment Program: Variables Affecting Inmate Behavior--H. D. Harrison II, Department Head, and R. W. Wood, Division Head, State Technical Institute at Memphis, Correctional Research and Evaluation Center, Tennessee (co-authors E. B. Blanchard, Professor, Psychology, Tennessee Psychiatric Hospital and Institute, and J. E. Bassett, Director, Self Management Program, Shelby County Penal Farm, Memphis, Tennessee)

The Actualization and Impact of Team Classification in State Correctional Institutions--John R. Hepburn, Department of Sociology, University of Missouri, St. Louis, and Celesta A. Taylor, Director, Criminal Justice Planning, Department of Corrections, Missouri Department of Social Services, Jefferson County

Seventh Step Foundation Program at Kentucky State Reformatory: Evaluation Report--Pat Sims and Mary Ellen Curtin, Department of Justice, Frankfort, Kentucky

A Reassessment of the Trustee System--William L. McWhorter, Assistant Professor of Sociology, Virginia Polytechnic Institute and State University, Blacksburg

Survey of Attitudes and Perceptions of Custodial Staff in Jails: Civilians vs. Sworn Officers--William Rankin and John Easterday, Program Evaluation Section, Wisconsin Council on Criminal Justice, Madison

Violence in Federal Prisons: The Effect of Population Density on Misconduct--Peter L. Nacci, Hugh E. Teitelbaum, and Jerry Prather, Research Office, U.S. Bureau of Prisons, Washington, D.C.

PANEL 12: EVALUATING ALTERNATIVES TO TRADITIONAL POLICE PATROL

On-Scene Arrests in Relationship to Police Response Time: A Preliminary Analysis of Part I Crime Data--William Bieck, Kansas City Police Department, Kansas City, Missouri

Community Service Officer Patrols as a Deterrent to Crime in Housing Projects--George F. Phelan, Planner, Southeastern Massachusetts Criminal Justice Agency, Fall River

Police Evaluation Research: An Analysis of the Significance of a Crime Prevention Technique--John F. Schnelle and Captain Robert E. Kirchner, Jr., Police Department, Metropolitan Government of Nashville and Davidson County, Tennessee

An Evaluation of Specialized Police Patrols--Barbara Sowder, Institute for Human Resources Research, Bethesda, Maryland

Wilmington Split Force Police Patrol--James M. Tien, Public Systems Evaluation, Inc., Cambridge, Massachusetts

PANEL 13: EVALUATION SYNTHESIS

Synthesizing Evaluative Information for the Congress--Joseph Delfico, Assistant Director, Program Analysis Division, General Accounting Office, Washington, D.C.

Evaluation Synthesis--Joe Nay, The Urban Institute, Washington, D.C.

Phase I Results: Washington State Youth Service Bureau Study--Jack O'Connell, Law and Justice Planning Office, Olympia, Washington

Recidivism and Research Design: Limitations of Experimental-Control Research--Robert Martinson, City College of the City, University of New York

The Three Faces of Evaluation: What Can Be Expected to Work--Herbert C. Quay, University of Miami, Coral Gables, Florida

PANEL 14: USING MODELS AND SIMULATION IN EVALUATIONS, II

Models and Policy Evaluation--Gordon Cassidy, Coordinator/Evaluator, and Marcel Laniel, Senior Program Analyst, Ministry of the Solicitor General, Ontario, Canada

Ingredients for Successful Implementation of a Model--Jan Chaiken, The Rand Corporation, Santa Monica, California

Probation Prediction Models and Recidivism--Robin C. Ford, Director, and Shelley R. Johnson, Kane County Diagnostic Center, Geneva, Illinois

Methodology for Comparative Evaluation of the Relative Effectiveness of Various Programs--Carl M. Harris, Syracuse University, New York, and Stephen Stollmack, President, Maximus Management Consultants, McLean, Virginia

Base Expectancy Models: A Non-Parametric Approach and an Evaluation Methodology--David V. Mastran, Maximus Management Consultants, McLean, Virginia

PANEL 15: EVALUATING PAROLE PROGRAMS

Something That Works: Financial Aid for Released Prisoners--Kenneth J. Lenihan, New York City

An Evaluation of the Short-Run Impact of Vocational Counseling and Job-Search Support on the Employment and Law Violation Experience of Released Offenders--Robert W. Gillespie, Department of Economics, University of Illinois at Urbana

Kentucky's Adult Halfway House: Six Month Follow-Up--Mary Ellen Curtin, Patricia L. Sims, and Richard Thompson, Kentucky Department of Justice, Frankfort

Two Studies of Adult Recidivism and Their Implications for the Use of Parole Prediction Devices--Ralph W. Smith, Office of Research, Department of Social and Health Services, State of Washington, Olympia

Parole Impact: A Prototype in Inside/Outside Parole Supervision--Ellen M. Bard, Parole Impact Program, MCI, Concord; and Marian Hyler, Research Unit, Parole Board, Commonwealth of Massachusetts, Boston

Evaluation of a Study Release Center--William Hodge, Department of Sociology; and Karen Morell, Prison Education Program, University of Washington, Seattle

Objective Selection Techniques as a Logical Prerequisite to the Evaluation of Work Release Programs--Douglas E. Scoven, Duane E. Brookhart, and J. B. Ruark, Division of Adult Services, Division of Probation and Parole Services, Virginia Department of Corrections, Richmond

PANEL 16: JUVENILE DIVERSION PROGRAMS

Methodological and Theoretical Issues in Juvenile Diversion: Implications for Evaluations--Dennis Berg, California State University, Fullerton; and David Shichor, California State College, San Bernardino

Recidivism Rates of Diverted Juvenile Offenders--Suzanne B. Lincoln, Kathie S. Teilmann, Malcolm W. Klein, and Susan Labin, University of Southern California

The Diversion of Juvenile Offenders: Initial Success and Replication of an Alternative to the Criminal Justice System--William S. Davidson II, Michigan State University, East Lansing, with Edward Seidman and Julian Rappaport, University of Illinois

Police Diversion in California--Arnold Binder and Ted Palmer, University of California, Irvine, and California Youth Authority

Positive Direction Program (A Police Youth Diversion Project)--Cindy Winslow, Mayor's Criminal Justice Council, San Francisco, California

Community Accountability Program--Kenneth E. Mathews, Jr., and Arlene M. Geist, Seattle Law and Justice Planning Office, Washington

#### PANEL 17: MEASURING THE IMPACT OF LEGISLATIVE CHANGES

Deterrence Regained: The Cheshire Constabulary's Breathalyser Blitz--H. Laurence Ross, Program in Law and Social Sciences, National Science Foundation, Washington, D.C.

Police Roadblocks to Decriminalization: A Multiple-Time-Series Analysis of Law Enforcement's Response to Changes in Public Drunkenness Statutes--David E. Aaronson and Thomas C. Dienes, The American University, Washington, D.C.; Michael C. Musheno, University of Minnesota, Minneapolis

Impact of the Minnesota Community Correction Act on Sentencing Alternatives--K. E. Larimore, Department of Corrections, Minnesota

Evaluating the 1973 New York State Drug Laws--Philip Richardson, Association of the Bar of New York City, Drug Law Evaluation Project

Measurement of Crime Pattern Modification: An Example, Gun Control--Stuart Jay Deutsch, Georgia Institute of Technology, Atlanta

#### PANEL 18: MEASURING PERFORMANCE OF THE COURTS

Indicators of Justice: Measuring the Performance of Prosecution, Defense and Court Agencies Involved in Felony Proceedings--Sorrel Wildhorn, The Rand Corporation, Santa Monica, California

Evaluating the Effectiveness of Defenders--Roberta Rovner-Pieczenik, The Police Foundation, Washington, D.C.

Measuring Prosecutors' Performance--Joan Jacoby, Bureau of Social Science Research, Washington, D.C.

Priority Prosecution Project in Denver--John D. Carr, Denver Anti-Crime Council, Colorado

Expanding the Perspective of Crime Data: Performance Implications for Policymakers--Kristen Williams, Institute for Law and Social Research, Washington, D.C.

Court Information Systems Phase I Evaluation--Burton Kreindel, Justice Systems Group, The MITRE Corporation, Bedford, Massachusetts

PANEL 19: EVALUATION OF EDUCATION AND TRAINING PROGRAMS

The Development and Use of Evaluation Designs in Criminal Justice Training Programs: The Cook County Model--Gad J. Bensinger, Loyola University, formerly with Cook County Criminal Justice Training and Leadership Development Section, Chicago, Illinois

The Analysis and Definition of Critical Training Requirements for Apprentice Sheriffs' Deputies in Texas via Job-Task Analysis and the General Linear Model--Doug Goodgame, Texas A&M University, College Station

Police Effectiveness in Handling Disturbance Calls: An Evaluation of Crisis Intervention Training--Jack B. Pearce, Planning and Research, El Monte Police Department, California; and John R. Snortum, Claremont, California.

An Exploratory Analysis of Personality Differences Among Police Recruits and Its Implication for Training and Performance--Louis A. Quatrano, Association of University Programs in Health Administration, Washington, D.C.

Changing the Attitude of Campus Police Toward Student Activists and Militants: Training in Community Relations and Emotional Control--Khayrallah M. Safar, Florida Agricultural and Mechanical University, Tallahassee

Correctional Staff Development in Virginia: A Tentative Evaluation--Ronald J. Scott, Roseann Evans, and Betty Keck, Virginia Commonwealth University, Richmond

Evaluation of a Program Training Police in Interpersonal Communication Skills--Edwin C. Susskind, University of Maryland, Baltimore; Derek Roemer, Southwestern Community Mental Health Center, Baltimore; Joel M. Lazar, Spring Grove Hospital Center, Baltimore; Oresto R. Digiondomenico and Steven L. Wise, University of Maryland, Baltimore

PANEL 20: EVALUATION DESIGNS FOR COMMUNITY CRIME PREVENTION PROGRAMS

Methodological Problems Facing the Crime Prevention Program Evaluator--John Richardson and Peter Hartjens, Governor's Commission on Crime Prevention and Control, St. Paul, Minnesota

Evaluating Street Lighting Programs--James Tien and Vincent F. O'Donnelly, Public Systems Evaluation, Inc., Boston, Massachusetts

Evaluating Citizen Crime Prevention Efforts--Robert Yin, The Rand Corporation, Washington, D.C.

Evaluation of Security Survey Programs--Charles Girard, International Training, Research and Evaluation Council, Falls Church, Virginia

Methods for Evaluating Citizen Crime Reporting Projects--Paul J. Lavrakas and Leonard Bickman, Loyola University of Chicago and Westinghouse Electric Corporation, Evanston, Illinois

Volunteer/University Teams for Community Program Evaluation--Peter M. Kelly, Kelly Scientific Corporation, Washington, D.C.

#### PANEL 21: EXPERIENCE WITH BUILDING IN-HOUSE EVALUATION CAPABILITIES

An Emergent Model for Training Evaluation Researchers in Criminal Justice Agencies--Margaret K. Snooks and Howard C. Daudistel, Evaluation Research Training Program, The University of Texas at El Paso

Stimulating the Use of Evaluation and Its Results: A University-County Government Linkage Approach--Knowlton Johnson, Institute of Criminal Justice and Criminology, University of Maryland, College Park

The Challenge of Comprehensive Program Evaluation and Planning--Kristann S. Jones, Office of Research and Planning, Division of Correctional Services, Denver, Colorado

Evaluation Management and Organizational Strategies--T. P. Jones, Department of Offender Rehabilitation, Tallahassee, Florida

Management by Evaluation: Putting Research to Work--Dale Shears, Institutional Services Division, Michigan Department of Social Services, Lansing

The One Man Band--Jack C. Stillwell, Victim Witness Advocate Program, Office of the Pima County Attorney; and Dolores Abernethy, Pima County Adult Probation Department, Tucson, Arizona

Management by Objectives and Program Evaluation in the Department of Justice--James F. Hoobler, Management Programs and Budget Staff, Office of Management and Finance, U.S. Department of Justice

#### PANEL 22: EVALUATING PRE-TRIAL RELEASE

Evaluation Techniques for State-Wide Pretrial Release Programs--Burton W. Butler and Stephen F. Wheeler, Pre-Trial Services Agency, Administrative Office of the Courts, Frankfort, Kentucky

Pre-Trial Release in New Orleans--Stuart P. Carroll, Mayor's Criminal Justice Coordinating Council, New Orleans, Louisiana

Evaluating Pre-Trial Release--Dewaine (Nick) Gedney, Pre-Trial Service Division, Common Pleas Court, Philadelphia, Pennsylvania

Bail Reform and Bail Jumping: The Case of Peoria--Paul Lermack, Department of Political Science, Bradley University, Peoria, Illinois

Measuring the Effectiveness of the Bail Bond System as an Assurance of Trial Appearance--Helen Reynolds, Department of Economics, Southern Methodist University, Dallas, Texas

The Pre-Trial Release Process when the Victim and Defendant Live Together--Jack C. Stillwell, David A. Lowenberg, G. Cornevieux, and H. Cunningham, Victim Witness Advocate Program, Tucson, Arizona

Evaluating the Efficiency of Pretrial Services--Robert Wilson, University of Delaware, Newark

#### PANEL 23--COMMUNITY BASED CORRECTIONS FOR ADULTS

Prerelease Residence and Parole Outcome: Panel Study Methods and Pennsylvania Results--David Duffee and Peter B. Meyer, Division of Community Development, The Pennsylvania State University, University Park

Cost Effectiveness of Residential Community Corrections: An Analytical Prototype--Charles M. Gray, Governor's Commission on Crime Prevention and Control, St. Paul, Minnesota and Chris Johnston-Conover, Rand Corporation, Santa Monica, California

The Community-Based Diagnostic and Evaluation Project--Patricia A. Holm and Joe Lehman, Department of Social and Health Services, State of Washington, Seattle

Residential Corrections Programs in Minnesota: An Evaluation Report--Michael J. McMahon, Governor's Commission on Crime Prevention and Control, St. Paul, Minnesota

Impact of Psychosocial Consultation upon Criminal Activity--Robert E. Pelc, Andrew Czopek, and David Shern, Central Intake Unit, Denver General Hospital, Department of Health and Hospitals, Denver, Colorado

Evaluation of the Offender Aid and Restoration Program in Northern Virginia--Robert M. Rich, Rich Associates, Alexandria, Virginia

Factors Related to Outcome of Volunteer Intervention with Criminal Offenders--Michael R. Rosmann and Edward D. Sorel, Department of Psychology, University of Virginia, Charlottesville

#### PANEL 24: ALTERNATIVES FOR TREATING JUVENILE OFFENDERS

Evaluation of Major Youth Correctional Reform in Massachusetts--Lloyd Ohlin, Alden Miller, and Robert Coates, Center for Criminal Justice, Harvard University, Cambridge, Massachusetts

Juvenile Probation Programs in the Impact Cities--Joseph H. Sasfy, The MITRE Corporation, McLean, Virginia

Project New Pride--Stephen F. Browne, Denver Anti-Crime Council, Colorado

Continuous Evaluation of a Youth Services Bureau--Molly G. Schuchat and Leopold O. Walder, Behavior Service Consultants, Inc., Greenbelt, Maryland

Evaluating the Effectiveness of GGI--J. G. Yehl, R. F. Ahlering, P. D. Anderson, I. T. Silvergleit, and J. E. Grush, Northern Illinois University, De Kalb

Correctional Group Counseling: An Evaluation--Joan McCord, Drexel University, Philadelphia

#### PANEL 25: POLICE MANAGEMENT

Rochester Study of Managing Criminal Investigations--James Bell, The Urban Institute, Washington, D.C.

Studying Police Decentralization in Cincinnati--Sumner Clarren, The Urban Institute, Washington, D.C.

Team Policing and Traditional Patrol: Unanswered Questions--Talmage Day, Public Services Laboratory, Georgetown University; and Jane P. Woodward, The Urban Institute, Washington, D.C.

Plain Facts Through Quantitative Analysis--Robert P. Eckert and Peter M. Kelly, Kelly Scientific Corporation, Washington, D.C.

The Criminal Investigation Process: A Summary Report--Peter W. Greenwood, Jan M. Chaiken, and Joan Petersilia, The Rand Corporation, Santa Monica, California

An Evaluation of Civilian Aides in Worcester, Massachusetts--James Tien and Richard C. Larson, Public Systems Evaluation, Inc., Cambridge, Massachusetts

#### PANEL 26: USING MANAGEMENT INFORMATION SYSTEMS IN EVALUATION

Project Arson Detection and Identification--Lawrence J. Funk, Evaluation Unit of the Mayor's Criminal Justice Council, San Francisco, California

The Potential Impact of Information Systems in Evaluating Prosecution and Court Policy--Sidney Brounstein, Institute for Law and Social Research, Washington, D.C.

Use of an On-Line Computer System for Program Evaluation--Laura Winterfield, Judicial Department, Denver, Colorado

Systematic Operational Programs Evaluation--Richard R. Galbraith, Arizona State Department of Corrections, Phoenix

PANEL 27: EVALUATING COMPLEX PROGRAMS: CROSS-SYSTEM APPROACHES

Assessing the Performance of Attorneys as Activists: An Evaluation of the American Bar Association's BASICS Program--C. Ronald Huff, Department of Sociology and Anthropology, Purdue University, West Lafayette, Indiana; Ross F. Conner and Gilbert Geis, Program in Social Ecology, University of California, Irvine

Meta-System Evaluation: Capacity Building--Stuart Vexler, Community Assistance Programs, Texas Youth Council, Austin

Tracking Offenders Through the Juvenile Justice System--Linda Devin-Sheehan, Youth Bureau, County of Suffolk, Riverhead, New York

Saint Louis High Impact Crime Displacement Study--William Stenzel, Institute for Public Program Analysis, St. Louis, Missouri

High Impact Anti-Crime in New Orleans: The Target Area Experience--Robert Sternhell and Stuart Carroll, Mayor's Criminal Justice Coordinating Council, New Orleans, Louisiana

Criminal Justice System Capability Building: A System-Wide Approach in Local Government--Marcus D. Ingle, Judith Wilks, and Dale I. Parsons, County of Onondaga, New York, Criminal Justice Planning and Coordination Unit

Atlanta's High Impact Anti-Crime Program--Samit Roy, Crime Analysis Team, Atlanta, Georgia

PANEL 28: JUVENILE DELINQUENCY PREVENTION

The Environmental Roles of Delinquency Prevention Programs--Burton J. Cohen, Management and Behavioral Science Center, The Wharton School of the University of Pennsylvania, Philadelphia

Family Therapy with Status Offender Families at a Child Welfare Agency--Robert Green and Kenneth W. Michaels, Children's Service of York County, Pennsylvania

A Longitudinal Evaluation of a Police School Liaison Program--Rodney Mulder, Urban Studies Institute and Donald Williams, Grand Valley State College, Allendale, Michigan

Helping Delinquent Dropouts to Cope: An Evaluation of an Innovative Project--Lee H. Bowker, Department of Sociology, Whitman College, Walla Walla, Washington

Difficulties in Evaluating Community Sited Delinquency Prevention Programs--The BPlay Project--John S. Bis, Optometric Center of Maryland, Baltimore

Evaluation Methodology in a Large-Scale Program for High School Students: Delinquency, Drug Abuse and Attitude Change--William F. Soskin and Robert L. Fisher, Project Community, University of California, Berkeley

PANEL 29: THE MODEL EVALUATION PROGRAM: DEVELOPING EVALUATION CAPABILITIES

Evaluation and Decision-Makers: Perceptions of the Evaluation Process--Anne L. Schneider and Peter R. Schneider, Institute of Policy Analysis, Eugene, Oregon; and Robert Henderson, Law and Justice Planning Office, Office of Community Development, Olympia, Washington

The Development of a Local Criminal Justice Evaluation Capability--Tom Long, State of Florida Bureau of Criminal Justice Planning and Assistance, Tallahassee, Florida

Community Based Research to Improve Methods of Evaluation--Larry E. Goodman, Darrel H. Tiller, and Henry P. Clark, Association of Central Oklahoma Governments, Oklahoma City

The Alameda Model Evaluation Program--Bruce Kern and John Lenser, Alameda Regional Criminal Justice Planning Board, Oakland, California

New Hampshire Model Evaluation Program: An Overview--Joann C. Vaughn and Virginia E. Garrell, Model Evaluation Program, Governor's Commission on Crime and Delinquency, Concord

Lessons Learned from the Model Evaluation Effort in Illinois--Richard F. Sullivan, Illinois Law Enforcement Commission, Chicago

The Michigan Model Evaluation Program--Marilyn Hall and John Snyder, Office of Criminal Justice Programs, Lansing, Michigan

PANEL 30: EVALUATING INNOVATIVE ALTERNATIVES TO ADJUDICATION

An Overview of Alternatives to Conventional Adjudication--David J. Saari, Center for the Administration of Justice, The American University, Washington, D.C.

The Problems of Comparative Evaluation--Earl Johnson, Jr., University of Southern California and Social Science Research Institute, U.S.C., Los Angeles, California

Alternatives to Incarceration: The Importance of Judicial Performance to Diversion and Referral Programs: The Case of TASC--Robert Sternhell, Criminal Justice Coordinating Council, New Orleans, Louisiana

An Alternative to Court: An Evaluation of the Orange County (Florida) Bar Association Citizen Dispute Settlement Project--Ross F. Conner, University of California, Irvine; and Ray Surette, Florida State University, Tallahassee

Victim-Defendant Relationships in an Adult Diversion Program--Jack C. Stillwell, Victim Witness Advocate Program, Office of the Pima County Attorney, Tucson, Arizona

Some Preliminary Results of the Impact Assessment of the Community Arbitration Program--Merry A. Morash, Community Arbitration Program, Annapolis, Maryland

PANEL 31: EVALUATING THE PUBLIC DEFENSE FUNCTION

Legal Delivery Systems to Prisoners: A Preliminary Evaluation--Goeffrey P. Alpert, University of Texas at Dallas

Statistical Failure and Subjective Success: An Evaluation of Criminal Defense Services--Richard Block, Loyola University of Chicago; Carolyn R. Block, Illinois Law Enforcement Commission; and LaMarr Billups, Loyola University Law School

Alternative Methods for Evaluating Indigent Defense Systems--Howard F. Feinman and Anne L. Schneider, Institute of Policy Analysis, Eugene, Oregon; and Robert Henderson, Law and Justice Planning Office, Office of Community Development, Olympia, Washington

Criteria for Measurement of Defender Office Effectiveness--Gustav Goldberger, National Center for Defense Management, Washington, D.C.

Evaluation of the Legal Aid Society Defender Division, Cincinnati, Ohio--Shelvin Singer, National Legal Aid and Defender Association, Chicago, Illinois

Developmental History as a Method of Organizational Evaluation--Jackwell Susman, Institute for Advanced Studies in Justice, The American University, Washington, D.C.

PANEL 32: EVALUATION PARADIGMS: OTHER APPROACHES

Experimental Design and Causal Analysis in Criminal Justice Evaluation--John G. Heilman and Marshall T. Miller, Auburn ALEPA Criminal Justice Evaluation Project, Auburn University, Alabama

Meeting the Demand for Repeating Successful Criminal Justice Projects by Using Economics--Thomas J. Nagy, Center for the Study of Law and Society, and Robert L. Fisher, Project Community, University of California, Berkeley

Ethnographic Approaches to Evaluation: Qualitative Evaluation Research Design for Criminal Justice--George W. Noblit, Memphis State University, Tennessee

The Need for Hybrid Evaluation Models in Multi-Module Criminal Justice Projects--Subhash R. Sonnad and Paul C. Friday, Western Michigan University, Kalamazoo

A Framework Model of Evaluation--Charles W. N. Thompson, Northwestern University, Evanston, Illinois

Evaluating a Community Based Diversion Project: A General Systems Strategy--Eduard A. Ziegenhagen, Center for Social Analysis, State University of New York at Binghamton

Conducting Evaluative Research and Implementing Its Results: A Dilemma for Both the Administrator and Researcher--Denis Szabo, University of Montreal, International Centre for Comparative Criminology, Canada

PANEL 33: IMPROVING EVALUATION USE

Utilization of Evaluation by Federal Agencies--Eleanor Chelimsky, The MITRE Corporation, McLean, Virginia

Evaluation, Feedback and Policy--Lawrence A. Bennett, Center for the Study of Crime, Delinquency and Corrections, Southern Illinois University at Carbondale

Framing the Questions in Criminal Justice Evaluation: Maybe You Can Get There from Here If You Ask the Right Questions--Phyllis Jo Baunach, Office of Research Programs, Corrections Division, NILECJ, Washington, D.C.

What Works in Criminal Justice? Some Uses of Program Evaluation--Peter C. Buffum, Pennsylvania Prison Society, Philadelphia

The Impact of Evaluation on Program Development: An Analysis of the Monitoring Function of the Minnesota Ombudsman for Corrections--Randall K. Halvorson, Ombudsman for Corrections, St. Paul

The Conceptualization of Program Evaluation and the Identification of Pitfalls in Some Current Evaluation Efforts--John K. Hudzik, Criminal Justice Systems Center and Associate Professor, School of Criminal Justice, Michigan State University, East Lansing; and Larry K. Gaines, College of Law Enforcement, Eastern Kentucky University, Richmond

Evaluation Analysis for the Reconceptualization of Urban Law Enforcement--Jeanie K. Meyer, Operations Resource Unit, Kansas City Police Department, Kansas City, Missouri

Classification of Social Services for Evaluation: A Dilemma and a Solution--Peter B. Meyer, Division of Community Development, The Pennsylvania State University, University Park

PANEL 34: WITNESS ASSISTANCE AND COURT MANAGEMENT

Introductory Remarks--Marc A. Nerenstone, Office of Regional Operations, LEAA, U.S. Department of Justice, Washington, D.C.

An Evaluation of the New York Victim/Witness Assistance Project's Court-Based Services--Jeremy Travis and Robert Davis, Vera Institute of Justice, Brooklyn, New York

Milwaukee County Project Turnaround--Richard Knudten, Evaluation Policy Research Associates, Ltd.; and James Jensen, Price Waterhouse and Company, Milwaukee, Wisconsin.

Overview of Witness Assistance Programs--William F. McDonald, Institute of Criminal Law and Procedure, Georgetown University, Washington, D.C.