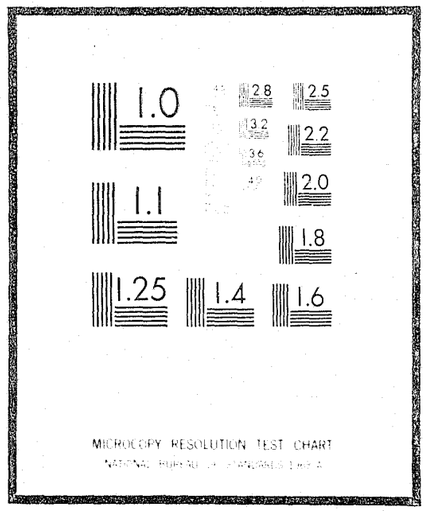

EVALUATIVE RESEARCH IN CORRECTIONS: A PRACTICAL GUIDE

By
STUART ADAMS, Ph.D.

This project was supported by Grant Number 73–TA–1001–G, awarded to The American University Law Institute, Washington, D.C. by the National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Department of Justice, under the Omnibus Crime Control and Safe Streets Act of 1968, as amended. 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.

March 1975
U.S. DEPARTMENT OF JUSTICE
Law Enforcement Assistance Administration
National Institute of Law Enforcement and Criminal Justice

LAW ENFORCEMENT ASSISTANCE ADMINISTRATION

Richard W. Velde, *Administrator*

Charles R. Work, *Deputy Administrator*

NATIONAL INSTITUTE OF LAW ENFORCEMENT AND CRIMINAL JUSTICE

Gerald M. Caplan, *Director*

Library of Congress Cataloging in Publication Data

Adams, Stuart Nicholas, 1909–
Evaluative research in corrections—a practical guide.
(Prescriptive package—National Institute of Law Enforcement and Criminal Justice)
Bibliography: p.
Includes index.
1. Corrections—Research. I. National Institute of Law Enforcement and Criminal Justice. II. American University, Washington, D.C. Law Institute. III. Title. IV. Series: National Institute of Law Enforcement and Criminal Justice. Prescriptive package—National Institute of Law Enforcement and Criminal Justice. [DNLM: 1. Crime. 2. Rbilitation. 3. Research. HV6024.5 A196e]

HV9274.A3 364.6'07'2 74–28348

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402 — Price $2.00
Stock Number 2700—00270

FOREWORD

In recent years the commonly accepted belief that research can be a productive administrative tool in almost any activity has been challenged within the field of corrections by a variety of social analysts. They assert that evaluative studies have provided little evidence that any correctional programs "work."

In *The Practical Guide to Evaluative Research in Corrections,* Dr. Stuart Adams examines the bases and validity of the assertion and then defines and describe techniques for using research to make corrections more effective. The resulting guidelines are designed to make research a more effective instrument in the management and improvement of the correctional subsystem of criminal justice.

The main emphases of the *Guide* are two: the correctional administrator has several responsibilities to fulfill if he is to benefit from research, and the researcher must command a variety of techniques if he is to meet the descriptive and analytical needs of his agency. The administrator must not only know the capability and requirements of objective research; he must also provide support, define his research needs, and make use of the products of his research staff. The researcher, in addition to comprehending the strengths and weaknesses of the many evaluation models, must also understand the strategies that aid in translating information into action in operating agencies.

The *Guide* covers both traditional and contemporary approaches to evaluation—from surveys and controlled experiments to operations research and simulation. Within this range are techniques that any agency—regardless of size or research sophistication—should find useful.

GERALD M. CAPLAN
Director

GOT A MOMENT?

We'd like to know what you think of this Prescriptive Package.

The last page of this publication is a questionnaire.

Will you take a few moments to complete it? The postage is prepaid.

Your answers will help us provide you with more useful Prescriptive Packages.

PREFACE

The guidelines set forth in the volume are general and exploratory. Since they are addressed to both administrators and researchers, they can deal only briefly with the concerns of either. And since corrections is a field in rapid transition, the directions presented here are in a sense initial sightings on a moving target.

For administrators, the sections on management, support, and utilization of research presume at least a passing familiarity with the results of research. For researchers, some training in methodology and statistics and some acquaintance with applications is presumed, although an effort has been made to keep the presentation concrete and elementary. The volume is seen as an early effort in a field where more exhaustive documentations of research procedures will soon follow.

The substantive materials and general concepts in this volume have come largely from organizations and persons that the author has worked for or with over the past twenty years. The organizations include the California Special Study Commission on Correctional Facilities and Services, the California Bureau of Criminal Statistics, the California Youth Authority, the Los Angeles County Probation Department, the School of Criminology of U. C. at Berkeley, San Quentin Prison, and the District of Columbia Department of Corrections.

Persons whose influences are reflected in various ways in the following pages include Ronald H. Beattie, Kenneth L. Hardy, Karl Holton, Joseph D. Lohman, Austin H. MacCormick, Richard A. McGee, and Heman G. Stark. As sources of operational or evaluative concepts, these individuals have played important roles in the search for greater rationality in corrections and criminal justice.

For reviews and comments on this volume, the author is indebted to Harold B. Bradley, James J. Franczyk, Don M. Gottfredson, Harland H. Hill, Leon Leiberg, Neal Miller, M. Robert Montilla, and Marguerite Q. Warren. Special acknowledgements are due Walter R. Burkhart, Daniel Glaser, and John J. Henning for detailed critiques of the manuscript.

STUART ADAMS

CONTENTS

PART ONE. STATUS AND IMPACT OF EVALUATIVE RESEARCH

PART TWO. ROLE OF THE AGENCY ADMINISTRATOR

PART THREE. RESEARCH METHODS AND STRATEGIES

PART ONE.
STATUS AND IMPACT
OF
EVALUATIVE RESEARCH

CHAPTER 1. INTRODUCTION

The purpose of this volume is to give practical direction to evaluative research in corrections. The basic strategy followed here is to identify concepts and procedures that "work," in the sense of producing operationally useful results, and to emphasize the use of such concepts and procedures in the planning and implementation of further research.

A "practical" approach to evaluative research does not deny the relevance or importance of two other well-known approaches to research: the methodological, and the theoretical. In actuality, the practical approach builds upon either or both of the others. But by bringing in additional information, particularly of kinds ordinarily considered unimportant for either method or theory, it sometimes makes discoveries or achieves impacts that would not have occurred under the other approaches.

The proposed approach has thus far seen little use in corrections, primarily because correctional research is a relatively new discipline, and its achievements have as yet been too limited to suggest new alternatives. Now that results have started to accumulate, alternative approaches become possible. We can begin to formulate different types of evaluation strategy by focusing on results as well as on method and theory. This should enable us to move beyond the "managerial efficiency" stage, where accepted method and accepted theory provide a limited basis for strategy, to the "effectiveness" stage, where combinations of method, theory and results provide sounder bases for strategy.

Since correctional research is a new discipline, and since the practical approach as defined here is even newer, it might be presumed that much of the present activity within this approach is exploratory. This is in fact the case. We are concerned in this volume not so much with constructing precise models for solving specific evaluation problems as with developing something more general. The result, in brief, is not a "cookbook," nor even a "handbook," but a "guide."

The aptness of the term "guide" becames evident when we recall the general status of evaluative research. We note a great deal of confusion over objectives, criteria and methods; disagreement over whether evaluative research shows programs to be "efficacious" or not; and, in the latter case, controversy over whether the inefficacy should be attributed to the research or to the correctional programs.

These contradictions and confusions suggest that our primary need in correctional evaluation at present is not precise formulas for the application of research to particular situations but rather general guidelines. We also need larger quantities of operationally relevant research results. How to attain these objectives is the question.

This volume sets itself four tasks, each taken up in a separate part of the volume. *Part One* is concerned primarily with a review of the present state of evaluation in corrections and with identifying the kinds of research that have had an impact—i.e., have "made a difference." *Part Two* examines the responsibilities of the agency administrator for the direction, support and utilization of evaluative research. *Part Three* takes up the methods, strategies and models of the evaluator, with emphasis on the changing techniques and goals of research. *Part Four* explores some possibilities for "evaluating evaluation" and some proposals for making evaluation more productive in the future.

These four areas of inquiry, along with their rationales, are discussed in more detail later in this chapter. As a preliminary to that discussion, it will be useful to examine one of these rationales: the rising pressures for evaluation in corrections.

A. Pressure for Evaluation: A Historical Note

The demand for evaluative research in corrections has surged in recent years. A generation ago, studies of the effectiveness of the correctional process were low-key activities, centered in several universities and a few correctional agencies that had directors with inquiring turns of mind. In the past twenty years, the scene has changed markedly. Rising crime rates, crowded prisons, fear for public safety, budgetary problems and the spread of new management methods into social agencies have had

their effect. Strong pressures for evaluation and growing commitments to evaluative study are now evident everywhere.

A notable early instance of this pressure occurred in California in 1957. The Legislative Auditor directed the state's Department of Corrections to request special funding for the purpose of systematizing and accelerating its ongoing research activities.[1] The Auditor observed that corrections was becoming ever more costly, yet there was apparently little impact on the behavior of ex-prisoners, judging by trends in recidivism rates. This was a condition that might be remedied by a more effective use of evaluative research.

The Auditor's recommendation led to the creation of research divisions in both the California Department of Corrections and the California Youth Authority. These divisions took over various evaluative studies that were already underway, shifting the research responsibilities from knowledgeable administrative or treatment staff to full-time professional researchers.

At about the same time, and possibly for similar reasons, departments of correction in several other states also established units for carrying out research. By 1966, there were 19 such units reported among 48 states that were surveyed on this point.[2]

In the mid-1960s, correctional self-assessment received a new impetus, this time from the planning-programming-budgeting movement (PPB). PPB made its appearance in the Department of Defense, spread quickly to other federal agencies, then moved out to state and local agencies.[3] It served not only as a stimulus to self-evaluation but also as a source of new measurement procedures, particularly cost-effectiveness and systems analysis.[4] By 1968, discussions of cost-benefit analysis began to be heard at the major correctional conferences.[5]

The late 1960s and the early 1970s saw still more pressure for evaluation in corrections, particularly from the field of criminal justice itself. In 1965 the President established the Office of Law Enforcement Assistance (OLEA) under the Attorney General, and some of the assistance grants made by OLEA were for the evaluation of correctional programs.

The basis for a greater impact was laid in 1968. The Omnibus Crime Control and Safe Streets Act set up the Law Enforcement Assistance Administration (LEAA), successor to OLEA, also under the Attorney General. LEAA was given responsibility for disbursing funds to criminal justice planning agencies in every state for worthy action programs in corrections, the courts and police departments.

It soon became evident that good management of these grants required some attention to assessment. In many instances, intermittent monitoring was perhaps sufficient; in others, rigorous evaluation of outcomes to establish both effectiveness and efficiency seemed desirable. The State Planning Agencies (SPAs), which served as channels for moving LEAA funds into local action projects, were charged with seeing that the project designs included statements about objectives, plans for measuring attainment of those objectives, and allocations of funds for accomplishing the measurement. The guidelines for evaluation that began evolving in the early years of LEAA's existence were formalized in the federal Crime Control Act of 1973.

We now see that an evaluation task of major proportions has been set for corrections by state legislatures, by rising professional standards in corrections itself, by LEAA and by public expectation. State Correctional Agency (SCA) administrators and SPA directors are being asked for continuing assessments of kinds and in volumes that are unprecedented.

The demands for program and system evaluation have clearly run ahead of the research capabilities of corrections and also ahead of the state of the evaluative art. This means that corrections is now under pressure to engage in massive and sustained evaluation, but the necessary research staff, analytical procedures, organizational support and study guidelines and models are either lacking or ineffectively mobilized. As a result, corrections faces, if not an evaluation crisis, at least some severe and lasting evaluation problems.

B. Two Key Elements in Evaluation

Of the several elements of the evaluation problem just noted, the present volume addresses itself primarily to two:

- The organizational support that is required to facilitate evaluation and to utilize its results;
- The research procedures and guidelines that are needed for successful evaluation.

In format and content, the volume is directed at two major audiences: administrators of correctional systems and criminal justice planning agencies on the one hand, and correctional researchers or evaluators on the other. These two audiences play critical, interrelated roles in the production and use of evaluative research in corrections.

1. *The Administrator.* The correctional administrator or criminal justice planner decides, knowingly or not, how extensively he will support and direct research in his organization. He also decides how seriously he will use research as a management device. He provides a structure and creates a climate that can either encourage or discourage the production, reporting and use of study findings.

Hitherto, if the administrator headed an SCA or a local correctional agency (LCA), he may have been the sole determinant of whether his agency "bothered" with research at all. In the future, particularly if he heads an SPA, he will almost certainly be required to maintain an evaluation unit. However, the accomplishments of this unit, and its ultimate impact on his agency and on the criminal justice system generally, will be under his control.

There is at present no curriculum in "Research Organization, Production and Utilization" for administrators of SCAs, SPAs and LCAs. Their skills in this area must come primarily from self-education: directed or casual reading, interchange with researchers, and observations of the production and use of research in other agencies and fields.

To cope with the evaluative task now confronting him, the administrator requires something more direct and substantial to supplement this haphazard road to learning: An administrator's guide to evaluative research is not the sole answer to the problem, but it appears to be an essential part of the answer.

2. *The Evaluator.* The evaluator may be better prepared than the administrator to assume his particular responsibilities for ascertaining the effectiveness of corrections. He has had, presumably, formal training in research, and perhaps some background in research planning and execution. He also has at hand a body of literature on research concepts, procedures and findings. Furthermore, dozens of treatises on evaluative research are in existence, all produced in recent years, setting forth principles and methods for measuring the effectiveness of social programs.[6]

Most of these treatises or guides are relatively traditional in content, and for various reasons they are of limited utility to the criminal justice evaluator. They over-emphasize methodological or design considerations and pay insufficient attention to observed impact of research methods. They place undue stress on the role and potential of the controlled experiment. They provide almost no material on examples of applications and results in the field of criminal justice. And, finally, they underemphasize or omit reference to contemporary methods of measurement, from cost-benefit analysis to simulation.

The evaluator who has recently come out of graduate school will undoubtedly have absorbed many of the viewpoints that are developed in these treatises. The same may be true of the evaluator who has been out of school for some time but has been keeping up with the literature of evaluation methodology. In either case, practical guidelines derived from the special needs and characteristics of correctional research will be just as important to the evaluator as to the administrator.

C. Organization of the Volume

Reference has already been made to the fact that this volume is divided into four major parts. The subject-matter of the four parts are outlined in more detail below:

1. *Part One: Status and Impact of Evaluative Research.* This section, following some introductory comments, focuses on two areas: the present status of evaluative research in correction, and the impact of selected cases on correctional structure and function. The status of research is examined through reviews of evaluation published over the past several years by a number of observers, working independently from various frames of reference. The examination of projects with impact—studies or projects that "changed things"—is an attempt to understand better the characteristics of research that "works," and to ascertain to what extent environing factors may have had a part to play in the impact achieved by the research.

2. *Part Two: Role of the Agency Administrator.* This part deals with the special responsibilities, skills and attitudes that heads of correctional agencies have or need if research is to be facilitated and its products are to be used.

One may wonder at the heavy emphasis on the administrator's role in a guide to evaluative research. The reason may be stated simply: A large part of the deficiency in correctional evaluation at the present time stems from administrator noninvolvement. Administrators who have research units in their agencies tend typically to delegate to researchers

[1] See end of Chapter for Notes.

some of the responsibilities that properly belong to the director or to one of his chief deputies. The result is that the potential of research in agency improvement remains unrealized. Administrators who lack research units have a different problem. They tend to neglect the possibility of obtaining evaluations through various funding sources or of "borrowing" useful research findings as they disseminate through the various channels that carry such materials. This, too, is a loss of potential.

In the administrator's area of concern, in addition to extent of interest in evaluation as a management aid, we place such matters as research needs and priorities; source of research products—i.e., in-house or external; choosing evaluators and funding the research operation; creating organizational climates that support research; developing "experimental" stances on the part of administrative staff; and making use of research findings.

3. *Part Three: Research Methods and Strategies.* Part Three, the longest of the four sections, deals with basic research concepts and procedures—objectives, the problem of criteria, methods of measurement, and models. It also deals with research methods, old and new.

"Old" methods refers to the familiar categories of "non-experimental" (or "pre-experimental," as some methodologists prefer), quasi-experimental and experimental procedures. New methods include cost analysis, operations analysis, systems analysis and simulation. For both old and new methods, there are definitions, examples of application and results, and an examination of implications for near-term and long-term research planning.

To the extent that new research directions or emphases are present in this volume, they are perhaps three in number. First, there is more attention to the newer or more contemporary methods of measurement—from cost analysis to simulation. Second, there is the suggestion that the utility of research may be more important than its form. Researchers in corrections hitherto have been highly conscious of textbook methods and procedures, somewhat unaware of what methods "work" in the operational setting of corrections. Research designs that produce an impact are not necessarily "strong" designs. In time this awareness should sharpen, and both evaluators and administrators may take their cues more from what works than from what is "proper."

Third, there is the implication that researchers should become more active in program conceptualizing and planning. They should, in other words, move part way into development as well as into long-range planning. This combination of the scientist and social engineer is not a new concept, but the practical approach to evaluative research suggests that there is more merit to the concept than has been recognized.

4. *Part Four: Looking Ahead.* The fourth part of the volume has three main foci. One is the discussion of methods by which evaluative research in corrections may be improved. Possible methods range from the increased use of theory as a guide in evaluation to the development of a consortium of SCAs and SPAs to engage in cooperative studies of crucial correctional problems.

A second focus is the possibility of making systematic assessments of the evaluation process and its products for the purpose of achieving progressive improvement of evaluation. One example of a means of evaluating evaluative research is presented.

The final point of focus is future directions in evaluation: Where do we go from here? Several possibilities are presented for consideration.

NOTES

1. Office of the Legislative Auditor, *Analysis of the Budget Bill of the State of California, Fiscal Year July 1, 1957 to June 30, 1958.* Sacramento: California Assembly, 1958, p. 56.
2. Robert H. Fosen and Jay Campbell, Jr., "Common sense and correctional science," *Journal of Research in Crime and Delinquency,* 3 (July 1966), p. 75.
3. Robert L. Chartrand, *Systems Technology Applied to Social and Community Problems.* New York: Spartan Books, 1971.
4. Charles J. Hitch, *Decision-Making for Defense.* Berkeley: University of California Press, 1965.
5. Stuart Adams, "Is corrections ready for cost-benefit analysis?" Paper presented at the 98th Congress of Corrections, San Francisco, August 1968.
6. Some of the better known of these general volumes on the methodological and theoretical aspects of the measurement of social programs are Edward A. Suchman, *Evaluative Research,* New York: Russell Sage Foundation, 1967; Carol H. Weiss, *Evaluation Research,* Allyn and Bacon, 1972; and Joseph Wholey and others, *Federal Evaluation Policy,* Washington, D.C.: The Urban Institute, 1970.

CHAPTER 2. THE STATUS OF EVALUATION IN CORRECTIONS

One useful base for determining in which direction administrators and evaluators should move is a review of the state of the art. Where does correctional evaluation stand at present? In particular, what are the results of evaluation?

For a score of years, evaluative research has been carried on more or less steadily in corrections. The bulk of the effort has been that of research divisions in SCAs and LCAs, with additional contributions by university faculties, private consulting organizations, research institutes, and foundations. Numerous reports and papers have been produced, detailing methods and findings in many areas of corrections. At times the volume of material has grown so large that special efforts have been made to assemble, organize and disseminate it. Both public and private centers have been involved and continue to be involved in these latter tasks.

It would be very useful for correctional administrators if systematic reviews of this material were available, showing kinds of evaluative effort, areas of concentration, research methods used, and actions resulting from the research. if any. Information of this kind would be valuable in guiding both the planning of research and its utilization.

Although a comprehensive review of the total literature has not yet been attempted, several efforts to pull together and assess portions of the literature have been made. We have had, thus far, several small-scale "evaluations of evaluation." Since these undoubtedly contain some important lessons, the best known will be summarized and some implications drawn.

A. Nine Reviews of Evaluation

Nine reviews will be examined. Two are monographs, four are published articles, two are chapters or sections of books, and one is an unpublished paper. The bulk of the reviews are of reports that appeared in professional or academic journals or were released as agency reports or papers. One of the reviews, a monograph, deals with final reports by grant recipients on projects they carried out under LEAA grants.

1. *One Hundred Correctional Outcomes (1966).* Bailey examined 100 articles or reports of outcomes of correctional projects or programs.[1] The studies were published or released, in most instances, between 1940 and 1960. The sample was a broad one, though probably not representative of all "evaluative" studies since Bailey selected for a) a base in empirical data, and b) a manipulated independent variable.

The studies were evaluated on research design, relative use of group as opposed to individual treatment, authoritative or nonauthoritative setting, researcher qualifications, and use of causal theory or models.

Bailey observed that there were some positive elements in the group of studies, including evidence of an increasing concern for demonstrating positive outcomes and also a progressive improvement in the quality of research over time. He noted that reports of successful outcome were common in the studies. For example, 9 of 22 experimental designs in the sample claimed positive effects. However, he regarded the reports of success as unduly optimistic in view of the quality of the data, the types of analyses, and the quality of the interpretations. His final judgment was that "evidence supporting the efficacy of correctional treatment is slight, inconsistent, and of questionable reliability."

2. *Correctional Caseload Experiments (1967).* Adams reported on a review of 22 experimental studies of probation and parole caseloads in California.[2] The experiments had been conducted over the preceding ten or twelve years to learn whether increased effectiveness resulted when caseloads were reduced in size or were differentiated by client personality or service-need types. Reports were drawn from the research divisions of three California correctional agencies, state and local, and from the School of Criminology, University of California at Berkeley. The studies included all known reports of experimental projects carried on by these agencies within the subject-matter field.

In the review, a determination was made as to whether the experimental cases showed statistically significant improvement in behavior (avoided parole revocation or arrest) or whether there were lower correctional costs for the experimentals than for the controls during the first year of follow-up.

Of the 22 projects, 13 showed either behavioral improvement or monetary benefits for the experimental cases. Some of the projects, primarily those with small numbers, showed monetary benefits without reaching statistical significance in behavioral improvement. Positive gains were relatively more frequent in projects dealing with youth than in those dealing with adults.

3. *Effectiveness of Correctional Programs (1971).* Robison and Smith examined several studies that related to five decision points in California corrections: a) sentencing to probation or prison, b) length of term or stay in prison, c) kind of treatment in prison, d) intensity of supervision on probation or parole, and e) whether to discharge from prison or release on parole.[3] The studies, approximately ten in number, were predominately controlled experiments.

The authors concluded that differences in recidivism rates between experimentals and controls were often difficult to interpret because measurement procedures were confounded by the reporting methods of the correctional systems. Such variations in rates as could be satisfactorily established were believed to result primarily from initial differences among the types of offenders that were processed. The remaining variation was judged to be the "result of differences in the way agency staff interpreted or defined violative behavior of offenders."

This review of the major California correctional programs that have been rigorously evaluated, the authors observed, suggests that ". . . there is no evidence to support any program's claim to superior rehabilitative efficacy."

4. *A Survey of 231 Rigorous Treatment Evaluations (1971).* Martinson, over a period of several years, assembled an extensive collection of published and unpublished reports on correctional treatment programs.[4] His monograph provides a critical summary of "all studies published since 1945" that assess rigorously the effects of any kind of treatment applied to convicted offenders.

The survey found 231 studies that met the methodological and empirical criteria for inclusion in the analysis. The author's main conclusion was that ". . . there is very little evidence in these studies that any prevailing mode of correctional treatment has a decisive effect in reducing the recidivism of convicted offenders."

5. *Implications of Negative Research Findings (1971).* Kassebaum, Ward and Wilner, at the conclusion of an experimental assessment of group counseling in a California prison, reported book-length on their findings.[5] Included in the treatise was an assessment of evaluative research in corrections generally. They noted that there has been a lack of evaluative research in corrections, and that there were few published studies on correctional programs reported either by agency research divisions or by independent investigators.

One noteworthy feature of the evaluative research that exists, the author observed, is that it brings generally disappointing findings. These negative findings have been appearing in growing numbers in the last few years, as correctional administrators try to respond to requests for evidence on the effectiveness of their programs. The authors see this trend as boding ill for the recent push toward use of evaluation as an aid to management. Corrections is possibly going to be less hospitable toward outside researchers and less willing to disclose its research findings, since, "as the reports come in, there is a dearth of good tidings for both the treatment specialists and the program administrators."

6. *Effectiveness of Psychotherapy With Offenders (1972).* Speer, seeking to develop a crisis intervention model for the rehabilitation of criminal offenders, reviewed the literature on psychotherapy in corrections.[6] He found 26 references to 21 controlled experimental studies which examined the effects of group and individual psychotherapy on offenders. Seven of the 21 studies involved adult offenders and 14 involved juveniles. Fifteen of the studies were institution-based; 6 were community-based.

The review disclosed that in only 11 of the 21 studies were follow-ups conducted and recidivism data collected. Six of the 11 were institution- and 5 were community-based.

Of the 6 institutional studies, 4 produced positive effects and 2 null effects. Among the 5 in-community studies, 2 showed positive and 3 showed null findings.

The only "moderately definitive conclusion" suggested by these results, Speer asserted, related to the age of the treated offenders. Of the 3 studies with adults, none showed significant reduction in recidivism through treatment. Of the 8 studies involving juveniles, 6 resulted in significantly less recidivism after treatment.

Overall, slightly more than 50 percent of these experimental studies showed significant reductions in recidivism associated with treatment.

7. *Experimental Treatments with Institutionalized Offenders (1972).* Shireman and others reported on 12 experimental studies of treatment of juvenile and adult offenders in state and local institutions.[7] The treatments included milieu therapy, group counseling and plastic surgery. Follow-ups into the community showed that in five of the 12 studies, the experimental subjects performed better than the controls. In two studies, there was no difference in performance but the treatment cost for the experimentals was lower because of shortened institutional time. In the remaining five studies, there was no advantage of any kind for the experimental groups.

In one of the experiments, the significantly better parole performance at 12 months began to diminish and at 24 months the recidivism rates of experimentals and controls were about the same. The authors concluded that ". . . certain forms of institutional treatment may have impact sufficiently powerful to influence postrelease performance positively."

8. *A Summary of SPA Project Evaluations (1973).* Berkowitz reviewed and evaluated 38 reports on projects undertaken in California in 1971–72 with LEAA support.[8] The projects were among 400 carried out during this period under the auspices of the California Council on Criminal Justice, the planning agency for the State's criminal justice system. The 38 studies were judged to be generally representative of the 400 projects, which employed a wide range of evaluative methods. Five of the 38 studies were controlled experimental designs.

The reviewer identified 154 measurable objectives that had been specified in the 38 projects. About 40 percent of these objectives were judged to have been accomplished, taking the substance of the final reports as evidence of accomplishment. The reviewer also identified 73 methodological deficiencies in the projects, ranging from lack of a control group to use of an unvalidated testing instrument. Goal attainment was highest and deficiency rate lowest in the five experimental projects.

9. *Evaluation for Corrections (1973).* Ward examined the general field of evaluative research in California corrections over the past 15 or 20 years.[9] The research covered the areas of group conuseling, parole, halfway houses, community treatment and psychotherapy. He noted that there were repeated failures to demonstrate conclusively that positive gains came from any of these varieties of "treatment." Under pressure to produce results after occasional premature claims of effectiveness, administrators often found themselves in embarrassing positions, which led occasionally to excuses, circumlocution and other doubtful defenses.

To improve correctional evaluation, Ward proposed targeting research funds more selectively, improving the coordination of projects undertaken in different locations, and relying more heavily upon the faculties of state universities for evaluative tasks. Asserting that continual assessment is essential to effective and efficient operation of agencies, he concluded by stating that "Evaluative research is one of the few ways of keeping the corrections business honest."

B. Inferences from the Nine Reviews

These nine evaluations of evaluative studies present a number of contradictions, some ambiguities and occasional agreements. A few of the more conspicuous of these results are worthy of comment, since they give us some of our first insights into how experienced observers sum up the status of evaluative research in corrections. These insights are necessarily tentative, and some of them might be modified if a more systematic review of the total body of evaluative material in corrections were to be undertaken.

1. *Objectivity.* Some of the reviews were polemical or ideological in tone; others were factual and objective. The former emphasized the "inefficacy" of either correctional programs or correctional research, saw some lack of candor on the part of administrators or researchers in reporting findings, feared a possible ban on outside researchers, and were concerned about keeping corrections "honest." The latter specified the ground rules of their reviews, and reported, primarily in quantitative terms, the results of their surveys or assessments.

2. *Criteria of Inclusion.* Some reviewers considered studies that others omitted or failed to discover. Speer included the PICO project in his review of experiments in psychotherapy.[10] Robison and Smith did not include PICO, despite the fact that it was a very successful demonstration of the efficacy of treatment in a California prison. Cases such as this disclose that the reviewers differed by the criteria by which they selected projects for review as well as the standards by which they judged a study for negative or positive findings.

3. *Scarcity of Experimental Studies.* There was

agreement that relatively few evaluative studies were experimental in design. The review that sampled most representatively found about 10 percent of the designs to be controlled experiments. However, the reviewer noted that this type of design was over-represented in the sample, and we may estimate a lower value—perhaps five percent—for the proportion of experiments in that sample.

4. *Treatment More Efficacious with Juveniles than with Adults.* Three of the reviewers examined adult and juvenile treatment programs separately and found that programs for juvenile tended to have higher likelihood of positive results. This finding is of interest for a number of reasons, including the suggestion that a comprehensive assessment of evaluations would be useful as a means of identifying structural or functional areas in which research might be the most productive.

5. *Few Positive Findings.* Some reviewers reported that no findings were convincingly positive. Others reported that as many as 50 percent of the controlled experimental studies in their surveys indicated positive gains from treatment. Some of the high estimates of percentages positive would probably drop appreciably if the studies were screened rigorously for selection bias or other conditions that might impair validity. However, even if half of all the controlled experimental studies "paid off" in terms of improved community performance, this would still mean that only a small percentage of all evaluative studies were productive.

C. Comparison with Evaluation in Other fields

Given the apparent lack of rigor in correctional evaluation and the apparently low percentage of positive findings, it should be useful to make some comparisons with research efforts in other fields. Comparison is difficult because few fields report systematically on the results of their evaluative research, or their research and development activities. Some of the high-technology industries are an exception.

Lessing, in 1950, quoted a former president of du Pont as estimating that not more than one in twenty of du Pont's research projects eventually paid off.[11] Boehm, in 1957, cited a Booz, Allen & Hamilton survey of 120 companies which reported that two-thirds of all research projects fail to produce useful results. Failure rates ranged from 50 percent in some companies to 99 percent in others.[12] Kay, in 1965, quoted a vice president of research, development and basic engineering for Monsanto as saying, "No one can know the precise figure . . . but I'm sure that more than half the money this country is spending for R & D is wasted."[13]

A more recent statement comes from Blood, who estimates that ". . . an average of four out of five engineers and scientists work on projects that do not reach commercial success."[14] Still more recently, White, speaking of the field of medical research, states that during three decades of intensive biomedical research, there has been no improvement in life expectancy of adults, and no discovery of ". . . effective means . . . for coping with the stubborn complex of social illnesses that now predominate in the economically advanced countries."[15]

These are rough comparisons, and they need to be interpreted with care. Nevertheless, they suggest that evaluative research in corrections may be making a commendable showing, given the difficulties of the context in which it must work and the newness of the greater part of the evaluation effort.

This showing is all the more impressive when one compares evaluation in corrections with that in the courts, police activities, welfare, employment or education. The quality of research design, the complexity of research activity, and the prevalence of significant findings appears somewhat higher in corrections than in these other fields.

If these are accurate perceptions, they may reflect not so much the brillance and dedication of correctional administrators and researchers as the fact that corrections has been under considerable pressure, both self-imposed and external, to evaluate and improve its operations. These pressures are one of the penalties of being, in a sense, the Cinderella (without a magic slipper) among social agencies.

Since it is desirable that perceptions be realistic as well as accurate, the fact remains that evaluative research in corrections has not yet resolved its greatest problems; it has just begun to face up to them. To make progress, it must decide where it needs to go and what techniques and strategies will get it there.

REFERENCES AND NOTES

1. Walter C. Bailey, "Correctional outcome: an evaluation of 100 reports," *Journal of Criminal Law, Criminology and Police Science*, 57 (June 1966), 153–60, p. 158.
2. Stuart Adams, "Some findings from correctional caseload research," *Federal Probation*. 31 (December 1967), pp. 48–57.
3. James Robison and Gerald Smith, "The effectiveness of correctional programs," *Crime and Delinquency*, 17 (January 1971), 67–80, p. 80.
4. Robert M. Martinson, *Treatment Evaluation Survey*, New York: Division of Criminal Justice Services, 1971, unpublished study. Cited in Gene Kassebaum et al, 1971, p. 311.
5. Gene Kassebaum, David A. Ward and Daniel M. Wilner, *Prison Treatment and Parole Survival*, New York: John Wiley, 1971.
6. David C. Speer, "The role of the crisis intervention model in the rehabilitation of criminal offenders," Buffalo: Erie County Suicide Prevention and Crisis Service, 1972, unpublished paper.
7. Charles H. Shireman, Katharine Baird Mann, Charles Larsen and Thomas Young, "Findings from experiments in treatment in the correctional institution," *Social Service Review*, (March 1972), 38–59, p. 54.
8. Francine Berkowitz, *Evaluation of Crime Control Programs in California: A Review*. Sacramento: California Council on Criminal Justice, April 1973.
9. David A. Ward, "Evaluative research for corrections," in Lloyd E. Ohlin (ed.), *Prisoners in America*, Englewood Cliffs, N.J.: Prentice Hall , 1973, 184–203, p. 202.
10. The report on the PICO Project was not included in any of the "pessimistic" reviews of evaluative research thus far published. When asked why he did not include this highly efficacious project among those he evaluated, one reviewer stated that it fell outside his designated time interval. For information on this project, see Chapter 12.
11. Lawrence P. Lessing, "The world of du Pont: how to win at research," *Fortune*, 42 (October 1950), 115–34, p. 115.
12. G. A. W. Boehm, "Research management: the new executive job," *Fortune*, 49 (October 1957), p. 164.
13. Hubert Kay, "Harnessing the R & D. monster," *Fortune*, 71 (January 1965), p. 160.
14. Jerome W. Blood (ed.), *Utilizing R & D By-Products*, New York: American Management Association, 1967, p. 16.
15. Kerr White, "Life, and death and medicine," *Scientific American*, 229 (September 1973), 23–33, p. 23.

CHAPTER 3. RESEARCH WITH IMPACT: SIX CASE STUDIES

In the preceding chapter, reviewers of evaluative studies were quoted as saying that programs were found to be ineffective, or that research showed "null effect," or that there were "significant reductions in recidivism." There were few if any references to the "impact" of research.

It is essential to explore the concept of impact and to illustrate it. A project or study that has impact brings about some change in the situation. It modifies the structure of processes of correction, or it alters the behaviors of offenders in the correctional system. Impacts may vary considerably in magnitude, ranging from dramatic and extensive to barely perceptible alterations. Impacts may also have theoretical and methodological as well as practical consequences. Our concern here is with impacts of operationally significant kinds, particularly those that might be described as major in magnitude.

Since it is part of the function of administrators and researchers to achieve impact—to change their offenders or to improve their systems—the question arises: How are these changes planned and brought about? What kinds of projects have impacts? What are the organizational and community conditions associated with impact?

As in the case of the *status* of evaluative research, there has been little or no exploration of the *impact* of correctional research. There has been awareness that impact was a goal of research, and a well-established conviction that the "best" research was likely to achieve the greatest impact. This conviction was sometimes stated as a judgment that "strong" research designs, such as controlled experiments, should be emphasized in program evaluations.

These kinds of issues have not been studied in corrections, although they are obviously important for correctional organization and procedure. It would be helpful to administrators, planners and evaluators if the scores or even hundreds of effective projects in the history of correctional evaluation had been followed up to learn what the impacts had been—if any—and also the factors that were related to impact.

Since there has been no systematic study of impacts, we now proceed to do as was done for status of evaluative research—examine the matter of research impact through a number of case studies. These are not reviews of impact, since there appear to be none thus far, but examinations of the studies or projects themselves.

The number of studies that can be covered in the space available here is clearly limited. The next several pages present six case studies of projects or studies that led to important changes in their host agencies or in the clientele they served. Since the extent to which these projects represent the universe of useful correctional evaluations is not known, our ability to generalize will be limited. Nevertheless, there should be much of value to be learned from these presentations.

A. Six Case Studies of Evaluative Research Impact

Of the six studies to be presented, one was a state correctional commission study, four were studies by research staff in state departments of correction, and one was a study by a non-profit organization funded by the U.S. Department of Labor. The six studies that were chosen met two criteria: a) the availability of full information about the study and the impacting process, and b) a present or prospective impact of considerable magnitude.

1. *Probation Subsidy in California.* In 1956, the California Special Study Commission on Correctional Facilities and Services undertook a study of the 60 county probation departments in the state. The purpose of the study was to ascertain the status of the probation function, to note problems and deficiencies in that function, and to make such administrative and legislative recommendations as the evaluation indicated were desirable.

The study was based on a field survey design. Structured interview schedules were devised to obtain descriptive and judgmental information from probation department representatives. Other schedules were employed in interviews with county judges, whom the probation departments served. Some workload and organizational data were gathered by mailed questionnaires. Data on the performance of probationers while under supervision were obtained from one of the larger counties that maintained a punchcard record system containing both background data and performance information.

The survey report was primarily a summary description of the structure and functioning of the 60 county probation departments, with additional material on attitudes and judgments of probation staff and judges.[1] Research staff added judgmental evaluations of probation operations, using published and consensual standards in the field of probation as bases for comparison.

Of the several administrative and legislative recommendations in the survey report, the major item was that the counties should be given a subsidy by the state, so that they could bring their operations in line with professional standards.

The legislative recommendation failed of passage in the Assembly in 1957. A few years later the survey was repeated, on a smaller scale, and the recommendation was resubmitted, this time with a variable formula that tied payments to the county to rates of diversion of convicted offenders from state institutions to county correctional programs.[2] This proposal was approved by the Assembly.

The initial effect of the program was to slow down the rate of commitments to state institutions. One state official estimated in 1973 that during the first eight years since enactment of the Program, about 40,000 adult and juvenile offenders had been diverted from California correctional institutions.

2. *The Preston Impact Study.* In 1959, a research analyst in the California Youth Authority's division of research raised questions as to the "impact" of CYA's Preston School of Industry (PSI) on its male deliquent wards. The inquiries developed into an exploratory study that ended with evaluative judgments about the PSI operation.

The study design was a panel interview procedure. A series of tape-recorded interviews was held with 15 youths who were sent by the juvenile courts to the Northern Reception Center Clinic and who were later assigned to Preston. Interviews occurred at two-week intervals during the six-week clinic stay and at two-month intervals at PSI.

The purpose of the panel interview data was to trace the effects of the clinic and training school experience on the youths. Interest centered on possible shifts in delinquency identification, attitudes toward authority, peer relationships, values, goals, and criminal language and skills. After extended reviews of the transcribed interview records, the researcher formed qualitative judgments about the developmental trends of the group. He concluded that the net effect was probably unfavorable to the youths and also to the communities to which they would return.[3]

The principal outcome of the study was a recommendation to the Youth Authority executive staff that a controlled experiment be conducted to compare the effects of institutional and community treatments—the latter to be carried out in small caseloads by special professional staff.[4] It was hypothesized that community treatment would be both more effective and more economical than institutional treatment. Such an experiment by the Youth Authority would be an invasion of the county probation function. If successful, however, it could be turned over to the counties for implementation on a permanent basis.

The Youth Authority executive staff made a decision to carry out the proposed experiment. An 18-month period of reformulating the original proposal and trying to get it accepted by the state legislature, funding agencies, and county officials ensued. This series of events, but especially the decision by the Youth Authority staff to test its basic program against an alternative that might eventually make the Youth Authority obsolete, illustrates a different kind of impact than Case No. 1. In the present instance, the impact was primarily the generation of a decision. In the normal course of events, however, that decision might be expected to lead to further major consequences.

3. *California Youth Authority's Community Treatment Project.* The Community Treatment Project (CTP) provided the experimental test called for by the Preston Impact Study. During 1961 to 1974, CYA used an interpersonal maturity (I-Level) typology, differential treatment, ward-staff matching on personality and behavioral characteristics, and a complicated experimental design to study the comparative outcomes of community and institutional treatment.[5] The project was organized into three phases, 1961–64, 1964–69, and 1969–74, each focusing on one or more of several intervention and knowledge objectives.[6] The research component of the project was funded by the National Institute of Mental Health.

The impact of CTP on the correctional community appears to have been appreciable, even though difficult to specify precisely.[7] Since the first release

of findings, there has been wide dissemination of the results through project reports, by institutes and seminars, and through the operation of a differential treatment training center. The latter was set up, with assistance from NIMH, to teach the techniques of the project to persons from interested agencies and disciplines.

In California, CTP as a formal structure appears likely to disappear from the Youth Authority—as was anticipated in the first proposals. The concept of community treatment is now being heavily implemented in the California counties where subsidy funds make it readily possible. It is worthy of emphasis that both CTP and the Probation Subsidy Program have reduced the institutional population of Youth Authority, but the most dramatic effect appears to have come from the probation subsidy.

4. *Parole Work Unit Program, 1965–1974.* In 1964 the California Assembly approved the Parole Work Unit Project, which added 76 new parole agents to the existing force of about 80 agents in the California Department of Corrections. This permitted the assignment of about half the Department's 10,000 parolees to Work Unit caseloads, which averaged 35 cases per agent and provided intense, regular or minimal supervision as needed. The other parolees continued in regular 75 to 80 person caseloads.

During the first six months of the project, returns to prison, particularly for technical parole violations, were higher among work unit parolees than among regular parolees. In terms of original expectations, the experiment was a failure. However, supervisory and operational staff reconceptualized the objectives and conditions of the project, emphasizing avoidance of unnecessary returns to the system and more effective use of community resources in retaining parolees in the community.

The balance of performance soon shifted, and the work unit parolees began showing lower rates of return for both new felony offenses and technical violations. Over the next several years the reduction of rates of return continued; it also spread to the conventional parole units through imitation, competition, or other influences.[8]

In March 1974, the Department reported to the Assembly that returns to prison had dropped 52.7 percent since 1965, keeping an estimated 5,827 parolees in the community and saving the Department about $23,000,000 in operating costs.[9]

5. *Pretrial Diversion: Project Crossroads, 1968–71.* This project was planned and implemented by a non-profit organization, aided by federal funds, to explore the possibility of diverting young first-offenders from adjudication by providing counseling, job-finding, educational placement, and other services during a 90-day period after arrest and before trial. Successful participation led to dropping of adjudication.[10]

The project made use of a quasi-experimental design to evaluate client performance and a cost-benefit analysis to ascertain the economic consequences of the new procedures. It was learned that recidivism rates for participants during the 90-day project period were reduced significantly below those of a comparison group chosen from the court files to match the participants. Participants also showed relative up-grading in job status and earnings. Both the recidivism and earnings gains persisted during a twelve-month follow-up period.

Economically, the project showed benefit/cost ratios of about 2 to 1, based on estimates that projected recidivism and earnings benefits about five years into the future.[11] The project has now become part of the operations of the D.C. Superior Court, where it handles about three percent of the court intake. Along with the Manhattan Court Employment Project, Project Crossroads has served as a model for many pretrial diversion projects across the nation. It has also prompted the drafting of Congressional bills (S. 798, Burdick; H. 9007, Railsback) to provide a legislative base for the pretrial diversion programs.

6. *Community-Based Management of Narcotic Offenders.* In early 1969, the associate director for planning and research in the D.C. Department of Corrections became curious as to the trend of intake of narcotic-involved offenders into the Department. A time-series study of the dates of commitment of offenders presently in the institutions was made. The study showed an exponential trend in intakes—i.e., a rapid upturn in recent months from a gradually rising base.[12]

To check on possible spuriousness, the trend was reexamined by use of admissions data from the past 15 years. These data also showed a rapidly rising trend in recent admissions. The final report of the study included a recommendation for prompt development of community-based treatment in lieu of prison for offenders of this type. The alternative appeared to be a rapidly rising prison population and pressures for more construction—with poor prognosis for rehabilitation.

The result, within six months, was two halfway houses for narcotic-involved offenders. Within twelve months, these were expanded into a District-wide Narcotics Treatment Administration—now apparently necessary for what looked like a full-blown heroin epidemic in the district's high-risk population. Within two years the program grew into (proportionately) the largest methodone maintenance program in the nation.

B. Characteristics of Impacting Evaluations

It may be premature to seek conclusions from only six evaluative studies that carried marked impacts, but it is important to begin somewhere. What can be said of these studies?

1. *Varied Methodologies.* Several different research methodologies are represented in these six case studies: survey research, in-depth interviews leading to qualitative findings, a time series study, a quasi-experiment, and a complex controlled experiment based on random assignment to control and experimental statuses. Impact does not appear to be the prerogative of any type or types of research method.

2. *Heavy Impact from "Weak" Designs.* The heaviest impact came from the crudest design—a field survey; the lightest impact may have come from the strongest design—the community Treatment Project (excluding the Preston Impact Study, which was a decision-posing rather than a people-changing design). This indicates, tentatively, that strength of design and magnitude of impact were inversely related or perhaps unrelated in these studies.

3. *Emphasis on System Change.* The six studies tended to emphasize system-changing rather than offender changing. Probation subsidies led to a shift in the correctional balance, toward the community and away from state institutions. The Preston Impact Study proposed quite directly the replacement of institutional treatment by community treatment. Both offender change and system change were implied in the design of the Community Treatment Project. The pretrial diversion project focused on structural modification of the arrest-adjudicate sequence. And the narcotic-involved offender study resulted in a substitution of non-correctional management for correctional management of a certain category of potential prisoners.

This finding has one or more major implications. It may imply that in the present state of the art, it is much easier to change correctional systems than to change offenders. On the other hand, it may indicate that there is greater interest in and support for system change. There are other possible interpretations, but these will have to await more extensive analysis of impact data.

4. *Relation to Planning and Development.* Studies with impact tend to show efforts in developing appropriate recommendations and follow-through plans, usually on the part of the researchers who performed the evaluation. This suggests that impact may occur more readily from the work of researcher-planners than from the work of researcher specialists.

For some projects, the explanation of important impacts may lie in the subsequent developmental work rather than in the initial evaluation. And in some cases impact may come as a complete surprise, not at all predictable from the results of the evaluation. The Parole Work Unit Project, for example, at first showed no advantage in small parole caseloads, even when organized according to the intensity of client needs. An important impact was achieved only when staff made improvisations in operations to counter the unexepected failure of the project to achieve hoped-for outcomes.

5. *Inter-Agency Involvement.* In the six cases, considerable inter-agency or inter-group communication was involved. This may have been necessary for these particular impacts to materialize, since most entailed system change. Interaction with or the cooperation of other agencies is usually an essential ingredient in such change.

6. *Research sponsorship.* The six cases of impacting research were sponsored by public agencies, a public commission, or a non-profit private agency. It is not clear what meaning this has for the effectiveness of two other possible sources of correctional evaluation: the faculties of universities, or for-profit consulting and research organizations.

7. *Exploratory Research.* Two of the six impacting studies were exploratory, arising out of research staff member curiosity rather than as a planned item in a departmental agenda of research. The recommendations and proposals followed directly out of the exploration.

C. Discussion

Considering the fact that most evaluative studies in corrections are non-experimental (as Bailey and Berkowitz have shown in Chapter 2), and also the fact that level of impact correlated negatively with strength of design in the preceding cases, it would appear that much of the current emphasis on experimental studies and strong designs may need review and re-evaluation. At the same time, we need

to inquire further into the relation between impact and research characteristics.

If it is eventually established that only a small percentage of studies should be experimental, and that most should be of a variety of other designs, the present mix of evaluation designs may be appropriate to the present state of corrections. It is conceivable, of course, that other conclusions could be reached. The current phase of corrections may be transitional, and loose study designs may be the most productive in a transitional phase, whereas in a later period of stability, experimental designs may prove generally more productive.

For the present, it appears reasonable to conclude that surveys, case studies, and time series analyses may be capable of greater impact on corrections than controlled experiments. This follows partly from the nature of the decision-making process in corrections and partly from the kinds of problems that are amenable to attack by the various research methods. There are obviously types of questions that can best be answered by experiments, but they may not at this time be the most important questions. It would appear that in the next decade or two, at least, evaluative research in corrections may call more for flexibility and resourcefulness than for rigor and certainty.

REFERENCES AND NOTES

1. Stuart Adams and Milton Burdman, *Probation in California.* Sacramento: Special Study Commission on Correctional Facilities and Services, 1957.
2. Robert L. Smith, *The Board of Corrections Probation Study.* Sacramento: California Board of Corrections. See also, Robert L. Smith, *The Quiet Revolution.* U. S. Department of Health, Education and Welfare, DHEW Publication Number (SRS) 73–26011, 1973.
3. Stuart Adams, *The Preston Impact Study.* Sacramento: California Youth Authority, 1959. Unpublished.
4. ————, *A Proposal for the Comparison of Institutional and Community Treatment for Juveniles.* Sacramento: California Youth Authority, 1959. Unpublished.
5. Marguerite Q. Warren, "The community treatment project," in Norman Johnston, Leonard Savitz and Marvin E. Wolfgang (eds.), *The Sociology of Punishment and Correction.* New York: John Wiley & Sons, 1970 (2nd ed.), pp. 671–83.
6. Ted Palmer, *The Community Treatment Project in Perspective: 1961–1973.* Sacramento: California Youth Authority, 1973.
7. An impact study will be desirable at some time in the future to trace the influence of the Community Treatment Project upon correctional thought and action. For the present, the significance of the project may have been obscured somewhat by discussions of the relevance of some differences in decision-making procedures between supervisors of experimental and control cases. See Paul Lerman, "Evaluating the Outcomes of Institutions for delinquents," *Social Work,* Vol. 13, July 1968, pp. 55–64.
8. Walter R. Burkhart, "The parole work unit programme: an evaluation report," *British Journal of Criminology,* April 1969, pp. 125–47. See also Lawrence A. Bennett, "Should we change the offender or the system?" *Crime and Delinquency,* Vol. 19, July 1973, No. 3, pp. 332–42.
9. For the latest statement on the Parole Work Unit Program see *The 1973 Report on the Work Unit Parole Program,* Sacramento: California Department of Corrections, March 1974.
10. Leon Leiberg, Roberta Rovner-Pieczenik and John F. Holahan, *Project Crossroads,* 3 vols., District of Columbia: National Committee for Children and Youth, 1970–71.
11. A later review of the cost-benefit analysis by the National Pretrial Intervention Services Center indicated tentatively that the actual benefit/cost ratio was closer to 1/1. See Neal Miller, *Evaluation of Research on Pretrial Diversion,* District of Columbia: NPISC, American Bar Association, 1974, p. 41. Unpublished.
12. Stuart Adams, Charles Reynolds and Dewey F. Meadows, *Narcotic-Involved Offenders in the Department of Corrections.* District of Columbia: Department of Corrections, Research Report No. 12, 1969.

PART TWO.
ROLE OF THE AGENCY ADMINISTRATOR

CHAPTER 4. SIGNIFICANCE OF THE ADMINISTRATOR'S ROLE

The correctional agency administrator is a key figure in the evaluation process. Whether or not research is an effective instrument in the management of the agency may depend more on the attitudes, perspectives and intellectual positions of the administrator than on those of the chief of research.

The validity of this proposition has to a large extent already been established by studies of the role of the administrator in relation to computerized information systems, operations research units, and planning-programming-budgeting activities. In each of these instances, the extent to which the activity is an asset or a liability to the agency or firm depends largely on how the chief executive uses or relates to the activity.

Most firms that have adopted electronic data processing systems, for example, have found that their original expectations of payoff have not been met. A study of factors associated with payoff indicates that installations in which organization and direction is "expert dominated" are less productive from the agency point of view than those in which organization and direction are integrated with and supportive of executive decision-making.[1] The latter occurs primarily when top executives seriously involve themselves in the planning, organization, and utilization of the system. Similar conclusions have been reported for the manner in which agency directors or top executives relate to their operations research units or planning-programming-budgeting activities.[2]

Not all correctional agencies have in-house research units—the latest survey of such units shows them to be present in about twenty of the fifty state correctional agencies.[3] And in many states with such units, the activity is of such limited scope that the foregoing observations are of only academic significance. Even in such cases, however, and also in cases where there is no research unit whatsoever, there is reason for the administrator's concern. He still has to consider what to do about research findings that originate in other systems, and whether he will introduce innovative programs into his system, and to what extent he will contract for evaluation by outside researchers.

A. The Administrative Dilemma in Evaluation

Part of the administrator's concern about evaluative research comes from the dilemma that research creates for him. The evaluation process casts him in contradictory roles. On the one hand, he is the key person in the agency, and the success of its various operations, including evaluation, depends on his knowledge and involvement. On the other hand, evaluation carries the potentiality of discrediting an administratively sponsored program or of undermining a position the administrator has taken. The question then arises: To what extent can the administrator give support to this potent management instrument if it has the power of calling into question his decisions and his beliefs?

Some administrators have shown they can give support only to a limited extent. Research reports that are unfavorable or research plans that appear to invade politically sensitive areas sometimes run into administrative censorship. The reports are filed away and the research plans are dropped. In time the research unit may become a statistical reports unit, isolated from the decision-making function.

If evaluative research is in fact an important part of good correctional management, how is this potential to be realized, particularly in relation to the role of the agency administrator?

We can start by postulating that the administrator is indeed the key to whether research will be useful in making a department of corrections more effective. We can then inquire as to how this works out in practice. What is it that the administrator does that makes evaluative research pay off? Does he have to do more than hire a competent chief of research and delegate to him responsibility for evaluation?

B. Evaluation-Relevant Roles of Administrators

There are clearly several matters that need attention beyond the hiring of an evaluator or a chief of research. First, there is need for the administrator to develop an appropriate stance with respect to research—to become an "experimental" rather than a "trapped" administrator, as Campbell has put it. There is need to foster a departmental climate in which research staff can operate at high effectiveness. There is need to define research requirements and establish research priorities that relate closely to constructive policy. There is need for continuing interaction with researchers, exchanging ideas, discussing new correctional concepts and emerging research findings, and learning how to penetrate the communication barriers that separate researchers from other agency staff.

There is need for the administrator to react appropriately to research results, to consider the best uses of positive findings, and to weigh fully the implications of negative findings. There is need to avoid over-reaction to either positive or negative findings, and at the same time to make suitable uses of both types of findings. There is need to involve researchers in decision-making and in planning. There should be active research input into the planners' activities, or direct involvement of researchers in the planning process if there is no independent planning unit.

More detailed comment on some of these needs follows in the paragraphs below and in the remaining chapters in Part Two.

C. Trapped and Experimental Administrators

Campbell has sketched briefly two recognizable administrative styles in relation to the handling of evaluative research.[4] The "trapped" administrator is a committed man. He believes in the efficacy or the rightness or perhaps the inevitability of his program, and if an evaluation indicates it lacks effectiveness, the administrator is in difficulty, emotionally and perhaps also officially. He may be inclined to question the validity of the evaluation, or he may simply shelve the research without comment. This creates a problem for his research staff, and it interferes with the orderly development of programs in his agency.

The "experimental" administrator sees himself as committed not to a particular program but merely to the concept of program or agency improvement. If his old or new program is found to be ineffective, he is ready to plan another program that might be more successful. He will be disappointed if evaluation shows a lack of effectiveness, but he will not be disorganized by the findings.

The experimental administrator is pragmatic, forward-looking, and more interested in finding solutions to problems than in justifying a particular choice of a solution. How do agencies secure administrators with experimental turns of mind?

One possible route is to seek administrators among younger members of staff who have had exposure to research methods and products, in school, in other agencies, or in the same agency. Given the growing importance of evaluation in human agencies, ability to make effective use of evaluation will undoubtedly become an essential characteristic in the agency administrator.

Another route is that of on-the-job training at the administrator level. Continuing interaction between administrative and research staff is, presumably an instructive process for the administrator as well as the researcher. For this process to be operative there will need to be a research unit in the agency, and the researchers will need to be capable of meaningful communication with the administrator at the required levels.

Still another route is that of institutes or seminars for administrators. Data processing equipment firms frequently conduct week-long seminars for administrators, seeking to impart basic concepts of management information system use. These events are to some extent self-serving, since they may stimulate demand for a proprietary product. However, they also enable management to deal with the problem of under-comprehension and under-utilization that is evident in the majority of EDP installations.

There are as yet few comparable seminars in research for correctional administrators, but there are some precedents,[5] and these could grow in importance as the need becomes recognized.

There is an adage sometimes heard when the problem of implementing evaluative research in correctional agencies is discussed: "You can't evaluate your own boss."[6] This statement has credibility if the boss is a "trapped" administrator. It is not necessarily true of experimental administrators. Since there are many reasons why it is desirable to organize evaluative research on an in-house basis, it is advantageous to train administrators to view evaluative research from the experimental rather than from the trapped or committed perspective.

Glaser cites an example of a correctional agency that has made noteworthy progress in evaluating its programs because it operates under the experimental rather than the committed strategy.

> "The California Youth Authority has been able to maintain evaluation research without as much interruption as has occurred in other . . . organizations, I believe, largely because it has persistently started its innovations as limited trials. Generally, these were controlled experiments, (initiated) at only one or two of its facilities or districts.
>
> "Such trials have assessed, much more rigorously than is customary in corrections, psychiatric teams for youth institutions, diagnosis and prognosis by a personality inventory, intensive small unit treatment for younger wards, intensive short-term institutional treatment for older wards, intensive differential treatment in the community, special narcotic addict counseling and nalline testing, intensive reading instruction, job placement programs, small parole caseloads, community-involving parole centers, and other variations in treatment practice. Research staff were involved with operations personnel in planning the trials so that their consequences could be measured. It is also noteworthy that these projects were not launched with great fanfare, and they were called experiments or trials from the start, so that there was no promise of demonstrable effectiveness."[7]

These remarks are encouraging in that they demonstrate the ability of a correctional agency to set up an in-house research and evaluation unit, well-staffed, and operating under experimental department directors, which carried out a vast quantity of rigorous research—in effect, "evaluating its own boss." By ordinary standards, the Youth Authority research division has set enviable records in the amounts and quality of research accomplished.

These remarks also have their doubtful aspects, since they give added weight to experimental design at a time when there may be reason to re-evaluate the role of this research design in corrections. This doubt may carry over as well as to the fact that so much of the CYA research focused on offender change in institutional settings when the primary need of the day may have been system change, at both the institutional and community levels.

D. An Effective Climate for Research

A correctional agency will have taken an important step toward an effective climate for research if the administrator starts with or soon adopts a genuinely experimental stance toward evaluation. By this step he will have made it possible for members of his staff to "evaluate their own boss." That old adage will have lost its validity, to the extent that it ever was valid, and an important barrier to effective program and system evaluation will have disappeared.

There are other steps toward freeing up the climate for research. One of them is the acceptance of the research unit as an essential part of the organization, giving it responsibility for major undertakings, and valuing its efforts or products fully. Research reports will be objectively assessed, and if they point to action, plans will be made accordingly. If the reports are negative or inconclusive, open discussion and rational action rather than suppression or destructive criticism will be the typical modes of response by administration.

The achievement of such climates and relationships implies research staff of competence and integrity, and a research program with continuity. Heavy responsibility and full participation by research in the evaluation and development of programs and systems requires something more than apprentices or novices in research roles. This in turn presupposes adequate fiscal and organizational support for the evaluative activity, and the persistence of climates and structures in which able persons with appropriate training can perceive worthwhile career opportunities in correctional research.

Agencies which maintain an open mind toward evaluation, which encourage interchange between their own research units and those of other agencies, are more likely to attract and develop effective research staffs and to increase the utility of research as an instrument of management. They are also more likely to be exposed to new and potentially productive ideas. Consequently, a policy of free reporting of projects completed, hypotheses tested and observations made appears desirable. This may occasionally seem unwise from the public relations point of view, but as a consistent, long-range policy openness appears to be the most sensible procedure.

The development of an experimental, research-valuing, open climate in correctional organizations is only part of the administrator's concerns. He also needs to think about research needs and priorities, about in-house versus outside evaluation, about the

manner of staffing an in-house research unit if the decision is to proceed in that manner, about the funding of research, and about the utilization of research. These concerns are taken up in the pages that follow.

REFERENCES AND NOTES

1. Alfred McKinsey and Co., "Getting the most out of your computer: survey of company approaches and results," *Data Processing Yearbook, 1963–64.* Detroit: American Data Processing, 1963, pp. 15–18.
2. Philip M. Morse, *Operations Research for Public Systems.* Cambridge: Massachusetts of Institute of Technology Press, 1967, p. 12; Alain C. Enthoven, "The systems analysis approach," in Harley H. Hinrichs and Graeme Taylor (eds.), *Program Budgeting and Benefit-Cost Analysis,* Pacific Palisades: Goodyear, 1969, p. 166.
3. Robert H. Fosen and Jay Campbell, Jr., Common sense and correctional science," *Journal of Research in Crime and Delinquency,* 3 (July 1966), p. 75.
4. Donald T. Campbell, "Reforms as experiments," in James A. Caporaso and Leslie L. Rose, Jr. (eds.), *Quasi-Experimental Approaches.* Evanston: Northwestern University Press, 1973, p. 224.
5. The Center for the Administration of Justice, American University, Washington, D.C., has offered a number of institutes for correctional administrators, including an Institute on Perspectives on Research in Community-Based Corrections.
6. Evaluation Symposium, *Proceedings.* Raleigh: North Carolina Division of Law and Order, June 1973.
7. Daniel Glaser, *Routinizing Evaluation: Getting Feedback on Effectiveness of Crime and Delinquency Programs.* Rockville, Maryland: National Institute of Mental Health, 1973, p. 158.

CHAPTER 5. RESEARCH NEEDS AND PRIORITIES

Corrections is in many ways a very inefficient enterprise. It locks up in costly facilities many offenders who would perform better if kept under supervision in the community, as various studies are beginning to show.[1] Many of those who are locked up deteriorate under the experience—turning into even worse "enemies of society" than they were before, or changing into social or economic underperformers. Corrections does not bear sole responsibility for these occurrences; much of the problem lies with the courts, the police departments and the penal code, not to mention other elements in the social structure. Within its own domain, however, corrections appears to be a very ineffective means of achieving worthwhile social objectives.

Under such circumstances, it should be easily possible for corrections to make important practical gains by using research and development principles and techniques to improve its efficiency. To achieve these gains, it is necessary to give considerable thought to research needs and research priorities. Something better than a trial-and-error or "shotgun" approach to research is required.

In principle, research monies should be directed primarily into projects that the administrator deems important—that reform his clients, improve his procedures, increase the overall productivity of his agency, and help him function more effectively in relation to larger systems of agencies. Since all such research needs cannot be met simultaneously, they must be given a ranking—most important first. However, areas of greatest need are not necessarily areas of payoff. Some correctional and criminal justice problems appear insoluble, at least in the short run, no matter how much money is poured into them. Consequently, evaluative studies should reflect both perceived needs and perceived possibilities of achieving practically important results. How are decisions about these matters arrived at, and by whom?

There are several potential sources of definition of research needs, probabilities of payoff, and priorities. One of these is the mission of the agency. Another is agency requirements for self-maintenance and self-improvement. Still another is the perceived role of corrections in relation to other agencies or systems in the community. Finally, there are such matters as correctional experience and correctional theory. These areas give the administrator his main clues as to what research is needed and what is likely to yield practical results.

A. Agency Mission

The primary missions of the correctional agency are to provide a penalty or punishment, render the offender temporarily (or, sometimes, permanently) incapable of harming others, and rehabilitate or reintegrate him so that he can function acceptably and productively in the community.

1. *Offender Rehabilitation* appears to be the mission or objective that is given the most emphasis when corrections personnel talk about their work. This emphasis is reflected in the kinds of evaluative research that has been done in corrections over the past two decades. We hear much about the evaluation of treatment programs, but very little about evaluating punishment or incapacitation.

If rehabilitation is the principal objective of corrections, it might be assumed that this identifies a major research priority. However, there has been much criticism of the performance of corrections in this area. Some of the reviewers of evaluative research point to the absence of evidence of rehabilitative efficacy, which would indicate absence of practical payoff, which in turn might indicate the need for a low priority for treatment rehabilitation research.

This has now become a controversial point, since a "second look" at evaluative research on programs shows that many programs bring significant improvement in behavior or reduce the costs of treatment or both. The second look also shows that some of the "unproductive" research was poorly done, and a repeat of the studies in better manner might disclose that the treatment was in fact efficacious although not recognized as such at the time.[2] If this

is a widespread occurrence, it would be premature to drop treatment evaluation to a low priority.

It would also be premature to give low priority to treatment evaluation if there are promising treatment procedures or structures that have not yet been tried and evaluated. Since there *are* good ideas that need to be tried, there is still reason for planning to use some research funds in the evaluation of programs based on these ideas.

One of the options open to the correctional administrator at the present time is that of focusing on his present programs and trying to improve the quality of research as a step to upgrading his programs. Another is to introduce new programs and to evaluate these, preferably with better research than was available in the past. Still another is to focus on major system change.

With respect to change, he might adopt the rationale that with a different system, he would be more able to carry out what he considered his main objective—changing the behavior of his charges. This is a rationale that needs further scrutiny, although it is already supported by two or three bodies of evidence.[3] There is one further rationale for focusing on system change: the evidence from our six case studies of impacting research suggests that it is easier to change the system than it is to change the offender. Since the present correctional system is not distinguished for productivity, continuing search for alternative systems appears reasonable.

2. *Punishment*, or the imposition of penalties as part of the correctional process, has not been extensively studied by correctional evaluators. This neglect may have resulted in part from an intensive preoccupation with evaluation of positive treatment techniques, which were regarded as the "real" business of corrections.

One reason for considering a higher priority for punishment is that it may be practically more important than has hitherto been acknowledged. Punishment equitably and rationally administered may have a socially useful effect. On the other hand, punishment incorrectly used probably has a dysfunctional effect. It can be counterproductive as well as expensive. Both these possibilities have been discussed in the correctional literature, but systematic analysis of punishment and penalty applications in corrections has not yet occurred in any real sense. The practical consequences of such studies can only be guessed at presently.

3. *Incapacitation*, like punishment, has been little studied for its practical relevance. Its importance is quite obvious in some cases, as when an offender shows a "propensity" for injuring others; its necessity is not so obvious in other cases, as when long prison terms are served for some "victimless" crimes. A long term in some of these instances appears to be a very irrational use of community resources. Studies of incapacitation, possibly by combinations of cost-benefit and behavioral analysis, may deserve considerably higher priority that has hitherto been accorded them.

B. Agency Maintenance and Improvement

The administrator of a correctional agency has a recognized obligation to operate it effectively and to modify its structures and processes as the state of the art improves. These obligations imply that the administrator will strive for some form of quality control and that he will seek to keep his agency abreast of development in the field—resources permitting.

1. *Quality control*, or maintenance of standards, involves routine or periodic evaluation. This can be either judgmental or in the form of statistical monitoring on relevant indicators. The alert administrator will know what his return-to-system rates are, for example, and whether his yearly admissions are getting "tougher" as the years go by.

For small departments, or departments with budget problems, there may be no research unit and no information system on punch cards or in a computer. In such cases, staff will have to monitor its effectiveness by judgment or intuition unless an administrative staff member tabulates from samples of file cards some minimal data on intake and departures and returns.

If the department can procure one or more research or statistical personnel, it can have the data on yearly releases, intakes and returns collected, tabulated and interpreted. The addition of a punch-card or computer-based record system would increase the capacity of the research and statistics personnel. From such a unit the administrator could expect periodic information on general outcomes (return to system, at least), outcome by type of individual, type of offense, program experience, trends in outcomes, and trends in type of intake.

Such monitoring of outcome and intake is perhaps the first priority of the agency insofar as evaluative research is concerned. How extensive the analysis of data of these kinds should become as the research staff and information resources of the agency increase is not clear. The payoff from such research is hard to estimate, but the need for at least a modicum of this kind of research is inescapable.

2. *Agency Improvement* is a constant source of concern to administrators in a field such as corrections where visibility is high and criticism is abundant. Along with pressures from the outside there is an additional impetus toward improvement in the form of published correctional standards. Both these sources create need for research, either in the betterment of old programs or in justification of new programs.

a. *Modification of old programs* can be done in several ways: on the basis of administrative judgment, with the aid of elementary research, or with the aid of more elaborate studies including controlled experimentation. Prior to the era of modern management, administrative judgment was the only method, and in most correctional agencies, this is still the predominant method. Research resources are too scarce to use on most old programs, although in a few instances we have seen costly controlled experimental evaluation of old programs. Experimentation on parole caseloads in California during the 1950s and 60s is perhaps the outstanding example of attempted program modification through research.

From today's perspective, the fifteen or twenty years of parole caseload research has not been as productive as one might have hoped,[4] and there seems to be no great incentive to perform similar series of experiments on other old programs. Some cases of research and development of ongoing programs have undoubtedly worked better than others, and systematic studies of old program modification might disclose which methods have proved more successful thus far.

b. *The introduction of a new program* is a logical occasion for making an evaluation. The effectiveness of new programs is generally unknown, and it is prudent to plan for the retention of such programs only if they can be shown to be effective. Otherwise, they may become established and persist despite the fact that they contribute nothing to the overall effectiveness of the agency.

The practice of evaluating new programs is a break with tradition, since it has been customary to introduce them without planning for formal evaluation. Prison colleges and halfway houses illustrate the traditional procedure. Prison colleges have been introduced into at least 168 prisons since the first opening at Leavenworth in 1953, but it was not until 1972 that the first rigorous evaluation of a prison college was carried out.[5] Similarly, an estimated 300 halfway houses for criminal offenders have been established in the nation's communities since about 1950, but there have been only a few rigorous evaluations of these programs, and those only in recent years.[6] How differently corrections might have developed had there been intensive research and development on halfway houses in the 1950s is an interesting speculation.

The pretrial diversion movement illustrates a possible new trend in program evaluation. Since the inauguration of the Manhattan Court Employment Project in 1967, approximately 30 pretrial diversion programs have been started. Approximately half of these have been evaluated by either a quasi-experiment or a controlled experiment, with cost-benefit analysis as an added feature of research design in about half the experiments.[7] This is a heavy emphasis on early evaluation of new programs, compared with the prison college and halfway house examples.

Although pretrial diversion programs may be a special case, it seems more likely that the concept of prompt and early evaluation of new programs is catching on. The generally favorable results reported in most of the studies—diversion from adjudication and lock-up, job status and earnings improvement, recidivism reduction, and benefits in excess of costs—have apparently hastened the adoption of the diversion concept in other localities.

The pretrial diversion studies disclose a number of advantageous features. One is the early start on evaluation, which facilitates rejection of poor procedures and selection of good procedures before the program becomes crystallized. Another is the possibility of between-program research comparisons, which may find validity problems in some of the studies and permit reinterpretation of the results. Still another is the possibility of structural and procedural comparisons between the programs, which facilitates the identification of good procedures.

In the drive to obtain early evluation of new programs, no guidelines have emerged as to the kinds and numbers of evaluations that should be made. If 200 pretrial diversion programs are to be established eventually, how many of these programs should be evaluated? And for how long should developmental study be carried forward to insure that the programs are brought to relative perfection, not left to solidify in crude or primitive states? Questions such as these must be deliberated about and acted on by adminis-

trators of agencies, program supervisors, funding entities and coordinating personnel.

C. Corrections as a Subsystem

A long-range approach to research-need definition is that of visualizing the correctional agency as a subsystem within the criminal justice system or within the larger social system. This requires us to examine corrections' objectives, structures and functions in broader perspectives.

1. *The criminal justice perspective* forces the correctional administrator to think of his subsystem in all its relationships with the courts, police, bail agency, public defender system, and other activities. He needs to devise plans, organize information, make recommendations and take actions that are more conscious of the objectives and needs of the criminal justice system as a whole. Obviously, this approach to agency operation requires reciprocal behavior on the part of the other agencies or entities concerned.

Research under these circumstances will include increasing numbers of collaborative studies. It will cross-cut agencies, and its findings and recommendations will be of concern to more than one of the subsystems. In format, such research is likely to show new features. Initially, the "systems" studies may be primarily case studies, surveys, and cohort followups. In time there will be an increasing number of systems analyses and simulation studies. Attention will shift somewhat from criteria of performance to system rates.

2. *The social system perspective* places the criminal justice system alongside the neighborhood, the family, the school and other social entities as subsystems that interact with the developing individual to make or break him as an effective, functioning member of the community. These subsystems are at once interdependent and limited in their capabilities. A faulty criminal justice system will work against the best efforts of home and neighborhood. On the other hand, a perfect criminal justice system cannot make up for the lacks in the other social subsystems. When these several entities work together on common objectives with good understanding of what works and does not work in the management of deviance, first at younger ages and later with mature persons, then the efficiency of social control of delinquency and crime will be maximized.

In an ultimate sense, of course, the deviance control system will include all the institutions or agencies of society. Not only the family and the school but also economic and governmental organizations define the "opportunity structure" and the "control systems" that impact on the rate of social deviance. So corrections and criminal justice must depend heavily on the total social structure for rational solutions to the problems of delinquency and crime and reintegration of the offender.

The primary responsibility for research and development in the larger systems does not lie with the correctional administrator. As time goes on, however, he will become increasingly involved in the definition of these research needs and in otherwise placing the tasks of punishing, incapacitating and rehabilitating in proper perspective.

D. Experience as a Definer of Research Needs

Experience work in numerous ways to define research needs and priorities. The emergence of a problem in operating the correctional agency will immediately suggest a study of some kind—an administrative study, descriptive research, or an evaluative research project. Some of these events are the "brush fires" that researchers often lament because they see this as digression from the "real" work of evaluation.

The accumulation of ordinary experience, in the form of casual information or hunches, will sometimes spark recognition of a jumping-off point. An exploratory study, or something more elaborate, may be called for to resolve or to clarify an emerging issue of practical significance.

Systematized experience, such as reports and studies, generate further studies. Completion of phase one of a study may indicate that a phase two is needed and probably worthwhile. The first evaluation of a program may suggest some definite modifications in structure and process, and a re-evaluation of the modified program.

These proddings by experience may be entirely intra-departmental, or they may be aided by reported experiences in other agencies that arrive by various dissemination processes.

E. The Role of Theory in Defining Research Needs

On its face, correctional or social-behavioral theory should have an important role in defining research needs and priorities. The vital contributions of theory to the rise of the scientific tradition make this view plausible. And some researchers even assert that theory is the key to productive research in corrections. As Glaser states it, "The primary cause of a poor yield from criminal justice research, I believe, is more often a poverty of theory than a dearth of methodological skill." [8]

Glaser's point of view is a defensible one, widely held by academicians, although some academicians would give a higher priority to method than to theory. There is, of course, a third alternative: the key to productive research may be simply the practical relevance of the research. The six case studies in Chapter 3 described research that was highly productive yet generally simple in method and theory. The studies might, in fact, be described as atheoretical and amethodogical, even in the case of the relatively sophisticated Community Treatment Project.[9]

This appeal to fact suggests that the claim to the importance of theory, while deserving a hearing, is overstated from the standpoint of the agency administrator. There is every reason to suspect that the primary cause of a poor yield in correctional and criminal justice research, in the present era, is lack of practical relevance, not lack of theory or method. Whether this situation will change in the future is not clear. It might be anticipated that as a research enterprise matures both theory and method will become relatively more important in its functioning. But what theory, and what method? Systems theory or behavioral theory? Operations research or controlled experimentation?

There is, admittedly, a certain minimum of unsystematic theory and basic research method that every researcher makes use of in carrying out his varied responsibilities in a correctional agency. Beyond this ill-defined level, however, both the correctional administrator and the correctional evaluator seem perfectly justified in asserting that their experience indicates that practical relevance comes first, with theory and method filling in where they can. At some future date, theory may indeed demonstrate its key importance in making research productive.[10] In the meantime, the administrator should be wary of invitations to invest heavily in theory testing or in trial applications of elaborate methodologies.

F. Discussion

It is clear that decisions on kinds and amounts of research in a correctional agency pose some real issues for the administrator and his evaluative staff, if any. If the agency lacks research staff, and has no budget for research, the best the administrator can hope for, internally, is some monitoring of the nature of his population intake and the quality of his outcomes.

Such monitoring should be a part of every agency's activity. The extent of the monitoring, and the complexity of the evaluations it will permit, will depend on the staff that can be committed and the quality of the record system that is or can be made available. With moderate resources, the monitoring can routinely relate measures of outcome such as return to the system or parole success to population subtypes and to experience in the system.

At higher levels of research support, program improvement and system improvement can become matters of focus. Evaluation of old programs may occasionally appear desirable and feasible. Evaluation of new programs should be taken as a necessity—unless the programs have been thoroughly tested elsewhere and their adoption in the agency appears justified without further evaluation.

Evaluative studies that point toward extensive system change are appropriate activities if research resources, departmental interest and capability, and political climate are all favorable. Since offender change has thus far been an elusive goal of corrections, system change may deserve higher priority. New systems may be able to accomplish what the old have failed to do.

System change at the criminal justice level also has acquired higher priority, and much new activity in the field is occurring as a result. This activity will undoubtedly entail new roles for the SCA administrator and evaluator.

REFERENCES AND NOTES

1. The California Youth Authority's Community Treatment Project and the Provo Experiment by Empey and Erickson have shown the superiority of in-community management for some types of juvenile offenders. The District of Columbia's Narcotics Treatment Administration has provided similar evidence on one type of adult and young adult offender. For the Provo material, see LaMar T. Empey and Maynard L. Erickson, *The Provo Experiment*, Boston: Heath, 1972. On the effect of the Narcotics Treatment Administration, see Robert L. DuPort and Mark H. Greene, "The dynamics of a heroin addiction epidemic," *Science*, 181 (August 24, 1973), 716–22.
2. The PICO Project is a classic example of the dangers of accepting the first evaluative findings on a particular type of program. The PICO data were first analyzed by one correctional agency and a report of negative findings was returned. The same data were later reanalyzed by another correctional agency, using simpler methods, and the report this time showed highly sig-

nificant and practically important returns. The second report stood up under critical review, and thus findings that were apparently negative turned out positive under a different research strategy.

3. The three cases cited in Note 1 are all examples of significant system changes, from treatment in an institution to treatment in the community, each accompanied by superior performance on the part of at least some definable types in the treated population.
4. In one sense, the several phases of the SIPU experiment (1953–64) and the Parole Work Unit Project (1965) and the follow-on Work Unit Program (1965 to the present) ended in the high impact of the latter program. It is not clear, however, whether this is an illustration of productive "research" or productive "improvisation" by "research wise" managers and practitioners who have been confronted by a massive case of null findings.
5. Colleen Barros, William Caldwell and Stuart Adams, *Effectiveness of the Lorton Prison College Project: Third Interim Report*. District of Columbia: Department of Corrections, Research Memorandum 72–4, May 1972.
6. Judith A. Hecht, *Effects of Halfway Houses on Neighborhood Crime Rates and Property Values*. District of Columbia: Department of Corrections, Research Report No. 37, November 1970.
7. Neal Miller, *Evaluation of Research on Pretrial Diversion*, Washington, D.C.: American Bar Association, 1974, p. 41.
8. Daniel Glaser, "The state of the art of criminal justice evaluation." Keynote speech at the Second Annual Meeting of the Association for Criminal Justice Research (California), Los Angeles, November 9, 1973.
9. Malcolm W. Klein, "Collaboration between practioners and researchers; relevant knowledge in corrections," *Federal Probation*, 37 (December 1973), 42–46, discusses the "almost totally atheoretical" nature of the Community Treatment Project (p. 43).
10. Wilkinson has introduced an interesting note into discussions of the importance of theory by commenting that contemporary social theory has "no more predictive power than astrology." This might suggest that we need to develop better social theory, but Wilkinson sees the answer to the problem of social decision and control in "mathematical languages" rather than in the "natural language" theory of the social sciences. See John Wilkinson, "A proposal for recycling our junk heap society," *Center Report*, 7 (June 1974), 7–12, p. 9.

CHAPTER 6. STAFFING AND FUNDING THE EVALUATION EFFORT

Correctional administrators may follow one or more of several possible routes to obtain staff to perform evaluative research. They may hire an in-house research staff, contract with a private non-profit organization, employ university faculty members, or seek the services of a for-profit consulting organization. These are the best known of a larger number of possibilities. All have their advantages and disadvantages, and it is important for the purchaser of research services to be familiar with both.

A. In-House Research Staff

The in-house staff is in many ways the best single solution to the evaluative research problem if funds can be budgeted and suitable personnel can be recruited. Since only about one half the nation's SCAs have taken this route, it may be presumed that some administrators have not felt a need for research capability within the agency or there has been difficulty in persuading the legislature to allocate funds for this purpose.

Among the twenty or so states that have correctional research units, only a small number, possibly two or three, have what may be termed a "large research capability." The California Youth Authority, with a research division budget of 1.2 percent of the 1971 annual operating budget, appears to be the most amply staffed in relation to resources.[1] The California Department of Corrections, with a budget over twice as large as that of the Youth Authority, commits only about 0.5 percent of its operating budget to research activity.[2] In 1971, the District of Columbia Department of Corrections allocated 2.4 percent of its operating budget to planning and research (including data processing).[3] This was a sharp climb from 0.7 percent in 1967. The 2.4 percent in 1971 was a peak, and following a reorganization of the department in 1972, the allocation to research dropped sharply in the next two fiscal years, falling to about the 1967 level.

The sizes and functions of the majority of the twenty in-house research units in SCAs are nowhere clearly documented. It appears that size ranges from one or two professionals in the smaller units to 20 or 30 in the larger. The smaller units serve primarily as statistical reporting sections; the larger engage in a wide variety of reporting, exploratory, and evaluative research, including studies aimed at bringing about major system modifications.

The work of these units, relatively speaking, is good. A review of the work suggests that the best evaluative research on agency programs in the past twenty years has come from this source. On characteristics such as quality of research design, objectivity, and relevance to operational needs, some in-house studies are clearly superior to the best agency program evaluations by university faculty, private non-profit, or for-profit research organizations. We see in this category such studies as the Community Treatment Project,[4] the Youth Center Research Project,[5] and the PICO Project.[6]

In addition to the superiority of the research, there are other characteristics of the in-house units that are advantageous. The staff is familiar with the objectives, philosophy and problems of the agency and does not have to undergo a costly familiarization process at the beginning of each new project. It is able to hasten utilization of relevant research findings by taking the lead in the planning and development of new programs. And it communicates more effectively with administrative and operations staff because of its familiarity with the department and its greater commitment to departmental objectives.

There are a number of disadvantages to such units, some important to administrators and some to researchers. Many of the units, because of location or salary scales, have difficulty attracting qualified staff. The settings are often isolated and limited in opportunity for professional growth, so there is strong incentive to move on to more interesting or rewarding situations. The units are within the civil service, which makes it difficult to remove staff members who prove unqualified for their tasks. And the environment of the units sometimes converts staff members into opponents of the "establishment," and they may reflect this actively in their research or in their general discourse and publications.

Perhaps the most critical question for the administrator is not whether he should have an in-house unit but how large it should be. All agencies need an evaluative capability, if only one person to monitor trends in agency outcomes and to examine these trends for relationships to intake, programming and historic factors. How far to go beyond this is an open question. Some observers have proposed that research staff be built to a level that requires about 5 percent of the operating budget for support.[7] This raises issues, such as how many of the 50 SCAs should strive for this level of support, the economics of a total research effort of this magnitude, and the relation of a large SCA evaluative effort to the growing SPA effort. Should there be an enlargement of SCA research efforts at a time when emphasis in planning and evaluation appears to be shifting from agencies as subsystems to criminal justice as a system? These are issues not so much for the agency adminstrator as for the criminal justice field in general.

B. University Faculty

University faculty members have long played a conspicuous role in correctional research, primarily descriptive and theoretical, but in some instances evaluative. By interest and training, these individuals have been frequently well suited to this role. A number of outstanding theoretical and evaluative studies have thus been produced, and both the concepts and the empirical data have influenced the thought and practice of corrections.

In addition to the interest and training, the informational background, and the theoretical and methodological skills, university faculty members have one other advantage for the administrator: they are present everywhere. With thousands of colleges and universities dotting the nation, it is easy to make contact with a faculty member who has an interest in corrections. This has made possible the establishment, over the past 20 years, of the 168 prison college programs taught in person by members of college and university faculties. It also facilitates the recruitment of faculty members for consulting and evaluative tasks. In the past five years, there has been a very rapid expansion of faculty member involvement in evaluative research in corrections, and this may now be the primary staff resource for project and program evaluation.

In practice, the research performance of university and college staff is marked by a number of difficulties. These need to be understood by a correctional administrator who seeks assistance in evaluation. These difficulties might best be summarized through the comments of specific agencies and persons who have made use of or have observed the work of university staff.

LEAA, in a conference on evaluation called by the U.S. General Accounting Office in 1973, reported that it had experienced disappointment with the granting of monies "hands-off" to a number of universities. "Although a variety of methods was used to carefully select the universities, LEAA was (hard) pressed to identify any results from the research.[8]

Etzioni, in 1972, commented on his impression that university faculty members were not good prospects for applied research tasks. They tended to turn the tasks into basic research projects to bring them into line with their own academic interests—a practice that Rossi has described as "Robin Hooding." [9]

Brooks, who spent some years as a director of a program evaluation center, saw the problem as partly that of the academic reward structure. This structure was not designed to encourage faculty members to participate in the evaluation of agency action programs. There were other problems: the time limits on grant-funded projects, the need to gather and analyze data on projects designed by others, and the limited opportunity to produce scientific papers out of evaluation research. These obstacles discouraged most faculty members who were qualified to do such research and they remained aloof, leaving the field to the inexperienced and unqualified.[10]

The latter point may help explain a remark made somewhat candidly by Berkowitz. She commented that her state planning agency has disbursed a large amount of money to faculty researchers, "And all we get back is junk. It's because the researchers, or the pseudo-researchers . . . don't seem to know the applications of the basic tools of research. Or they work up a beautiful research design, and they mess it up because of contamination of data . . . It's definitely a technological problem." [11]

If the problem with the academician is self-interest, inexperience, or methodological rigidity, a slightly different problem may arise with the for-profit consulting organization. Roos notes that whereas the academician occasionally pays too little attention to the administrator's needs the consulting organization sometimes pays too much attention to the administrator's wishes. "Instances have been observed where a consulting organization asked to evaluate a program provides its client with a whitewash which the evaluator assumes, or has been told, the client expects." [12]

The National Science Foundation, in a report released in 1969, before the evaluation "crunch" in criminal justice reached its most critical stage, diagnosed the staffing problem in a different way:

> "When social scientists have been called upon as consultants in practical situations, many have failed to deliver. Two general complaints have been voiced about social scientists as consultants. First, communication has been impeded because social scientists speak in a jargon incomprehensible to the layman . . . (Secondly,) when faced with a problem that has no ready-made conceptual answer, they frequently retreat to the laboratory for more research and more facts. But the client would ordinarily settle for a less than scientifically adequate answer. He simply wants the consultant to apply his trained intelligence, and give help based on the information at hand." [13]

One final comment will round out this commentary on academicians as applied researchers. Abelson, the editor of *Science*, remarked on the difficulties that industry experiences in securing the right kinds of researchers from the universities. "Industry has found that the typical Ph.D. tends to be inflexible. He usually wishes to keep on redoing his thesis." [14] As a consequence, Abelson added, industrial firms have occasionally found it advantageous to hire persons at lesser levels of education and to finish training them within the firm's research unit.

The typical correctional administrator ordinarily lacks the resources of the facilities to train his own researchers. Usually he must either contract with university faculty for research tasks, or recruit one or more relatively inexperienced researchers and trust that they can learn by doing what appear to be needed evaluations. If he adopts the former course, he will find it advantageous to specify quite clearly what the agency expectations are in research objectives and the form of final products, and to obtain definite assurances that these expectations will be met.

C. Private For-Profit Research Firms

A source of research products somewhat less readily available than university faculty is the for-profit consulting firm. These enterprises have in recent years begun to diversify into areas other than defense or industry and to include social program evaluation in their work. They often include a wide range of specialists on their staffs, including social scientists, mathematicians, economists, computer experts, behavioral scientists, anthropologists, historians and others.

These firms emphasize the use of computers, systems analysis, advanced statistical techniques, and, where the data and problems justify it, simulation. Their work tends to be problem-oriented, strong in technique or methodology, and weak in correctional knowledge and theory.

The performance of these firms ranges widely, from very good to very poor. One of them has in recent years adapted one of its techniques to produce a prediction instrument that scores prisoners for probable success in work release. The instrument appears to be markedly better than any of the existing prediction instruments in corrections. Developed for the District of Columbia Department of Corrections under a $15,000 contract with LEAA, it shows an efficiency (measured by the Mean Cost Rating method) twice that of the linear discriminant function and configuration instruments already widely known in corrections.[15]

Another of the for-profit firms recently completed an evaluation of identical correctional programs in several locations that appears to have been an expensive research disaster. Although well designed, the study was directed from a central location, and control over data collection and design integrity was poorly maintained. One or two concessions to all the program managers, who seemed either unconcerned about or irked by the evaluation, destroyed the possibility of rigorous measurement at every site. Extensive and elaborate statistical analysis of the truncated data failed to repair the damage done by poor project execution.

At their best, the for-profit consulting organizations offer varied experience, methodological expertise, an interest in the client's problem, business-like methods, and a concern for doing work that will be judged satisfactory by the client. At their worst, they show great naivete about the agency's objectives and procedures, a single-minded concern about making a profit on the undertaking, need for assistance in research design and execution in an unfamiliar setting, and a consulting firm jargon that is as incomprehensible as that of some academic researchers.

As in the case of the university faculty, the correctional administrator who seeks evaluative services from a consulting firm should specify clearly his research objectives and the anticipated products of research and have these spelled out in the form of a contract.

D. Private Non-Profit Research Organizations

Several organizations engage in correctional research on a non-profit basis. These include such entities as the National Council on Crime and Delinquency, the American Justice Institute, and the American Correctional Association. In some cases the organization includes standard-setting for the field as part of its function. Funds for the research come from federal or private grants or from the budgets of the agencies requesting the research.

The staffs of these organizations are typically former officials of correctional agencies and researchers who are potential careerists in correctional evaluation. Both administrative staff and researchers usually have had extensive experience in the field of corrections, and they display a commitment to the field.

Although the evaluators in these organizations often have training in the social or behavioral sciences to the doctorate level, they are not usually specialists. They sometimes make use of academic or research firm consultants in evaluative studies requiring high expertise.

Some of the work done by these organizations is excellent. One of the best of the fifteen evaluations of pretrial diversion studies referred to earlier (Chapter 5) was carried out by a non-profit organization.

Because of the limited number and size of these organizations, their part in the total correctional evaluative effort is necessarily small. They make a special contribution to the effort by reason of their strong commitment to advancing the field, by identifying and carrying forward pioneering kinds of studies, by seeking to coordinate various program development efforts, by promotion of operating standards, and by offering several kinds of technical assistance to state and local agencies.

E. Discussion

The staffing of evaluation in a correctional agency poses a number of interrelated issues for the administrator. The number and complexity of the issues will depend upon how the administrator defines his research needs and priorities. The fewer the needs, the fewer the issues.

If the principal need of the administrator is for a modest amount of data compilation and some year-to-year comparisons of intakes and outcomes, a small statistics and research unit will suffice. Such a unit can provide various kinds of data for management purposes as well as some elementary monitoring of results. For most SCAs and LCAs this might be the most suitable arrangement. These agencies can leave to larger and more heavily budgeted organizations the role of building sizable research staffs and serving as pioneers in testing existing programs and evaluating new programs. When such programs have been shown to be effective, not only on the original site but also in new locations, they can then be widely diffused. This reduces the need for research staff and avoids wasteful duplication of evaluations at numerous sites.

The administrator who wants to participate in the refinement of old processes and structures and the testing of new concepts will need to consider whether to add research staff or to seek research services from the outside. Building a research staff takes time, since budgeting, recruiting, and familiarization must occur before research planning and implementation can begin. Two or three years after the budget request has been submitted, the first evaluative research products may being to issue from the unit, and it may be longer before highly useful findings appear. The realistic administrator will keep in mind that only a very small percentage of research projects pay off. The payoff is apparently higher from in-house units than from academic or consulting firm research, although the latter have the advantage of being easier to start up when there is a decision to go and easier to disband when there is no reason to continue further.

If the administrator decides to proceed with process or program testing, he will presumably have reviewed his research priorities to define important areas for evaluation. He will also have made preliminary checks to ascertain whether other agencies are about to finish what he is about to start. One of the unsatisfactory aspects of correctional research is the vast amount of uncoordinated and duplicative effort that is currently underway. In the absence of good channels of communication and definite rules of procedure, much of this waste is inevitable. It is now quite clear, however, that there is a crying need for someone—possibly the research units of correctional agencies, possibly some national office or organization that has a fundamental interest in increasing the efficacy of evaluation—to seek coordination of effort.

There are several reasons for this. One is simply the need to conserve research resources. Another is to make possible standardization of effort. Since some studies need to be duplicated to make sure that initial findings are not erroneous, when the duplication is by different agencies, the several designs and implementations should be as similar as possible. Still another is the need to perform comparative studies in which the same design is applied at several sites simultaneously, possibly by a single research organization. The comparative data that come from such a study are likely to be much more informative than replications or duplications spread over time and agencies in several independent efforts.

The correctional administrator is, of course, not responsible for initiating this process of coordination. For the present, his primary concern may be only to avoid starting an evaluation project that some other agency has already completed in a competent manner. In time, however, the administrator whose agency is heavily involved in evaluation will come to appreciate the need for more effective coordination of evaluative studies between agencies.

If the administrator's evaluative ambitions run as far as making dramatic changes in his agency, or in the criminal justice system as a whole, the staffing requirements and related actions are much harder to anticipate. Changes on the larger scale tend to be political as well as administrative and scientific. While research can precipitate the change, it may play only a small role in its execution and in the maintenance of the new state of affairs. Furthermore, changes that go beyond the boundaries of the agency will inevitably involve other agencies and interests, so that initial planning for such change will involve public relations, education, and possibly political action.

REFERENCES AND NOTES

1. State of California, *Budget Report 1973–74 FY*, p. 184.
2. State of California, *Budget Report 1973–74 FY*, pp. 348, 378.
3. Stuart Adams, *Trends in Allocation of Resources to D.C. Correctional Functions: 1967–1971*. District of Columbia: Department of Corrections, Research Report No. 23, 1970, p. 9.
4. Marguerite Q. Warren, "The community treatment project," in Norman Johnston, Leonard Savitz and Marvin E. Wolfgang (eds), *The Sociology of Punishment and Correction*, New York: John Wiley & Sons, 1970 (2nd ed), pp. 671–83.
5. Carl Jesness et al., *The Youth Center Research Project*. Sacramento: American Justice Institute, 1972 (2 vols.).
6. Stuart Adams, *Interaction Between Individual Interview Therapy and Amenability Classification in Older Youth Authority Wards*. Sacramento: California Youth Authority, Research Report No. 20, January 1961. Reprinted as "The PICO Project," in Norman Johnston et al (see reference 4, above), pp. 548–61.
7. Norval Morris and Gordon Hawkins, *The Honest Politician's Guide to Crime Control*, Chicago: University of Chicago Press, 1970, p. 252. See also *The Organization and Effectiveness of the Correctional Agencies*, Washington, D.C.: American Correctional Association, June 1966, p. 739.
8. *Evaluation of Law Enforcement Assistance Administration Programs: A Conference Summary*. Washington, D.C.: National Academy of Public Administration and the U.S. General Accounting Office, February 22–23, 1973, p. 9.
9. Amitai Etzioni, "Redirecting Research Dollars," *Washington Post*, June 11, 1972, p. 133.
10. Michael E. Brooks, "Dimensions of and Constraints on Evaluative Research," *Proceedings*, Southeast Evaluation Symposium, Raleigh: North Carolina Division of Law and Order, 1973, p. 14.
11. Francine Berkowitz, in *Proceedings: Workshop on Research, Session II*, 43rd annual meeting of the California Probation, Parole and Correctional Association, San Francisco, May 30–June 1, 1973, pp. 26–27.
12. Noralou P. Roos, "Evaluation, quasi-experimentation and public policy,: in James A Caporaso and Leslie L. Roos, Jr. (eds.), *Quasi-Experimental Approaches*, Evanston: Northwestern University Press, 1973, p. 297.
13. National Science Foundation, *Knowledge Into Action: Improving the Use of the Social Sciences*. Washington, D.C.: Government Printing Office, 1969, p. 15.
14. Philip H. Abelson, "Additional sources of financial and political support for science," *Science*, 180 (20 April 1973), p. 259.
15. Fair, Isaac, Inc., *Development of a Scoring System to Predict Success on Work Release*. Washington, D.C.: Department of Corrections, January 1971.

CHAPTER 7. THE UTILIZATION OF RESEARCH

Researchers and the funders of research in corrections often complain that research products are not used by correctional administrators and their staffs. Weiss has spoken for numerous researchers in her remark: "The most common complaint (of evaluators) is that their findings are ignored." [1]

If there is truth in this complaint an obvious problem exists: for the researcher, the administrator, and the community. The researcher sees his accomplishments wasted, the administrator fails his obligation to keep up with advances in his field, and the community suffers from inefficiency in rehabilitation and reintegration.

Just how serious a problem nonutilization of research is has never been carefully examined. The extent of the problem has perhaps been overstated both by researchers, who tend to be highly articulate, and by funding sources, who have reason to be concerned about waste of their monies. If administrators are asked about the matter, they tend to see the problem as one of inefficiency or inutility or both. Their views, in composite, are essentially as follows: "Research is a process of using up far too much money to produce an incomprehensible and irrelevant report that is delivered long after it was promised."

Whether it is a matter of administrative default or ill-conceived research, the important thing is to understand whether there is a problem, what it is, and what solutions may be suggested. As presently defined, the problem appears complex. One aspect is the low payoff in research whatever the field. If research in high-technology is successful only five percent of the time, the researcher in corrections who wants more than one hit in twenty tries may be quite unrealistic. Another aspect is the doubtful quality of research. If it is Berkowitz' research contractors (p. 30) who complain about nonutilization of their products, their complaints may be unjustified. Still another aspect is the poor fit between researcher styles and administrator needs. If researchers strive toward higher levels of theory, method and certainty while administrators want merely relatively informed advice in uncertain situations, communication between the two groups will remain poor.

Some aspects of the problem relate primarily to the administrators' situations and operating philosophies. Many administrators do not welcome change; their anxieties, personal views, and preference for order and tranquility in their institutions may turn them against innovation, even when it appears reasonable, constructive, and butressed by scientific evidence. Others exist in precarious balance, and the operationalizing of new, even though tested concepts, increases the difficulty of their position. Still others have seen many promising new ideas come and go, leaving disillusion in their wake. Finally, new ideas may call for increased budgets, and the skepticism of legislative committees about budget expansion in corrections cuts deeper than researcher ire.

A. Basic Questions About Research Utilization

To clarify the matter of whether there is a serious research underutilization problem, it will be useful to examine some of the issues and to suggest how any substance in these issues might be dealt with. The key points of concern appear to be the receptivity of administrators to new ideas, the extent of nonutilization of worthwhile products of research, and factors that govern utilization or nonutilization of research products. A review of these issues may suggest solutions to what some perceive as a major problem in the evaluation of action programs.

1. *Administrative Acceptance of New Ideas.* It is possible to have two contrasting perceptions of change in corrections. One is that the basic structure of contemporary corrections, which is anchored to incarceration of offenders in large prisons remote from population centers, is tradition-ridden and resistant to change. The other is that corrections is a social system in transition, with changes occuring in rapid succession, particularly in the past twenty or thirty years.

Either of these perceptions comes forth strongly, given the right perspective. The person who looks for change sees in recent years the rise of halfway houses, prison colleges, work-release programs, study-release programs, furlough programs, family visiting, behavioral modification programs, co-educational prisons, diversion programs and numerous other departures from tradition.

In virtually all these instances, the changes were introduced as untested concepts that were believed by some administrator to be worth trying. Each concept incorporated a social value not initially accessible to prisoners but potentially able to influence the rate at which ex-prisoners adjusted in the community. During the initial trials of the concepts, there were opportunities to observe how the new programs worked, and favorable judgments were pronounced. Thereupon the ideas diffused more or less rapidly to other sites.

The case of co-correctional (coeducational) prisons is illustrative. In 1972 the concept was given a trial at a federal prison in Forth Worth. Since then the idea has spread to several states—Massachusetts, Virginia, Connecticut, and Texas—its spread hastened by qualitative judgments expressed by various kinds of observers. One corrections official remarked that the arrangement had a "civilizing effect" on the prisoners. A social scientist asserted that "It works as a defusing situation. There is an entirely different atmosphere." [2]

These instances of administrator-initiated change, arising out of concern for improvement of system effectiveness and diffused by qualitative judgments of approval, are evidence of the readiness of the field to accept new ideas and new structures or processes. If administrators will accept untested but reasonable operating concepts, why will they not accept tested and reasonable concepts that are presented to them by researchers? Or are researchers and funders of research projects misreading the situation?

2. *Nonutilization of Research.* There are some clearly established instances of nonutilization or even rejection of correctional research findings. The Provo Experiment is a vivid illustration of the rejection of scientifically significant and operationally useful findings in youth corrections.[3] The rejection process, initially documented by the authors, has been commented on by Glaser.[4]

Another instance of nonutilized research concerns a highly efficient method of selection for work release developed for the District of Columbia Department of Corrections in 1971.[5] This prediction device, when tested on a validation sample of work release candidates, exceeded considerably the efficiency previously shown by prediction instruments developed in California by discriminant function methods or based on the configuration analysis method publicized by Glaser.[6]

Despite the possession of a markedly superior instrument, DCDC officials failed to act on the recommendations for use that were included in the validation report. The officials spoke of uncertainty as to which of the alternative procedures to adopt, and of long-range plans for a pre-release program, which ought to be set up before the scoring system was put into effect.

An unmentioned factor that may have had a bearing was the process of reorganization under way in the department. This obviously took priority over an innovative selection procedure for work release. A second possible factor was the declining role of research in the department. Although in several years preceding the Department had worked actively to implement some major new programs recommended by research staff, the reorganization implied a reduced role for research, and this may have helped tip the balance against the selection device.

One significant sidelight of the deferred action on the device was the refusal of the department of corrections in a neighboring state to consider seriously the adoption of a modified form of the instrument for its own use. "If D.C. Corrections, who developed it, isn't going to use it, why should we?" This suggests that nonutilization of a research product by its owners or immediate purchasers may react adversely upon the process of technology transfer.

How extensively nonutilization reduces the potential impact of research and detracts from the possible productivity of corrections departments is not clear. It is probably a less serious problem than researchers imagine but very likely a more serious problem than one would conclude after talking with correctional administrators. Both sides react to this issue in rather subjective manners.

Since the vast bulk of correctional research has no apparent payoff, it would be wise to learn more about some of the possible obstacles to payoff. On the one hand we may have research that is useless for reasons of error, illogic or irrelevance. On the other hand, there may be factors that impede the utilization of definitely usable research. We have much speculation about this matter but few facts or concepts. It is obviously time to look more systematically at what can be agreed upon as "definitely

usable" research that is not being used, particularly as to why it is not being used and what might be required to change matters.

This gives us two areas of research utilization that are in need of closer scrutiny in the interests of increasing research productivity. The first is a review of projects that have an impact, focusing on characteristics, setting, types of effect, the magnitude of effect, and, if possible, explanations of the reason for impact. The other is the characteristics, setting, and reasons for failure of recognizably useful research which has not been put to use. An exploration of some of the factors related to impact was made in Chapter 3. A glance at factors relating to use of research in industry follows below.

3. *Factors Relating to Innovation.* At present it is difficult to identify responsibility for nonutilization of evaluative research in a clear and specific manner. Anecdotal reports such as those on Provo and on the DCDC instrument for selection for work release are illuminating, but they do not give us general principles. Consequently, to avoid the simplistic controversies in which researchers blame administrators and administrators indict researchers, it may be useful to explore in more objective manner the barrier (and gates) to research utilization. This may suggest, if not principles, some temporary benchmarks.

Utterback has examined extensively the literature on innovation in industry, and he makes a number of comments that appear relevant here.[7] He looks at environmental factors, sources of ideas, internal characteristics of organizations, the roles of basic research, and the diffusion of innovation. He defines innovation as "technology actually being used or applied for the first time." [8]

The primary influence on innovation in industry is "market factors." These he sees as accounting for "60 to 80 percent of innovations in a large number of fields." Next in importance are "scientific and technological advances and opportunities." Other factors are "expanding markets," and "rising costs of inputs." Some innovations are aimed at reducing the use of the more expensive inputs.

Ideas for innovation appear to originate in discussions about needs. These are followed by searches for the technical means to meet the needs. Basic research does not appear to be significant as a direct source of innovation, Utterback points out. Its primary role is in the production of knowledge. Thereafter, it enters the process of innovation by a roundabout route, through education, with a time lag of varying duration.

The importance of technology transfer is shown by the fact that no organization can respond to more than a fraction of its needs or problems that require innovative research. As a consequence, most of the ideas successfully developed and implemented by an organization come from outside that organization. Even in the case of "research-rich" du Pont, 14 of 25 major product and process innovations within that company originated entirely outside the company.

Several aspects of Utterback's analysis are of interest here. One is his identification of the "market place" as the primary influence on innovation or technology utilization. Another is the stress on organizational needs as a generator of ideas that lead to innovation. Another is the apparent insignificance of basic research as a direct influence on innovation. Another is the importance of the process of innovation transfer. The meaning of all these matters for innovation or research utilization in corrections is not clear although each seems highly relevant.

The emphasis on market factors rather than on "scientific technological" factors suggests that practical relevance may outweigh theory and method in determining the likelihood of innovation. The relative importance of organizational needs and basic research for direct innovation is not surprising yet at the same time instructive. Also instructive is the extent to which even major organizations must depend on ideas brought in from the outside to provide solutions to problems.

These are speculative inferences and ambiguous conclusions. It would be desirable to have a study similar to Utterback's, focusing on correctional operations, correctional innovations, and correctional evaluations. This would reduce considerably the need for conjecture as to how research products are channeled toward utilization or rejection in corrections.

B. Improving the Utilization Rate

Since we do not have the kinds of information on correctional agency innovation that Utterback pulled together on industrial innovation, we are not in a position to make strong recommendations on how to improve utilization. We can, nevertheless, offer some tentative suggestions.

1. *Focusing on Areas of High Priority.* The importance of this matter has already been attested in the discussion in Chapter 5, "Research Needs and Research Priorities." That discussion need not be recapitulated here, other than to point out that community and organizational needs are the starting points and probability of payoff is a modifying factor. Innovations that meet basic social requirements and are cost-effective should find a ready market.

2. *Relevance of Research.* Although high-priority status comes first, relevance is important. Relevant and significant findings must compete with high-priority findings for attention and for budgetary support, and they may need to compete also with organizational and attitudinal considerations. However, the appearance of relevance should improve the chances that a potential innovation will become an actual one.

3. *Exposure.* Some means of insuring that wide distribution is given to findings should enhance probability of adoption. It is evident from the Provo and DCDC examples that potential innovations may need to find support in other locations than the ones that generated the ideas. If, as Utterback states. most ideas that lead to innovations come from agencies other than the applier, the channels for dissemination should be kept open. This implies better reporting of findings, wide distribution of reports, effective information center services, training centers and institutes, and similar facilitating aids.

4. *Organizing Potentially Useful Findings.* Some potentially useful findings make their impact after considerable delay. *Probation in California,* with its recommendation of a probation subsidy, was released in 1957. It failed to achieve its goal of an Assembly act enabling a subsidy to the counties in 1957, but the survey report was not forgotten. It served as a model for an updating survey in 1964, and with an appealing formula for subsidy payments and a detailed action plan, the Subsidy Act was finally approved in 1965.

Many potentially useful ideas do not get a second chance, or the interval between appearance and utilization stretches out unreasonably, or the power of some ideas is diminished because related materials are not brought together in a way that capitalizes on their convergence. The latter problem is sometimes described as a failure to make research accumulative.

Making research accumulative is difficult even within a given research unit. Across units the task becomes almost overwhelming. As an interim solution to this problem, we see periodic releases of project abstracts that report all matter submitted; literature reviews that are selective, highly abbreviated, and strongly reflective of the author's training and values; computerized information services that are flexible and easily updated versions of the periodic abstract publications; and specialized groups that try to accumulate reference materials on particular programs, such as the Newgate Resource Center.[8]

The summary materials that come out of these services conceal more than they reveal, and there is some justification for calling the services graveyards of useful ideas. They have as yet found no satisfactory solution to the problem of identifying and flagging items of special significance, or of bringing coherence to materials that are related but not quite accumulative. The administrator or researcher who knows what he wants may use some of these services as one clue to potentially useful literature and follow up with an extended search and analysis of the original literature—if it can be assembled.

Glaser has suggested a "propositional inventory" as a solution to some of the problems that are evident here.[9] Such an inventory would provide a clearer framework for organizing hypotheses and tested generalizations than has hitherto been achieved. However, this scheme is more likely to appeal to the detail-conscious researcher than to the administrator with his need for more general perspectives. It also appears more useful in an era of stability than in an age of rapid change.

For the present, we may need to continue awhile the procedure followed by the National Pretrial Intervention Service Center, which recently assembled laboriously all known evaluative reports on pretrial diversion programs in order to screen out what might be useful for policy makers contemplating action in this area.[10] Both the assembly process and the screening were difficult experiences that fully support the argument that present-day dissemination procedures may be one of the foremost barrier to research utilization.

5. *Linking Research to Planning and Development.* In Chapter 3, it was noted that research with impact seemed to be the work of researchers who also made recommendations and wrote proposals or plans for action. This makes researchers promoters of ideas that develop from their findings. If the findings have utility and the researchers are capable, the procedure should increase the rate of utilization of research. It is of interest that Utter-

back reported that innovations in industry appeared to require the least time "when the inventor himself attempts to innovate than when he is content merely to reveal a general concept." [11]

Since research and planning are capabilities not likely to be found in the same individual, it may be advantageous to bring them together by establishing planning units within research divisions. We observe this occurring occasionally in corrections. The CYA research division, after about fifteen years, has recently taken on a new identity as the division of research and development.

6. *Utilization of Research at Place of Origin.* A sound proposal for action, arising from just-completed research, should be most convincing in the agency where the evaluation was done. The relevant information is most complete and motivated persons are already at hand. This sequence has been enacted many times. When the Preston Impact Study was completed, the researcher presented the proposal for a community versus institutional treatment experiment with conviction, and executive staff appeared to require no great amount of persuasion. The impact study had no direct influence elsewhere since it was not published and disseminated. Had it been disseminated, it probably would have had less influence on other youth agency staffs since its success in the California Youth Authority appeared related to factors unique to the Authority.

Project Crossroads, carried out in the courts of the District of Columbia, was readily implemented as a regular procedure by the same courts when evaluative staff reported the project to be a behavioral and economic success. The DCDC narcotic-involved offenders proposal experienced easy going through the District of Columbia and Congressional hearings because the DCDC researcher-planners had preceded with extensive staff work and provided ample documentation at the several hearings.

It appears, then, that utilization of findings has its best likelihood of occurring at the site of discovery. This likelihood might be expected to drop as the findings move away from home. Accessory information tends not to go along, the context of factors at other sites is perhaps less suitable, and the persons who would be the most able to promote the new idea are not present at other possible sites.

C. Discussion

Utilization of research findings is properly a matter of concern in view of the need for productive innovation in corrections, the high cost of research, and the low rate at which research products are paying off at the present time.

Researchers claim that their efforts are ignored, but it is difficult to take this claim seriously. It appears that the primary reasons for nonutilization of research are poor quality, irrelevance, fragmentation, incomprehensibility, nonaccumulativeness and inaccessibility. We need to add to this the fact that research staff are increasingly becoming outsiders as the demand for evaluation grows, and outsiders face special problems in comprehending what is relevant, in carrying through on recommendations and plans, and in communicating effectively with the best prospects for utilization of particular pieces of research. It may also be suggested that the main problem in nonutilization apparently does not arise from operating staff; historically operating staff has been notorious for its readiness to innovate. How else explain the long parade of operations-initiated programs over the past two or three decades?

There is, it should be added, one way in which administrators may be accountable for much of the nonutilization problem. To the extent that they contribute to the poor quality of inaccessibility or fragmentation of research by not providing support or effective coordination, they must share responsibility with researchers.

Since utilization is desirable on the grounds of need to improve corrections, more attention to the manner in which utilization occurs is urgent. In view of the scarcity of knowledge in this area, further studies of the conditions and mechanisms of utilization are needed. Just as we lack adequate studies of the conditions and levels of research impact, we also lack enough information on the conditions and processes of utilization, which bring the potential of research to the point where it might be able to have an impact.

The key to utilization seems to lie in the relevance of the research to the principal tasks or functions of the correctional agency. To be insured of utilization, evaluation results need to relate to offender change, system change, and perhaps cost reduction, since corrections seems to stand out as one of the most inefficient of all social agencies. All these are important goals of correctional administrators.

To the extent that any of these goals prove elusive, as is alleged to be the case with offender change, it may be strategic to increase emphasis on the other goals. The apparent roadblock on offender change may be only temporary, and improved research strategies may discover ways of changing offenders that are major improvements over present experience. The possibility of such discoveries with juvenile offenders has already been demonstrated by Empey, Warren and Palmer, and confirmed by Speer and Shireman *et al.*[12] Whether improved research will bring similar demonstrations of ability to change adult offenders is a matter for further exploration.

Further steps toward increasing the rate of research utilization include the expansion of in-house research capabilities, increasing the attractiveness of careers in correctional research by raising monetary and psychic rewards, recruiting researchers who are able to plan and develop as well as to describe and analyze, encouraging open and experimental attitudes on the part of administrators and operating staff, devising procedures for wider and more usable dissemination of research findings, establishing methods for identifying and focusing attention on significant but unused findings, and discovering ways of coping with or adjusting for the nonaccumulative character of bodies of evaluative materials in selected subject-matter areas.

REFERENCES AND NOTES

1. Carol Weiss, *Evaluating Action Programs: Readings in Social Action and Education*. Boston: Allyn and Bacon, 1972, p. 319.
2. Edith E. Flynn, cited in "Connecticut is going coed," *New York Times*, 27 May 1973.
3. Lamar T. Empey and Maynard L. Erickson, *The Provo Experiment*. Boston: Heath, 1972.
4. Daniel Glaser, *Routinizing Evaluation*. Rockville, Md.: National Institute of Mental Health, 1973, pp. 163–64.
5. Fair, Isaac, Inc., *Development of aScoring System to Predict Success on Work Release*. District of Columbia: D.C. Department of Corrections, 1971.
6. Stuart Adams, Wendell Plair and John D. Spevacek, *The Use of a Predictive Scoring System to Improve Selection for the DCDC Work Release Program*. District of Columbia: D.C. Department of Corrections, Research Memorandum 71–7, November 1971.
7. James H. Utterback, "Innovation in industry and the diffusion of technology," *Science*, 183 (15 February 1974), 620–26.
8. Ibid, p. 620.
9. Daniel Glaser, *op. cit.*, pp. 177–82.
10. Neal Miller, *Evaluation of Research on Pretrial Diversion*. District of Columbia: National Pretrial Intervention Service Center, American Bar Association, 1974, p. 41.
11. James H. Utterback, *op. cit.*, p. 622.
12. Lamar T. Empey and Maynard L. Erickson, *op. cit.*; Marguerite O. Warren, "The Community treatment project," in Norman Johnston, Leonard Savitz and Marvin E. Wolfgang, (eds.), *The Sociology of Punishment and Correction*, Second Edition, New York: Wiley, 1970, pp. 671–83.; Ted B. Palmer, *The Community Treatment Project in Perspective*, Sacramento: California Youth Authority, 1973; Charles H. Shireman *et al*, "Findings from experiments in treatment in the correctional institution," *Social Service Review*, March 1972, pp. 38–59; David C. Speer, "The role of the crisis intervention model in the rehabilitation of criminal offenders." Buffalo: Erie County Suicide Prevention and Crisis Service, 1972, unpublished paper.

PART THREE.

RESEARCH METHODS AND STRATEGIES

CHAPTER 8. DEFINITIONS, CRITERIA, METHODS AND MODELS

The basic task of the evaluator is to ascertain whether some event or situation, either real or conceptualized, is better than another. The event may be anything from an offender's action to an agency's decision. The situation may range from an offender's record of performance to the manner in which the criminal justice system relates to the larger social system. The task is essentially a process of comparison.

The evaluator may set himself a second task: explaining why a particular event or situation is better than another. Some research methods, such as the controlled experiment, are believed to be especially useful in this task. However, the great majority of correctional evaluations up to this time have been non-experimental. We may assume, therefore, that correctional evaluators have been more concerned with making comparisons than with explaining precisely the reasons for the differences they have found.

In setting directions for the evaluator, we need to examine some basic concepts and procedures. In this chapter we look at definitions of evaluative research, methods of making comparisons, standards and criteria, and methods and models of evaluation.

A. Definition of Evaluative Research

As used in this Guide, evaluation is considered to be an aspect of scientific method. It is applied research; it may be qualitative or judgmental, although it is more likely to be quantitative. It is largely descriptive, although it occasionally includes inferential or explanatory elements. The latter occurs in those rare cases where evaluation attempts to determine the causes of observed differences.

As a general and brief definition of evaluation, we can state that it is a procedure for ascertaining whether an event, process or situation is better than another. This definition does not adequately cover the case of the controlled experiment, hence, it may be useful to consider a more elaborate definition. Suchman quotes from the American Public Health Journal (February 1960) to offer the following: [1]

> "Evaluation: The process of determining the value or amount of success in achieving a predetermined objective. It includes at least four steps:
> - formulation of the objective,
> - identification of the proper criteria to be used in measuring success,
> - determination and explanation of the degree of success, (and)
> - recommendations for further program activity"

This definition is of limited utility, and it could be made more concise in logic. Its utility is limited because it focuses on the controlled experiment, which is only a small part of the total evaluation process. Furthermore, the experiment appears to be of significance primarily in minor rather than major kinds of decisions in correctional planning and administration. For these two reasons, the definition appears to describe neither the typical nor the important kinds of correctional evaluation. With reference to its logic, step four is perhaps more a social action principle than a step in evaluation procedure.

If we were to expand the first definition above to include explanation and measurement, we would have the following:

> "Evaluation is a procedure for ascertaining whether an event, process, or situation (real or conceptualized) is better than another. The procedure may include steps for measuring 'how much better' and for explaining the reasons for the difference."

This definition includes the notion of comparing events, processes or situations with concepts. This is done to cover those kinds of evaluations where comparison is made between real occurrences and standards or goals. In the survey of probation in California, for example, existing probation practices were compared with concepts of practice as set forth in documented professional standards or as spoken by members of the Special Study Commission that ordered the survey of county practices.

B. Making Comparisons in Evaluation

Since the main task in evaluation is that of making comparisons, it is important to consider the numerous ways in which comparisons may be made. Following are some of the better known:

1. *Real Condition versus Ideal.* In this procedure, existing conditions are compared with some perceived ideal or some "superior" concept. The Preston Impact Study traced the development of juvenile delinquents in the PSI facility and compared the resultant images with a sense of how the juveniles should have or might have developed in a "good" rehabilitating or reintegrating environment. The comparative judgments were wholly qualitative.

2. *Real Condition versus Published or Official Standards.* The survey of probation in California examined the functioning of probation departments against professionally promulgated standards for probation systems and procedures. These standards were augmented to some extent by the deliberations of the Special Study Commission. The comparative "measurements" that were set forth in the survey report were both quantitative and qualitative.

3. *"Before" Status versus "After" Status.* Persons, processes, and structures or systems may be compared at earlier and later times. Differences noted between the two times may be described either quantitatively or qualitatively. Special conditions or experiences that were introduced during the interval may be considered in relation to the changes, if any, that were observed. The comparison may be in either qualitative or quantitative terms.

4. *Program Persons versus Non-Program Persons.* Indicators of the performance of individuals who have been involved in a program may be compared with those of other individuals who have not been involved. The "other individuals" may be selected in several ways:

- All other persons in the prison (or other relevant population),
- A random sample of all prisoners,
- All prisoners who left the prison in a previous year,
- A "norm" group, on whom performance data are already available from other occasions

The comparisons may be either qualitative or quantitative. The comparisons in these cases will be rough and errors in interpretation will be easily possible. If more is desired, more exact comparison procedures may be used.

5. *Program Persons versus "Similar" Non-Program Person.* This procedure compares the performances of program participants with those of persons who are identified as "similar" to the participants by use of special selection procedures and verification techniques. These similar individuals comprise a "comparison" group, which is presumed to be a reasonable approximation of the participant group on essential characteristics. A comparison of this kind is a "quasi-experiment."

6. *Program Persons versus "very Similar" Non-Program Persons.* By establishing a group or "pool" of eligibles and randomizing eligibles to experimental and control statuses, a more precise basis for measurement of difference can be established. This procedure is the basis for the controlled experimental design, in which controls and experimentals are processed through essentially the same correctional experiences except that an "experimental" treatment is added for the program persons. In comparisons based on this procedure, it is possible, in principle, to make inferences as to whether the differences observed are caused by the experimental treatment.

7. *Real Outcomes versus Expected Outcomes.* Expected outcomes for a general population of offenders or various subpopulations can be established by analyzing pre-program and in-program experiences in relation to post-program performance. Techniques such as linear discriminant analysis, configuration analysis and direct search in mathematical space will identify meaningful predictor variables and provide scoring weights for the variables. Scoring tables associated with expected outcomes can be constructed. These tables permit the calculation of expected outcomes of future groups from the population on which the instrument was constructed. If a subgroup is then subjected to a new program or experience in addition to the usual or traditional program, the difference between expected and actual outcomes may be interpreted as the result of the new experience.

8. *Comparison of Agency Reactions to Participant and Behaviors.* This procedure focuses not on the behaviors of persons but on agency actions such as parole revocation, arrest, conviction, and incarceration. Instead of seeking psychological test scores or descriptions of overt behavior of individuals, the researchers seeks data on agency actions against the individual. Agency actions and agency records are taken as reasonably useful indicators of internal and overt behaviors of offenders.

9. *Participant Costs versus Control Costs.* Comparisons of participants and control or comparison of group members on costs of management during specified periods, such as equal lengths of time in an institution or equal lengths of stay in the community after release, may disclose differences in the rates at which the two groups generate new correctional costs. The cost differences may be interpreted as indicators of differences in the groups' performances.

10. *Comparing Participants and Controls on Both Benefits and Costs.* Participants and controls may be compared on both costs and monetary benefits from increased earnings, welfare costs avoided, police and court costs averted and reimprisonment costs averted) during equivalent periods after release. The cost-benefit analysis made from these data may yield a benefit/cost ratio whose function is to show whether ascertainable benefits outweigh costs.

These types of comparison vary widely on several characteristics. Some are easily and quickly made. Some yield more objective and more accurate data than others. Some permit conclusions about effectiveness only; others permit conclusions about both effectiveness in behavioral terms and efficiency in monetary terms. Some permit specific judgments about cause and effect.

Which of these comparisons will be used in particular cases depends upon the practical objectives of operating staff and the research objectives of the evaluator. If quick action based on relatively uncertain data is considered adequate by the administrator, comparison types 1, 2, 3 and 4 may be used. If efficiency in terms of the relation of outcomes to input is the focus of interest, comparison 10 will be used. Researcher-dominated strategies may center on comparisons 5, 6, 7, and 8, which emphasize complex methodology and greater certainty of results. Administrator-dominated strategies may focus on the two ends of the list.

C. Administrator Comparisons versus Researcher Comparisons

The improvement of correctional agencies over the past decades has been primarily an administrator-dominated process. The influx of significant innovations, especially since the middle of the century, has resulted more from the judgmental evaluations of administrators than from the empirical and theoretical evaluations of researchers. We have, as a consequence, more or less rapidly developing movements such as halfway houses, prison colleges, work release, study release, family visiting, week-end furlough, family counseling, and pre-trial diversion.

Nothing comparable has developed in corrections as a result of concepts originated by researchers, tested, developed, and pronounced ready for operational use. This suggests that administrators are capable of some rationalizing of the correctional system, with a moderate amount of pressure from and assistance by the humanitarian and managerial traditions in this society.[2]

This proposition has at least two implications for the researcher. One is that the administrator is able to act on the basis of limited or uncertain information. He has been acting thus for decades, frequently to good effect, and it has yet to be demonstrated that researchers can improve markedly on this process. Consequently, when an administrator asks for quick information, it is reasonable to give him quick even though uncertain information. He has made good use of such information before, and it is to be presumed that he can continue to do so in the future.

Another implication is that when the researcher has uncertain but possibly promising information, it is appropriate for the researcher to discuss such information with the administrator. The discussion can take several directions. "Here is some interesting information, and it can either be developed further or it can be acted on as a trial. Here are some action alternatives. Here are some research alternatives." If the administrator chooses an action alternative on fairly limited information, in a manner that allows for a pull-back if necessary, this is not a retreat to an inferior procedure. It is an advance over what the administrator made use of in his previous modifications of the system.

The researcher may find this perspective useful in seeking an appropriate role for himself in the rationalizing process. Thus far research has contributed little to the process, while administrators, using judgmental procedures, have contributed much. When researchers learn better how to make useful inputs into a fast-moving and flexible decision-making process, rationalization may proceed at a more rapid pace. This transition can be made more readily if the researcher develops facility in operating with both uncertain and certain data rather than looking only for increasingly certain knowledge.

D. Conceptual Models, Standards, and Performance Criteria

Comparisons or measurements in evaluation proceed from baselines such as concepts, written or

spoken standards, and indicators of performance. Administrators and researchers may form concepts—conceptual models—of innovative programs in brainstorming processes, letting imagination run free. Or they may deliberate about an existing program and come up with an idea for an alternative which they think will be an improvement. The researcher who periodically interviewed fifteen youths as they were processed through CYA's Northern Reception-Center Clinic and Preston School of Industry arrived at a conceptual model of a better alternative—treatment in the community—rather slowly.

Administrators are accustomed to this process of going from conceptual models to innovative operating programs. Traditionally, they have arrived at their innovations conceptually and thereafter have evaluated them judgmentally.

When standards of procedure or organization are developed, as happens commonly in professional or service organizations, these standards provide a more objective basis for judging or comparing procedures and structures. To the extent that written standards derive from conceptual models, the use of standards as bases for evaluating programs is not markedly different from the use of conceptual models. One practical difference is that standards carry some authority and the innovation that is proposed on the basis of the standards is easier to defend.

Indicators of performance become important as bases for comparison or measurement when the focus of interest is offender change rather than structural or procedural change. Concepts of good behavior and standards of good performance can be visualized and stated and a released offender can be judged against these concepts and standards.

In practice, the offender-judging process fixes upon a number of aspects of performance such as obeying the law, cooperating with officials, and meeting one's responsibilities as a citizen. Eventually, these areas of performance come to be represented by "criteria of performance," which are accessible and measurable behaviors or indicators of behavior.

The success of evaluation of offender change is determined to a large extent by whether valid criteria are identified and properly used. Following are some of the criteria that have been employed in correctional evaluation, listed without regard to quality or usefulness:

- Observed behaviors and attitudes
- Personality inventory scores
- Attitudinal inventory scores
- Probation or parole agent ratings
- Revocations, technical or new offense
- Arrests
- Time till rearrest
- Type (or seriousness) of offense
- Conviction
- Length of sentence
- Time served
- Costs of apprehension; court costs
- Supervision costs (probation or parole)
- Incarceration costs
- Criminal career costs
- Personal or social adjustment
- Job status and earnings improvement
- Benefits from pretrial diversion
- Benefits from recidivism averted
- Benefits from welfare costs averted
- Benefits from health costs averted
- Educational status improvement

Not many of these criteria are in general use in evaluation. Some are unreliable, some require costly data collection procedures, some are not valid measures of behavioral change, and some are not highly regarded by persons who control budgets.

Following are several of the performance criteria that have been widely used or are beginning to be used extensively, with comments on some of their essential characteristics:

1. *Arrests.* Arrests have both good and poor characteristics as criteria. They are official, they get into the record where they can be found years later, and they are usually the first "hard" evidence that the offender is not performing well in the community. On the other hand, many persons who are arrested do not go to trial, or are not convicted. Furthermore, arrests tell us little or nothing about the quality of the alleged offense—whether it was a traffic violation or a heinous murder. Finally, arrests and the presumably more valid criterion, convictions, may give discrepant indications, as in the case of the Lorton Prison College Project evaluation (Item 4, below).

2. *Type of Offense.* Because of the weakness of arrest as a measure of performance, it is important to have supplementary information such as "offense charged." If one wants to judge the quality of performance of a group of ex-prisoners at three to six months after release, arrests by offense charged are a better indicator than arrests alone. Collection of the offense data may take appreciably more time in some systems, but they make clear the seriousness of the failure in performance. They are also a necessity when a seriousness-of-offense scale is part of the research design.

3. *Time till Arrest.* Since one of the goals of offender rehabilitation is to reduce further involvement with the law, effective rehabilitation should reduce frequency of rearrest and increase time till next arrest. Consequently, it is useful to add time till arrest to the criterion "package." We say "package" because it is important when possible to use more than one criterion. By choosing a balanced group of performance indicators, one can compensate the weaknesses in some by the strengths in others. Also, the use of several criteria provides a broader and probably more valid picture of performance levels and performance differences.

4. *Conviction.* This criterion is one of the more plausible indicators of performance in the community. Adjudication has presumably screened out the more "invalid" cases resulting from questionable arrests. Some "valid" cases may have been screened out also, so that conviction, too, is an imperfect criterion, although it is probably more valid than arrest for most measurement purposes.

Weeks or months may pass between arrest and conviction, hence, the latter is a delayed measure of performance. The researcher who wants an indication of performance within a few months may have to use arrests alone, or arrests and charges, or include whatever convictions have occurred by the cut-off date for the measurement. This will leave some cases with charges pending, which creates problems in interpretation.

This problem is illustrated by a quasi-experimental evaluation of the Lorton (DCDC) Prison College program in 1972. At six months time in the community, participant parolees and comparison parolees each showed arrest rates of about 30 percent. However, conviction rates were 7 percent for the participants and 19 percent for the comparison group. There were 12 pending cases for participants and 3 for comparison group members.[3]

At 12 months out, the arrests were still even (30.5 versus 20.4 percent), but the pending cases were now cleared, and the conviction rates were 18 percent versus 25 percent.[4] These data are of interest for at least three reasons: first, they show different indications in the same population by two "hard" criteria, with arrests showing no difference and convictions showing appreciable differences; second, they suggest that arrests are less valid than convictions as a criterion; third, they disclose that the difference indicated by convictions shifts disproportionately from six months out to twelve months out, and an interpretation at six months that ignored pending cases could have been very misleading.

5. *Length of Time in Lock-up.* This criterion has about the same relation to conviction that seriousness of offense bears to arrest. Persons who spend three years locked up presumably have offended more seriously than those who spend one year. It is thus useful to record time spent in reconfinement as well as the fact of confinement when evaluating programs that permit lengthy follow-ups.

One of the desirable features of this criterion is that it provides an indication of the quality of performance prior to reincarceration. Another is that it permits "costing" performance by translating months served back in lock-up into the costs of that lock-up when the unit costs of prison maintenance are known.

6. *Costs of Correctional Treatment.* Costs of initial treatment and new correctional costs after release from treatment are useful measures of performance. One of these is an input measure; the other is outcome. Both are important when economic efficiency becomes a focus of concern in evaluation. The cost criterion has come increasingly into prominence in corrections in recent years. It appears likely to become much more important in the future as the techniques of benefit-cost analysis become better known.

7. *Benefits.* Like costs, benefits have become an increasingly important criterion in evaluation. Among the fifteen diversion program evaluations referred to earlier (p. 25), costs and benefits were included as part of the research design in every case. In about half the studies data proved difficult to obtain, so only half the studies reported benefit/cost ratios or otherwise compared costs and benefits in their final drafts.

Benefits include both cost reductions and earnings or savings augmented. In pretrial diversion, there are diversion benefits, earnings benefits, and recidivism benefits to be calculated. The diversion benefits are the averted costs of court appearances, police handling, trial costs, and incarceration when a participant is excused from adjudication because of successful performance in the program. Earnings benefits result partly from the higher rate of employment during the program period and some retained superiority in employment and earnings rates after the program period. Recidivism benefits accrue because the program participants generally show lower rearrest and reincarceration rates. In some projects,

it can be shown that additional benefits accrue from various other sources, as when narcotic offenders incur lower health costs after treatment, and families show diminished need for welfare payments when probation or work release are used as dispositions.

E. Methods and Models of Evaluation

Researchers speak of research designs, evaluation methods, and evaluation models. These concepts will appear numerous times in the following chapters, and it is pertinent to comment on them here. Our focus will be on research methods and evaluation models.

1. *Research Methods.* A research method is, generally, a set of procedures for a systematic and objective inquiry into an area of experience to increase knowledge or to facilitate problem solving. The variety of methods is great, and to avoid some detail we will use broad categories. We can talk first about traditional methods, such as experiments, quasi-experiment, and non-experimental methods. We can also refer to contemporary methods such as benefit-cost analysis, operations research, system analysis, and simulation. Neither of these sets of terms falls into a neat classification scheme. Some persons might prefer to use the term "technique" instead of "method" for some of our categories; others might regard some of our methods as clusters of methods rather than a single method. These are issues that need not be taken up here.

Six of these methods or groups of methods will be discussed in Chapters 9 through 14.

F. Evaluation Models

There is widespread reference to "evaluation models" in corrections, and for purpose of acquiring perspective, it is worthwhile going briefly into this concept. One way of describing evaluation models is according to the principle by which they are named. Some are named after the method they employ, as "experimental model" or "benefit-cost" model. Others derive their names from some aspect of the system or its functions, and as "process" model, or "outcomes" model. Still others focus on the characteristics or the style of the evaluator, as the "apprenticeship" model, or the "advocacy" model. Finally, some focus on the ultimate goal of measurement, such as effectiveness or efficiency; on the broader aims of evaluation, such as innovation or development; or on global aspects of the model, such as adherence to academic or industrial concepts of assessment.

Generally, a model is a physical, conceptual, or mathematical representation of something. In evaluation, a model is a conceptual framework for a set of comparison or measurement procedures. Empey, in the Silverlake experiment, speaks of a "field experimental model," a plan for comparing the performance of delinquents in a new program with that of similar delinquents in an old program to ascertain whether the new program is better.[5] The experiment was done in the community—the "field"—hence, the name.

1. *Methodological Models.* We note that for the purpose of this Guide there are at least six methodological models. Each takes the name of the method employed in the evaluation. There is the *non-experimental* model, which is a catch-all that includes several dissimilar research methods. It is distinguished by the fact that it is neither a quasi-experiment nor an experiment. It may also be stipulated here that the non-experimental model refers primarily to traditional methods of evaluation. It does not include such contemporary methods as benefit-cost analysis, systems analysis, and simulation.

By the same reasoning, we have also under methodological models the *quasi-experimental*, the *experimental*, the *benefit-cost*, the *operations research*, the *systems analysis*, and the *simulation* model. Each of these will be discussed in detail in later chapters.

2. *Subject-Matter Models.* Some evaluation models are thought of in relation to aspects of the subject under study. Thus we have the *outcome* model, which focuses on results; the *system* model, which focuses on the overall operation and structure, including its relation to the environment; the *input-output* model, which is concerned with results in relation to effort and initial materials; the *process* model, which evaluates procedures, or ways in which results are being sought; and the *means-end* model, whose concern is with the extent to which the processes or structures that make results possible have been provided.

The outcome model is perhaps the best known of the subject-matter models. It seeks to ascertain how effective or efficient an agency or program is in attaining its goals. For the correctional administrator these goals may include reduction of recidivism, reduction of violence, and reduction of returns to prison; or reduction in new correctional costs after treatment or release; or an increase of social benefits over costs as a result of correction.

The outcome model may use various research methods to determine the extent to which goals have been attained, and particularly to discover which of two approaches to goal attainment is better. It may use fast or slow methods, rough or precise methods, and its accuracy or validity may vary from excellent to poor, depending on factors ranging from the ability of the researchers to the interest and support of operating staff.

The means-end model is an "outcome" model of a special kind. Its principal objective is not a result such as the reduction of recidivism; it has instead the immediate purpose of setting up procedures or structures which are means to other ends such as the reduction of recidivism. The evaluation of such activties involves inquiry into whether the proposed mean or instrument has in fact been developed and in accordance with specifications.

The system model is concerned with neither the attainment of objectives nor the construction of means to attain objectives. It focuses on the working characteristics and relations of the agency under evaluation. This examination is concerned not only with how the entire system works but also with what would make it work better, according to some perceived standards or expectations.

The system model is more appropriate for use in the evaluation of a multifunctional entity than a single-function project or program. As seen by Etzioni, users of the system model are concerned with several kinds of activities: the achievement of goals and subgoals; the coordination of organizational or system subunits; the acquisition and maintenance of necessary resources; and the adaptation of the organization to its external environment and to its own internal demands.[6] By conceptualizing the possibility of various levels and combinations of these several activities, the system modeler can ask to what degree the organization can achieve its objectives under a particular set of conditions.

It is obvious that the system model views the subject of evaluation in large outlines. For the correctional evaluator, this model has its ultimate application in assessing the entire correctional agency. For the criminal justice evaluator the ultimate application includes the police, courts, corrections, the penal code, and related entities such as the public defender's office and the bail agency.

There is an even larger system, which includes the agencies of socialization—the family, the neighborhood and the school—as well as the criminal justice system and other social systems. The task of applying the system evaluation model to this "super-system" is as yet only vaguely conceptualized. At some future date, however, the maximum payoff in criminal justice system evaluation will come when it is possible to examine, with adequate theoretical and methodological tools, the operation of the socialization and criminal justice subsystems within the larger social system.

A curb-stone judgment on the importance of the larger system model was delivered recently by Louis Nelson, warden of San Quentin prison. Asked whether prisoners can be reformed, Nelson replied, in substance, that "If we can correct the home; if we can correct the schools; if we can straighten out the courts, and the police department; and if we can sweep out all the dark corners of society; then we can talk about reforming the prisoner."[7]

3. *Actor-Oriented Models.* The actor-oriented or researcher-oriented models include the *apprenticeship model* and the *advocacy* model. There are other possibilities, including the *traditional model,* in which professional researchers using the better-known methods of research carry out evaluations that provide knowledge but avoid recommendations.

The apprenticeship model focuses less on methodology and subject matter than on the process of developing and employing capable evaluators in corrections. The rationale for this method arises from the limited results obtained by correctional agencies when they depend upon outside researchers. The research team comes into the agency to perform a study, often promising much at the beginning, communicating imperfectly with administrative and operations staff, eventually delivering a research report that is frequently difficult to understand, and that contains only a small part of what the researchers learned. The researchers then depart, carrying much learning in their heads—learning which will gradually be lost.

The apprenticeship model would retain more of the learning from a study, keeping it in the agency where the evaluation was done. It would do this by having evaluation projects carried out by the normal administrative and operating staff of the agency. A researcher from outside the agency would teach or guide these staff until they achieved some mastery of techniques and strategies. As Emrich puts it:

> "The guide works more by asking the right questions than by giving the right answers. He is tolerant of research approaches that are not optimum for the problem, since agency personnel must frequently start

with fairly elementary research techniques and master them before they go on to more sophisticated techniques. In general, the guide believes that a well-done, but rather simple-minded solution to a research or evaluation problem, is much more likely to have impact on agency policy than a highly sophisticated solution which the agency personnel have almost no likelihood of understanding.

"The guide comes into the situation with a reportoire of approaches that are intended to get the agency involved in utilizing research and evaluation to help bring about major improvements. To the extent possible, the guide assists the agency in setting priorities and carrying out a sequence of projects which addresses the agency's problems according to these priorities. As each project is undertaken, the guide primarily with questions and suggestions, stimulates the agency personnel to move systematically through the stages of solving the problem.[8]

A variation of the apprenticeship model has been tried, with considerable success, in the California Department of Corrections and in the Los Angeles County Probation Department before the creation of the present research units in those agencies. Research of considerable importance was undertaken under this model, notably the SIPU and PICO experiments in the Department of Corrections.[9] The level of achievement was heightened, undoubtedly, because some of the participating "operational personnel" were treatment staff members with graduate degrees. These individuals found the research role a gratifying one and, with assistance from university statisticians and behavioral scientists, were able to inaugurate projects of major significance.

The advocacy model has been discussed by Guttentag (1972) as a possible answer to some of the weaknesses of the traditional model.[10] One of these weaknesses is that the agency is sometimes presented with a "highly sophisticated solution which the agency personnel have almost no likelihood of understanding," to use Emrich's phase. In the advocacy model, not only will the findings in the evaluation be presented but also, when indicated by the findings, recommendations for implementation will be made and possibly even plans for action will be formulated.

This model is visible in the work that led to the Probation Subsidy Act, the Community Treatment Project, and the Narcotics Treatment Administration. In each case, the researcher not only made an evaluation but wrote reports that included recommendations for action, and either alone, or in collaboration with others, wrote proposals for the carrying out of the action.[11]

In this model, the researcher, by being willing to promote actively a program or system that appears to be indicated by the findings of a study, reduces the "utilization gap" that exists between productive research and traditional operations in many agencies. The traditional evaluation model does little to close this gap.

The advocacy model calls for researchers with competence in research, understanding of the mission and structure of the correctional agency, and a capacity to define and promote desirable new alternatives. This is a large order, and there may be few persons available to fill the role. Furthermore, such persons can be effective primarily in agencies where there is an experimental administrator and a receptive operations staff. In other situations, the researcher might have to expend most of his energy in instructing or in "selling" operations staff on lines of action they are reluctant to follow.

4. *Goal Oriented Models.* The goal of evaluation is to discover superior *effectiveness* or *efficiency* in programs, processes or systems. An effectiveness oriented model is concerned primarily with whether a process or structure yields better behavior than another. An efficiency oriented model may focus on either *managerial efficiency* or on *cost returns efficiency*. In the former case, the efficiency model asks whether processes and structures are according to specifications, up to standards, and working as planned. In the latter case, it asks whether the new processes or structures are yielding high improvement in behavior per unit cost.

These three alternatives form a hierarchy of evaluation. At the first level, *managerial efficiency*, inspection of the operation will disclose whether there is a structure which is operating as planned. Are staff members aboard in all essential positions? Are their qualifications suitable? Are procedures in effect? Is there a record system? Are the records being properly maintained? Are procedures understood? Are they written down in accessible form? Is the process going smoothly? Are clients being processed in the prescribed manner? Are there procedures for effective resolution of problems?

Managerial efficiency may not accurately reflect outcomes. Client behavior may improve more in some managerially "inefficient" agencies than in "well-run" agencies. This means that managerial efficiency does not usually predict client outcome, which is the primary concern of the agency. Consequently, *effectiveness in terms of client behavior* generally takes priority over managerial efficiency as a goal of evaluation. This gives a second level in the evaluation hierarchy.

Effectiveness at low cost is better than equivalent effectiveness at higher costs. Consequently, evaluation eventually asks about both behavioral outcomes and cost outcomes. What is desired is that good behavior be maximized and that costs be minimized. Analyses that lead to this goal are *cost-effectiveness* studies, or benefit-cost studies. They are obviously of considerable importance to the administrator since they enable him to achieve the most rehabilitation for a given amount of resources. This is, in the present state of the art, the third and highest level in the evaluation hierarchy.

5. *Broad Strategy Models.* Exploration, innovation and adjustment are three kinds of activities in which evaluation is used, generally in quite different styles. *Exploration* is a search for leads or ideas, using non-experimental methods in critical or promising areas, with receptiveness to the possibility that "something may turn up."

Innovation is a firming up and testing of new ideas in the hope that they can be developed into better modes of meeting the objectives of the organization than previous modes. Innovation may use non-experimental methods, or a mixture of methods in sequence, as it moves from exploration to consolidation of an idea. It will usually organize its efforts in project form. When it has arrived at success, the project will become a program and future evaluations of the activity may be of a monitoring kind.

Adjustment is the process of making shifts in programs to effect improvements that are suggested by special evaluations or by gradual increases in information through observation or routine evaluation. Routinization of evaluation is important for established programs since it provides a relatively economical yet dependable means of watching over the effects of programs and accumulating the information needed for occasional refinements of procedures.

Innovation and routinization are contrasting activities in evaluation, each representing a fundamental need of correctional and other agencies. Innovation is necessary for creative change, and it becomes especially necessary at times when the agency is badly in need of improvement. Routinization is necessary for efficient management, but it tends to be difficult when the forces for innovation are working strongly. Over time, the history of corrections may be viewed as a swinging of the pendulum from innovation to routinization, with either of the activities diminishing somewhat when conditions favor the other.

The recent appearance of Glaser's *Routinizing Evaluation* highlights this interplay of the two tendencies.[12] Contemporary corrections is in ferment, if not in a state of revolution, and a call for the routinization of evaluation may appear like swimming against the tide. On the other hand, the correctional scene is crowded with new concepts, waiting to be adopted and developed into functioning programs. From this point of view, a call for routinization of evaluation may be extraordinarily good timing.

6. *Academic and Industrial Models.* One final perspective on models is the broadest of all: We can draw a contrast between an "academic" or "social science" model and an "industrial" or "policy science" model. The social science model consists of principles and methods that include a search for increasing certainty, concern for theory development, hypothesis testing, and high regard for controlled experimentation. The industrial or policy science model features operations research, systems analysis, simulation, and cost-benefit analysis. This model is somewhat atheoretical, although it is willing to make use of relevant theory; it is pragmatic, seeking practical results rather than knowledge for knowledge's sake.

Correctional evaluation has thus far been guided primarily by the social science model, but it is now beginning to borrow extensively from the industrial model. This trend will presumably continue more or less gradually until the advantages that inhere in the industrial model have been largely exploited by correctional evaluators.

REFERENCES AND NOTES

1. Edward A. Suchman, *Evaluative Research.* New York: Russell Sage Foundation, 1967, p. 28.
2. An extreme example of an administrator's attempt to rationalize the correctional system is David Fogel's proposal for a "justice model" of corrections. The model, whose custodial sector would be based on 300-person living units, would provide some self government, prison industries with prevailing wage rates,

accessibility to community services, civil legal assistance for inmates, independent ombudsmen, conjugal visiting, victim compensation and restitution, law libraries, no mail censorship, an extensive furlough program, open access to the press, a contract system for parole, and adversary and appeal procedures in parole revocation. *Criminal Justice Newsletter*, 5 (June 17, 1974), p. 3.

3. Colleen Barros, William Caldwell and Stuart Adams, *Effectiveness of the Lorton Prison College Project: Third Interim Report*. District of Columbia: D. C. Department of Corrections, Research Memorandum 72-4, May 1972, pp. 5-6.
4. Caroline Chalfont and Nancy Story, *The Lorton Prison College Project: Third-Year Final Report*. District of Columbia: D. C. Department of corrections, 1973, p. 30.
5. Lamar T. Empey and Steven G. Lubeck, *The Silverlake Experiment*. Chicago: Aldine, 1971, p. 9.
6. Amitai Etzioni, "Two Approaches to Organizational Analysis: A Critique and a Suggestion," in Herbert C. Schulberg, Alan Sheldon and Frank Baker (eds.), *Program Evaluation in the Health Fields*. New York: Behavioral Publications, 1969, 101–20, p. 103.
7. Louis Nelson, on the Merv Griffin Show (Show No. 498), "Prison Reform," Los Angeles, KTTV, 21 February 1974.
8. Robert L. Emrich, *A New Strategy for Public Policy Research*. Davis, Calif.: National Council on Crime and Delinquency, August 1971, pp. 5–6.
9. Of the SIPU studies, five are worthy of mention: *Special Intensive Parole Unit, Phase I, Fifteen-Man Caseload Study*. Sacramento: California Department of Corrections, November 1956; *Special Intensive Parole Unit, Phase II, Thirty-Man Caseload Study*. Sacramento: CDC, December 1958; Joan Havel and Elaine Sulka, *Special Intensive Parole Unit, Phase III*. Sacramento: CDC Research Division, Research Report No. 3, March 1962; Joan Havel, *Special Intensive Parole Unit, Phase IV (High Base Expectancy Study)*. Sacramento: CDC Research Division, Research Report No. 10, June 1963; Joan Havel, *Special Intensive Parole Unit, Phase IV (Parole Outcome Study)*. Sacramento: CDC Research Division, Research Report No. 13, September 1965.
   The principal PICO Project reports are five in number: Alvin Rudoff and Lawrence Bennett, *The Pilot Intensive Counseling Project, First Technical Report*. Sacramento: CDC, January 1958; Alvin Rudoff, *The Pilot Intensive Counseling Organization Project, Second Technical Report*. Sacramento: CDC, January 1959; Glenn Smith, *The Pilot Intensive Counseling Organization, First Special Treatment Report: The Utilization of Authority in Correctional Casework*. Sacramento: CDC, January 1961; Thomas L. Frazier, *The Pilot Intensive Counseling Organization, Second Special Treatment Report: A Casework Approach to The Treatment of Young Narcotic Users*. Sacramento: CDC, May 1961; Stuart Adams, *Effectiveness of Interview Therapy with Older Youth Authority Wards: An Interim Evaluation of the PICO Project*. Sacramento: California Youth Authority, Research Report No. 20, January 1961. See also Stuart Adams, "Interaction between individual interview therapy and treatment amenability in older Youth Authority wards," in *Inquiries Concerning Kinds of Treatment for Kinds of Delinquents: Monograph No. 2*, Sacramento: California Board of Corrections, July 1961, pp. 27–44. Reprinted as "The PICO Project," in Norman Johnston, Leonard Savitz and Marvin E. Wolfgang (eds.), *The Sociology of Punishment and Correction*. New York: Wiley, 1970, pp. 548–61.
10. Marcia Guttentag, "Models and methods in evaluation research," *Journal of the Theory of Social Behavior*, 1971, 175–95.
11. Two examples of the research report-action proposal sequence are the following: a) Stuart Adams, *The Preston Impact Study*, Sacramento: California Youth Authority, 1959, and Stuart Adams, *A Proposal for the Comparison of Institutional and Community Treatment for Juveniles. Sacramento:* CYA, 1959. The latter was elaborated in Stuart Adams and Marguerite Q. Grant, *A Demonstration Project: An Evaluation of Community-Located Treatment for Delinquents*. Sacramento: CYA, 1961. b) Stuart Adams, Dewey F. Meadows and Charles Reynolds, *Narcotic-Involved Inmates in the Department of Corrections*. District of Columbia: D. C. Department of Corrections, Research Report No. 12, February 1969, and Stuart Adams, *The Narcotic Addiction Rehabilitation Center: A Proposal for a Multi-Modal Community-Based Treatment Center for Addict Offenders*. District of Columbia: DCDC, April 1969. The first of these sequences led to the California Youth Authority's Community Treatment project; the second led to the District of Columbia's Narcotics Treatment Administration.
12. Daniel Glaser, *Routinizing Evaluation*. Rockville: National Institute of Mental Health, 1973.

CHAPTER 9. NON-EXPERIMENTAL EVALUATIONS

Non-experimental evaluations include a wide variety of research methods. The types most commonly used in corrections and criminal justice appear to be the *case study*, the *survey*, the *time series*, the *cohort analysis*, and the *before-after study*. Such studies can be either descriptive or evaluative. If they focus on presentation of a picture of what exists, they are descriptive. If they also make comparisons and indicate preferences that are related to policy formulation, they are in part evaluative.

There are several reasons for careful consideration of non-experimental studies. First, they are the most numerous of the several major categories of evaluative studies. Bailey reported that 22 percent of his sample of 100 correctional evaluations were experimental designs.[1] In Berkowitz' sample of 38 evaluative studies, about 13 percent were controlled experiments and another 13 percent were quasi-experiments.[2] In the first five years of its existence, the DCDC research division conducted about 50 studies, one of which was a controlled experiment and three of which were quasi-experiments.[3]

These examples suggest that non-experimental studies comprise at least 80 to 90 percent of the evaluative studies in corrections. These percentages would be different if we looked not at numbers of studies but at time invested or numbers of reports produced. For the present, however, numbers of studies are our only means of establishing a trend from the Bailey, Berkowitz and DCDC figures. On the basis of this trend, the non-experimental study is easily the most numerous.

A second reason for giving this type of study careful attention is that it seems to carry a heavier impact for corrections than experiments or quasi-experiments. This was suggested in Chapter 3, where it was noted that two surveys of county probation had had a more profound effect on court and correctional practice than the controlled experimental Community Treatment Project. Such evidence is sketchy, but it supports the belief that non-experimental studies have been the basis of the most important research-induced policy changes in corrections in the past two decades. And there is no compelling reason to believe that this condition will change in the near future. The non-experiment seems an especially useful instrument in times of rapid change, while the experiment loses value in such times.

There are other reasons for interest in non-experimental studies. They can be applied to poorly understood problems in ambiguous contexts—a situation that is not favorable for a controlled experiment. As a result, many of the more important problems in corrections are approached first by means of a non-experimental study, and only later, if at all, by an experimental study.

The non-experimental study appears more suited to executive decision-making styles and tempos, and its versatility gives it the lead in a variety of problem-posing situations. Before the experiment can be brought to bear, the important decisions have often been made and the center of interest is now new problems in new areas.

Non-experimental studies are usually quick of execution and generally inexpensive as compared with experiments. Also, they pose less of a threat or burden to operating staff, and they facilitate communication with practitioners since the concepts, techniques and manner of reporting are closer to common experience.

Some aspects of non-experimental studies are disadvantageous. Their value is determined to a large extent by the experience, judgment and objectivity of the researcher; improperly used, they may create more confusion than enlightenment. Their procedures lack standardization, their reliability is uncertain, and their interpretation is sometimes difficult. Many of these characteristics are more troublesome to researchers than to administrators. The latter are constantly faced with unreliable and uncertain data in their decision-making processes and they are more accustomed to acting on such information, though often with questionable effect.

A. Types and Examples of Non-Experimental Studies

Following are brief discussions of several of the

more common types of non-experimental studies. Some of the types, particularly the time-series and before-after studies, are occasionally grouped with quasi-experiments. They are shown here in applications or forms that are too uncontrolled to merit an experimental or quasi-experimental classification.

1. *The Case Study.* This type of study aims at a relatively intensive examination of a "case"—a person, group, place, event, program, or experience. It looks at the subject broadly or deeply or both, collecting and interpreting information in a variety of ways to learn what exists, what is happening, and possibly why. When it is purely descriptive, the case study examines its subject and reports in a non-judgmental manner. When it is evaluative, it brings in concepts, standards or studies of other cases for comparison and judgment.

Campbell and Stanley refer to the "one-shot case study" as the "weakest but most common evaluative design." [4] It should be noted that "weakest" as used here refers to standing in a hierarchy of study design characteristics. It does not refer to social policy impact, where the case study sometimes rates "strong" in its capability for influencing policy.

A recent example of the use of this design is the study, *Team Policing: Seven Case Studies*, recently released by the Police Foundation.[5] It reports findings in a two-year examination of the team approach to crime control by police.

The purpose of the case studies was to inquire whether team policing achieved the objectives of community and police better than conventional policing patterns. In brief, was it better to be "concerned, service-oriented and invisible" than "tough, enforcement-oriented and visible?" [6]

The concept of team policing has been influencing police department thinking since World War II, with isolated instances of application becoming visible in various localities. Following a recommendation by the President's Commission on Law Enforcement and the Administration of Justice in 1967, several American cities have been trying out the concept.

The team policing case study was undertaken in a number of cities simultaneously. Each case study described the site city and police department and noted the structure and operational experiences of the particular team program in that city. The studies were reportorial, not in-depth, and some subjective assessments were made by persons on the scene.

Efforts at a "more scientific evaluation" were made in some of the sites by the police departments, but in no case did the efforts at evaluation approach the controlled experimental method. And in no case did the team projects include an evaluation component to measure the amount of real crime and apparent crime change in the area.

The effect of the case-study findings on police administrator decisions at the end of project first phases was usually "quite small" for a number of reasons. First, the police chief claimed to know, intuitively, what the case-study staff reported to him. Second, the evaluators often lacked credibility, with the chief or the department or both. Third, the administrative decision as to whether to terminate the project or go on to a next phase was influenced by many "outside political factors," so that the evaluator's input, credible or not, was only one of several potential elements entering into the decision-making process.[7]

Since the completion of the first study by the team policing research staff, some cities have gone ahead with phase two projects, in some cases building in evaluation components and working toward more credible assessments of the team policing concepts. If the case study report has not impacted policy in any visible manner, there remain further opportunities for such studies to make impacts.

The Preston Impact Study was a simpler, smaller instance of a case study. It was a researcher-initiated follow-up of a group of youths passing through a CYA reception center and training school. Attention was centered on how the boys changed and why as they went through the process of institutionalization. The study continued over a period of a year, requiring three or four days a month for field work in the first two months, and in alternate months thereafter. Tape-recorded interviews collected information on several areas of clinic and training school experience, reflections on thoses experiences, and perceptions of the self and of the future.

The interview transcripts provided a basis for summarizing the developmental process, and the image thus created was largely negative. An alternative concept of developmental process gradually emerged, featuring lower levels of violence, reduction in peer pressure, a more diluted criminal lore, fewer anti-authority concepts, and milder patterns of exploitative behavior. This configuration was seen as possible primarily in a non-residential community treatment center. Consequently, the test of community versus institutional treatment was proposed to CYA executive staff.[8]

Unlike some of the team policing case studies, the Preston Impact Study had an immediate and direct impact upon the departmental administrator. He and his staff made a prompt decision to support the plan for an experimental test of the two treatment types. In this case, the researcher's credibility was high, even though the data presented to staff were qualitative and subjective. The administrator and staff had heard nothing new, except in the specific details. Intuitively, they were aware of the general content of the anecdotal materials and the conclusions that had been drawn. Now that a definite action was proposed, they were ready to support the recommendation. Hence, the Preston study had full impact and its recommendations were adopted even though the study had started as nothing more than an exploratory investigation undertaken by a curious researcher.

Quick Case Studies. Case study procedures vary as widely as the many subject matters they are concerned with. Some are loosely structured and produce highly subjective narrative reports. Others are formal, highly systematic, and seek to present structured, objective materials along with opinions and judgments.

One example of the latter type is described by a set of procedures developed by the Special Action Office for Drug Abuse Prevention. The method is described in SAODAP's publication, *Quick Evaluation Methodology*, as a "short, systematic framework for assessing a variety of drug abuse treatment programs." [9] It "enables an average treatment program to be evaluated by two people in two days, including one day spent on-site."

SAODAP's "quick evaluation methodology" is one more illustration of the fact that necessity is the mother of invention, even in research. With several hundred drug-addiction treatment programs now in existence and new ones being proposed every day, some means of making prompt assessments was needed to avert chaos and permit some semblance of managerial control.

The case study technique that was devised is at once speedy, comprehensive, and useful for the decision-maker. It is built around ten analytical criteria, several types of descriptive data, and a plan for subjective assessment by the evaluators. Given the documented results of the evaluation, administrators are in better position to make three kinds of decisions:

- Whether the program should continue to be funded
- Whether similar programs should be approved for start-up
- Whether technical assistance should be provided

The effect of the quick evaluation case studies on the decision-making process at SAODAP headquarters and their implications for existing and proposed drug treatment programs remain to be seen. Since it is an in-house effort, the problem of credibility may be minimal. What its ultimate effect on structure, process and outcomes will be is a matter for conjecture.

2. *Surveys.* Surveys are procedures for systematic collection of data of similar kinds from several sources. The usual purpose is to draw conclusions about the combined sources rather than about any one source or individual.

The data collection methods are highly varied. Telephone calls may be made to all the prisons in the nation to ask about the presence and characteristics of post-secondary educational programs in the prisons. A questionnaire may be mailed to directors of SCAs to ask how prisoner grievances are handled. A form may be distributed to all the cells in a prison to ask about interest in prison college courses and whether the inmate would like to enroll in a college program. Questionnaires and interviews may be combined to obtain extensive bodies of data on processes and structures, operating problems, and staffing ratios in several correctional systems of a selected type.

Surveys can be either descriptive or evaluative in purpose, although in the latter case they may still contain much data that is primarily descriptive. When used for evaluation, surveys may follow either or both of two procedures. First, the responding units may be grouped together and the responses in various categories may be averaged or otherwise summarized. This value may then be compared with standards or baselines for interpretation and possible action. Second, the responding units may be examined separately, their differences noted, and "better" units may be compared with "poorer" units. The better units thus provide the standard for comparison and they may also become a model for the next phase of the development of the units.

Procedure *one* was followed by the California Special Study Commission when it examined the summarized responses to its survey of probation in California. Its evaluation of the responses led it to

recommend a state subsidy program for the 60 county probation departments.

Procedure *two* turns each of the responding units into a case study, in effect, and the survey procedure becomes a multi-case study. The result will be similar to what the police team cases studies achieved when they made the between-city comparisons to see which procedures were better.

Sometimes a survey is conducted to provide an unofficial standard against which self-assessments can be made. An SCA director who feels a need to document his arguments for increased budgetary support may survey a number of adjacent SCAs to learn the level and range of services provided in some type of correctional effort. He can then compare his own agency's practice or capability with the survey result. If he is well below "average" or below some value that seems appropriate for his situation, he may use this discrepancy as an argument for increased support.

Survey results can also be used as a possible defense against criticism. SCA administrators who feel they are criticized improperly for "unfavorable" escape rates, recidivism rates, costs per man/year or other indices of performance may conduct surveys (if relevant data is not already published) to ascertain whether the criticism is warranted. If the criticism is warranted, then the survey results provide the administrator with a baseline against which he can measure his status and make plans for change. The interpretation of such baselines is often a difficult task, especially if the survey was quickly and informally conducted by relatively inexperienced persons. Actions based on such interpretations need to be carefully considered, preferably with the help of technical assistants or consultants.

3. *Time Series.* Time series are based on measurements of events at periodic intervals and thus provide a picture of trends over time. They show increases, declines, breaks, displacements, fluctuations and stabilities. Prison, parole or probation populations provide data for one type of time series. Prison administrators often anticipate their space needs by extending a time series line into the future, perhaps adjusting the estimates for the next two or three years for factors such as impending changes in law or changing crime rates.

One example of a time series that led to a recommendation for action was the study of narcotic-involved offenders in the D. C. Department of Corrections. This study, like the Preston Impact Study, arose out of researcher curiosity. The departmental administrator asked for information on how many drug offenders there were in the institutions, and the information was obtained on a computer printout, name by name with supplemental data. This answered the question.

In thinking about the numbers involved, a research staff member became interested in the dates of intake of the prisoners and the times of intake were tabulated. This disclosed an exponential curve—generally flat in earlier years, but breaking sharply upward in the last year or two. This could mean a number of things: narcotic-involved offenders were beginning to come into the department in greatly increasing numbers, or the curve was misleading, or there might be other explanations. If the curve was misleading—spurious—it could be reflecting the fact that most narcotic offenders stayed in prison only a year or two while no more than a few served long terms. To check against the possibility that this was the reason for the apparent upturn, a second body of data was obtained from the departmental admissions records. The admissions data showed the same form of curve—only higher up on the chart—hence it was concluded that the department was facing a new environmental condition, the rising use of drugs, which had very unfavorable implications for space requirements. The time series that described this trend was evaluated as "bad."

Research staff, in reporting its findings, recommended that the department consider establishing community-based centers for the handling of narcotic-involved offenders who did not appear to be a menace to the community.[10] The centers could try any of the several treatment modalities that were being tried in various metropolitan areas: therapeutic community, methadone maintenance, narcotic antagonist. The departmental administrator approved the plan, the city government and a congressional committee concurred, and within seven months of the completion of the research report, the first community treatment center was being opened. This was the first of many centers that eventually opened in the District, first under the Department of Corrections, then under the Narcotics Treatment Administration, which was formed to organize and speed the effort.[11]

Not all time series are evaluative. The great majority are primarily descriptions of trends—routine accumulations of data that help administration to monitor its operations and make year-to-year plans to handle relatively minor changes suggested by the data. Occasionally, however, there are signals in the data, either suggesting or forcing the consideration of alternatives to present policy.

4. *Cohort Analysis.* Cohort analysis is a process of developing a special kind of time series—the record of *performance* of a group over time. Cohorts are artificially constructed groups that share some experience such as being born in the same year, being admitted to a hospital in the same year, or being released from prison in the same year. Correctional administrators may want researchers to provide data on the performance of admissions cohorts, program cohorts, or release cohorts.

The most familiar cohort in corrections is the release cohort. The most common research use of the release cohort is to record its performance at specified intervals of time back in the community and to make comparisons with various baselines or standards to ascertain how well the cohort is "performing." This information may be taken as a measure of how well the department, a particular program, or a particular type of offender is doing. In working with release cohorts, the usual grouping is in yearly releases. Sometimes six-month or three-month cohorts are used to shorten the time required before the first data on performance are fed back.

One way of looking at release cohorts is in terms of how many have "returned to the system" by the end of the first year, second year, or *n*th year after release. This can be elaborated to ask whether the return was for a new offense or for technical violation of parole. The total cohort can also be broken down into offense types or program types on which identification is available and the return information can then be made specific to the type. Comparison of returns by type provides data that are useful for some kinds of planning.

One of the SCAs that has been most active in developing and maintaining information on its release cohorts is the California Department of Corrections. *California Prisoners,* an annual publication that is prepared by the statistics division of the Department, has been appearing for nearly thirty years.

Several aspects of cohort performance are practically relevant: performance pattern, performance level, performance distribution, and comparative performance.

The *performance pattern* of release cohorts is often shown as a rising curve of arrests or returns to the system. The curve climbs rapidly in the first few months after release, then slows and levels off after three or four years. Some observers have developed rule-of-thumb estimates, based partly on the California cohort data, that half the failures occur by the end of the first year, 75 percent by the end of the second year, and 90 percent by the end of the third year.[12] There is considerable variation from this pattern of values in many instances so the pattern may be of little practical value at present.

The shape of the pattern, aside from its value at specific points, is often useful in evaluating the influence of external events. In 1968, cohort performance curves of DCDC releasees showed sharp upward displacements that appear to be explained best by two social disturbance—the disorders following the assassination of Martin Luther King, and the rise and fall of Resurrection City, the "poor people's encampment" on the Capitol Mall. The curves of youth, women and adult male release cohorts all showed effects, some highly conspicuous, some more obscure.[13] The most visible effects came on curves that were beginning to level off after the second or third year. The result was that some DCDC performance curves now show clearly two periods of rapid rise: one at the time of release, the other at the time of the 1968 disturbances. The ultimate performance levels (expressed in "failures") for these "double rise" curves will be considerably higher than might have been expected had the disturbances not occurred, we may easily presume.

Departures from the expected pattern have also been noted after some treatments were applied. Treated populations often leave their institutions with an appreciably lower failure rate for the first few months than is usual for the institutional releases. In some instances, however, the second- or third-year rates are back to "normal" for releases from the institution. Any evidence of benefit from treatment has disappeared. Two notable instances of this effect are the Fricot Project and the Case Project of the California Youth Authority and the National Training School for Boys, respectively.[14] The first was a test of small versus large living groups in dormitories; the second was a demonstration of behavior modification by means of a token economy. In each instance the pattern of performance of the treated cohort suggests an "erosion of treatment effect," or, in a different frame of reference, that "the street is taking over."

The performance *level* of a release cohort depends on the length of time it has been out in the community, the types of persons in the cohort, and the types of environments in which reintegration is being

attempted. In some instances, the correctional administrator could add a fourth factor: the treatment experience that the releasees have received or are receiving. Persons who have graduated from the Community Treatment Project (in selected personality types) or from the Provo Experiment or from the PICO Project (in the "amenable" category) are members of cohorts that have significantly lower failure curves than would otherwise be expected.[15]

Since many factors impinge on cohort performance level, it is not informative to speak about a "general recidivism rate" for broad populations of ex-prisoners or ex-probationers. Talk about recidivism makes sense only to the extent that the length of time after release, offender history, kind of person, treatment experience, and intervening social environment are made clear.

The performance *distribution or* "array" of a cohort is a pattern of differential performances that appears when the cohort is analyzed on several criteria of performance. At three years out a cohort may show a high rate of arrests but a low rate of convictions that entail sentences of one year or more. The following tabulation shows the array of performances noted in the total release cohort of adult males from the DCDC reformatory in calendar 1965.[16] The length of follow-up was three years. The releasees totaled 432 in number.

Performance Criterion	*Percent "Failed" at 36 Months*
Arrested by police	60 to 80% (estimated)
Booked into DCDC Jail	46%
Booked and held for further hearing	41%
Booked and fined or sentenced	36%
Booked and sentenced to 1 day or more.	34%
Booked and sentenced to 30 days or more.	32%
Booked and sentenced for 90 days or more.	30%
Booked and sentenced to 360 days or more.	18%

We see here a release cohort with a return-to-system rate as high as 46 percent and as low as 18 percent at the end of three years after release, depending on which of the seven criteria we choose. Since there are many other criteria that could also be used to represent quality of performance in the community, it is obvious that the planning of an evaluation that involves offender cohorts requires some arbitrary decisions about a criterion or criteria.

Performance *comparisons* between cohorts have already been touched on in the discussion of "erosion" of treatment effect and performance levels. Performance comparisons may become rigorous in some types of evaluation, such as quasi-experiments and controlled experiments; they may be rough in other types of evaluation, as in some forms of operations research. Illustrations of these types of comparisons follow later in chapters 10, 11, and 13.

5. *Before-After Studies.* One way of evaluating a correctional treatment is to ascertain the condition or status of a group of clients before involving them in the experience and again after they have been through the treatment. Changes from before to after treatment may be interpreted as a possible consequence of the treatment. How confident this interpretation will be depends on whether other potential influences can be visualized as sources of part or all the observed change.

An illustrative example comes from a pretrial diversion project that was set up recently in Dade County, Florida. Two hundred and five youths were accepted for admission into the Dade County Pretrial Intervention Project during the year ending January 19, 1973. They received counseling, job placement, educational placement, and in some cases psychotherapy. The following tabulation shows status changes that occurred between time of entry and January 19, 1973:

Status	*At Entry*	*1–19–73*
Unemployed, not in school	77	7
Attending school	32	34
Employed	68	93
Attending school and employed	21	42
In vocational training	2	20
In armed forces	1	9
Receiving psychotherapy	4	47
Referred to narcotics rehabilitation	0	38

The members of the group showed notable changes in status while involved in the program, particularly in leaving the unemployed category and entering psychotherapy. To what extent the group might have been influenced by factors other than the program is not clear. The project report did not speculate on other possibilities.

The before-after study is used frequently in corrections, possibly more than the case study, although there are no definite data on this. One of its most extensive uses is in connection with psychological tests, given before and after program involvement. This type of measurement is easy to arrange, but its ability to make an impact on correctional practice is not readily evident.

B. Discussion

Non-experimental studies include a variety of techniques that are usually considered exploratory in nature and rough in the nature of their indications. They are not highly regarded for their precision or the certainty of their findings. They tend, nevertheless, to exert strong influence on the decision-making process in corrections. From the limited information available thus far on levels of influence of impact, such methods appear to be more effective than the experimental and complex statistical methods in bringing about needed changes in corrections.

These methods are good for exploration, well-adapted to quick discovery and quick action, easily understood by operations personnel, and economical. They have the disadvantage of requiring good judgment and wide experience on the part of the evaluator. Paradoxically, they need to be implemented by persons who have the capability of working with the more complex research designs but who might prefer the status and the precision of such designs.

The non-experimental methods are better for exploratory investigations and for suggesting innovations than they are for later quantitative validation and developmental work. Consequently, they may have their greatest utility in times of change.

REFERENCES AND NOTES

1. Walter C. Bailey, "Correctional outcome: an evaluation of 100 reports," *Journal of Criminal Law, Criminology and Police Science*, 57 (June 1966), 153–60. Reprinted as "An evaluation of 100 studies of correctional outcome," in Norman Johnston, Leonard Savitz and Marvin E. Wolfgang (eds.), *The Sociology of Punishment and Correction*, Second Edition, New York: Wiley, 1970, pp. 733–42.
2. Francine Berkowitz, *Evaluation of Crime Control Programs in California: A Review*. Sacramento: California Council on Criminal Justice, April 1973.
3. D.C. Department of Corrections, *A Summary of Completed Research: 1967–70*, January 1971, and *A Summary of Completed Research: 1970–71*, July 1971.
4. D. T. Campbell and J. C. Stanley, *Experimental and Quasi-Experimental Designs for Research*, Chicago: Rand McNally, 1963.
5. Lawrence W. Sherman, Catherine H. Milton and Thomas V. Kelly, *Team Policing: Seven Case Studies*. Washington, D.C.: Police Foundation, 1973.
6. Ibid, p. x.
7. Ibid, p. 102.
8. Stuart Adams, *The Preston Impact Study*. Sacramento: California Youth Authority, 1959.
9. Mary A. Toberg, Lee I. Dogoloff and Michele M. Basen, *Quick Evaluation Methodology: Special Action Office Monograph*. Washington, D.C.: Executive Office of the President, Special Action Office for Drug Abuse Prevention, October 1973.
10. Stuart Adams, Dewey F. Meadows and Charles Reynolds, *Narcotic-Involved Inmates in the Department of Corrections*. Washington, D.C.: DCDC, Research Report No. 12, February 1969; Stuart Adams, *The Narcotic Addiction Rehabilitation Center: A Proposal for a Multi-Modal Community-Based Treatment Center for Narcotic Offenders*. Washington, D.C.: DCDC, April 1969.
11. Robert L. DuPont, "How corrections can beat the high cost of heroin addiction," *Federal Probation*, 35 (June 1971), 43–50.
12. Stuart Adams, *A Comparative Study of Recidivism Rates in Six Correctional Systems*. Washington, D.C.: DCDC, Research Report No. 21, January 1970.
13. Several D.C. Department of Corrections cohort study reports show unexpected upturns in rates, apparently because of the 1968 disturbances. See, for example, Stuart Adams, Wanda S. Heaton, Dimitri Monos and Margaret Upchurch, *Parole Performance Trends Among Community Treatment Center Releasees*. Washington, D.C.: DCDC Research Report No. 7, September 1968; Stuart Adams, Wanda S. Heaton, Dimitri Monos and Margaret Upchurch, *Performance Trends among Youth Center Parolees*. Washington, D.C.: DCDC Research Report No. 8, October 1968; Stuart Adams, Wanda S. Heaton and John D. Spevacek, *Post-Release Performance of 432 Reformatory Releasees*. Washington, D.C.: DCDC Research Report No. 11, February 1969; W. Heaton and S. Adams, *Community Performance of Three Categories of Institutional Releasees*. Washington, D.C.: DCDC Research Report No. 15, June 1969; Colleen Barros, Virginia McArthur and Stuart Adams, *A Study of Post-Release Performance of Women's Detention Center Releasees*. Washington, D.C.: DCDC Research Report No. 31, August 1970; Stuart Adams, John D. Spevacek, Clarence Allen and Kathleen O'Brien, *Trends in Performance in the Community: Youth Center and Community Treatment Center Releasees*. Washington, D.C.: DCDC Research Report No. 36, November 1970.
14. Carl F. Jesness, *The Fricot Ranch Study: Outcomes with Small versus Large Living Groups in the Rehabilitation of Delinquents*. Sacramento: California Youth Authority, Research Report No. 47, October 1965; Harold L. Cohen and James Filipczak, *A New Learning Environment: A Case for Learning*. San Francisco: Jossey-Bass, 1971.
15. Marguerite Q. Warren, "The community Treatment project," in Norman Johnston et al (see reference 1, above) pp. 671–83; LaMar T. Empey and Maynard L. Erickson, *The Provo Experiment*. Boston: Heath, 1972; Stuart Adams, "The PICO Project," in Norman Johnston et al (reference 1, above), pp. 548–61.
16. Stuart Adams, Wanda S. Heaton and John D. Spevacek, *Post-Release Performance of 432 Reformatory Releasees*. Washington, D.C.: DCDC Research Report No. 11, February 1969.
17. Thomas K. Peterson and Richard C. Nichols, *Pre-Trial Intervention Project: A Rehabilitative Program for First Offenders*. Miami: Office of the State Attorney, First Annual Report, January 1973.

CHAPTER 10. QUASI-EXPERIMENTAL STUDIES

Quasi-experimental studies include several varieties.[1] The one to be discussed here makes use of a participant group and a comparison group to measure the effects of an experience. This is the one best known in corrections and probably the most useful for the correctional evaluator.

In principle, quasi-experiments fall below the "true" or controlled experiment in quality of information yielded. In practice, however, there is some overlap in quality. Well-done or well-situated quasi-experiments can provide more valid information than poorly implemented true experiments. Furthermore, quasi-experiments may be conducted in situations where true experiments are for various reasons difficult or impossible. In many instances, therefore, the quasi-experiment is the method of choice. And whether by necessity or preference, it is used more frequently than the true experiment, as the data from Berkowitz and DCDC indicate.[2]

A. Characteristics

For present purposes we define the quasi-experiment as a study in which a treatment process is evaluated by means of information on a treatment group and a "comparison" group. The latter is chosen in a way that makes it "similar" to the treatment group, especially on characteristics that are believed to be related to post-release performance in the community. The choice is not random selection. Rather, it consists of going to a file of records on persons in the correctional population and selecting cases that resemble, at least roughly, the members of the treatment group. If the first selection is rough, the comparison group members may later be screened more carefully, by hand or by computer, to insure closer similarity to the treatment group on key characteristics.

In addition to the similarity of the two groups, the quasi-experiment requires that both are to have had essentially comparable experiences in the correctional system except that the treatment group underwent the special program that is being evaluated while the comparison group did not. Any differences in performance that are later detected by the measurement process are then interpreted as a possible consequence of the treatment experience.

The manner in which the comparison group is selected in quasi-experiments sometimes creates a problem of interpretation when the program is one that is entered on a volunteer basis. Ways of dealing with this problem are discussed below under procedures, Section C.

B. Reasons for Using the Quasi-Experiment

This research design is used when a "rigorous" evaluation is desired but the controlled experiment is not feasible. The latter may be ruled out for several reasons:

- Objections to "denial of treatment" to control group members,
- Operational conditions too complex for a controlled experimental design,
- Treatment program no longer in existence,
- Decision deadline too near to allow for a true experiment.

1. *Denial of Treatment.* The controlled experiment, which randomizes treatment eligible subjects into experimental and control groups, is frequently objected to on ethical grounds. It is regarded as improper to withhold presumably beneficial treatment from eligible subjects, since they may now have to live out their lives suffering from a deficit that assignment to control status was responsible for.

The ethical argument is sometimes disregarded on the ground of the greater good of mankind, and also on the argument that treatment is available for only a portion of the eligibles and must be denied to some anyway. There is also the argument that treatment will be available to the controls at a later date, after the scientific objectives of the experiment have been attained. If the administrator is persistently opposed to "denial of treatment," however, the quasi-experiment is a feasible alternative method of measurement since its "controls" would not otherwise have gone into treatment.

2. *Complexity of Operations.* Many correctional or criminal justice processes are quite complex, and the randomization of cases into treatment and control statuses is often impossible in a way that will insure comparability. For example, in the assignment of individuals to several halfway houses from several different sources—prison, probation, courts, bail agency, and the parole board—the assignment disparities between experimentals and controls may become so great that the randomization process is no improvement over the equating process used in the quasi-experiment. In such cases, it is preferable to use the operationally easier quasi-experiment.

3. *Capturing Historic Events.* Some correctional events that occur only once may be candidates for rigorous measurement because they appear to have had profound effects upon the prisoners or correctional clients who experienced them. The Gideon Decision (U.S. Supreme Court, 1963) brought the release of 1,252 Florida State prisoners before their normal discharge or parole dates. The prisoners were indigents who had been charged with felonies and tried without counsel.

The Florida Department of Corrections in 1965 became interested in the possible effects of such a method of release and conducted a quasi-experimental study to measure the results. A sample of 110 of the early releasees was matched in considerable detail with 110 full-term releasees and both groups were followed up for 28 months after discharge. At that point in the follow-up, the Gideon releasees showed a 13.6 percent recidivism rate, while the full-term releasees showed a 25.4 percent recidivism rate.[3]

Although the interpretation of these results may pose problems, the results have important practical and theoretical implications—and the results could not have been obtained by means of a true experiment.

A different kind of reaching back into history occurred in the research office of the Los Angeles County Probation Department in 1964. It collected from the record systems of the Police Department, the Probation Department and the California Youth Authority old data on the criminal justice experiences of three troublesome juvenile gangs. The gangs had been selected as controls and experimentals in a "natural experiment," using information from area studies in the files of the Probation Department to find comparable groups from earlier years. The main objective was to supply "experimental evidence" in a short time for an urgent policy decision: Should the Group Guidance Program (street work with gangs) be discontinued, as the Police Department was insisting?

The criminal justice data permitted analysis of the effects of full treatment, partial treatment, and no treatment on the three gangs. The principal criterion used was costs and benefits—the arrest, incarceration and supervision data were not amenable to ordinary statistical analysis—and the findings were that street work with juvenile gangs was highly cost-effective. The cost-benefit findings were supported by personal history narratives from ex-gang members and by the criminal justice data, globally interpreted.[4]

4. *Need for Quick Decision.* Both the Gideon and the Los Angeles studies were completed in a matter of months, even though in the latter case the natural experiment extended over a period of six years. By going backwards in time, the quasi-experiment can reduce drastically the time required to make valid comparisons in situations where there are accessible and useful data. When urgent decisions are required, this characteristic of the quasi-experiment is highly valuable.

C. Quasi-Experimental Procedures

After the objectives have been defined, the primary steps in quasi-experimental procedure are 1) identification of the treatment (sometimes called "participant") group; 2) ascertaining the personal-social and criminal characteristics of the group; 3) identifying a comparison group with similar characteristics, making such tests and adjustments as are necessary; and 4) comparing the performances of the two groups in equivalent post-treatment times and situations. These steps are taken up in detail below.

1. *Identification of Treatment Participants.* The participants are, generally, the recipients of a treatment or correctional experience whose effectiveness has become a matter of interest. The D.C. Department of Corrections since 1969 has, in conjunction with the Federal City College, operated a prison college program for inmates. In 1971 it conducted a quasi-experiment to evaluate the college program.[5] A true experiment was not feasible because program staff objected to denying the college experience to eligible prisoners. Consequently, the more flexible quasi-experimental design was used. Participants were defined as men who had experienced at least two quarters of college courses in the prison, 13 or 14 hours per quarter, and had been released to the

community on parole at least six months prior to the time of cut-off in the community follow-up.

2. *Personal-Social and Criminal Characteristics.* In the absence of randomization to insure comparability of the participant and comparison groups, great care must be taken to insure comparability by other means. One way is to make a detailed characteristics profile of the participants and to try to match this profile when selecting the comparison group from whatever its sources.

In the DCDC prison college quasi-experiment, data were sought on about 25 characteristics of the participants. Only 11 of the selected characteristics were found to be adequately reported in the case folders of the participants. Consequently, these 11 characteristics, shown in the list below, became the profile categories on which matching would be attempted:

- Age
- Scholastic achievement test scores
- IQ
- Education completed
- Prior juvenile commitments
- Prior adult arrests
- Prior adult commitments
- Length of sentence
- Marital status
- History of drug use

3. *Identifying the Comparison Group.* The comparison group members in the prison college quasi-experiment were sought among inmates who had applied for the college program but had been released from prison or were otherwise diverted before being admitted to the program. The names were located by searching the applications file for names of men who had not been admitted to the college program, then selecting those non-admissions whose cases showed no prejudicial reasons for non-admission. Generally, these were men who had been paroled earlier than first expected, who entered alternative programs while waiting and did not want to drop them, or who changed their minds while waiting.

It was considered highly desirable, if possible, to draw the comparison group members from men who had volunteered or applied for the college program. This would insure comparability between participants and comparisons on existence of motivation to enter the program. It would probably not insure comparability on intensity of motivation, which was a variable that remained uncontrolled in the study.

No attempt was made to draw a comparison group from the general prison body, although such a group would have enhanced the project methodologically. It might have disclosed how much difference in outcomes, if any, appears to be associated with known motivation and unknown motivation in two comparison groups. If the quasi-experiment grows in importance relative to the true experiment (as now seems likely), it will be very useful to accumulate information on its operating characteristics.

Since participant groups tend to be screened at least twice during entry into a voluntary treatment program—once by motivation to enter and also by possession of the necessary background qualifications—it is important to try to match comparisons with participants on both types of screening factors. In the case of college program participants, few attempt have as yet been made to conduct rigorous evaluations, hence there has been little occasion to attempt matching. There has been no successful true experiment with prison college programs, and the DCDC study is the only successful quasi-experiment with such programs. If these evaluations become more numerous, it would be highly useful to test outcomes for differences that might relate to both the motivation-to-enter criterion and the formal qualifications criterion.

4. *Testing and Adjusting for Participant-Comparison Differences.* Since quasi-experimental studies often involve small numbers, and the sources of comparison group members are often quite restricted, participant and comparison groups may occasionally look quite dissimilar on some of the key characteristics chosen for matching purposes. These dissimilarities should be eliminated if possible in the initial selection and sifting of comparison group members. If little or no sifting is possible, differences in traits may need to be studied for possible effects.

One procedure is to examine each trait difference for magnitude and direction of effect. If the differences are small and the direction is random, there may be no bias in selection so far as the visible characteristics are concerned.

If significant differences or large differences suggest the possibility of biasing effects, tests for the effects of bias may be made and adjustments introduced if necessary.

5. *Comparing Performance.* Performance comparison in quasi-experiments and experiments alike is a difficult domain. In PICO I, a very well-designed true experiment, the researchers examined outcomes and reported no differences worthy of note.[6] A year later, different researchers examining the same data found scientifically important and operationally useful results.[7] Under further scrutiny, these results held up.

In the DCDC prison college quasi-experiment, the researchers found that participants and comparisons each showed arrest rates of about 30 percent each at six months in the community. Had arrests been the only criterion, the program would have been judged ineffective. However, three criteria were being used. The appearance of no difference soon disappeared because convictions proved to be much lower for the participants. The fact that the participants also had more cases pending obscured the results and made a longer follow-up necessary.

In the Los Angeles County Probation Department's quasi-experimental study of street gangs under its Group Guidance Program, the cost-benefits data, the personal interviews with ex-gang members, and the arrest-incarceration-supervision data all indicated that Group Guidance was a success. Another study of Group Guidance by university staff members, using arrest as a criterion and an own-control study design, indicated that Group Guidance was a failure.[8] The conflict in reports of performance in this case has not yet been resolved, although the studies were completed nearly ten years ago.

D. Uses of the Quasi-Experiment

No systematic information is available on the extent of use of the quasi-experiment. The three examples cited above are relatively well known because they have been published or their original reports have been widely circulated. Some additional applications of the method that are beginning to receive attention are in pretrial diversion.

Of the 30 or more pretrial diversion projects now in operation or under trial in the nation, at least fifteen have been the subjects of evaluation, primarily by means of quasi-experiments. Nine of the projects have been grouped into a cluster evaluation, with all the elements in the cluster being evaluated by the same research team and reported in a single report. This means, in essence, seven evaluation projects, six of them singles and one a cluster.[9]

Of the seven projects, two were initiated as controlled experiments, with eligible cases being randomized into project and non-project statuses. The other five studies, because of either organizational complexities or official objection to a randomization process used the simpler process of drawing a comparison group from the court files. In the two controlled experiments, the randomization process yielded very small control groups, approximately one-tenth the size of the treatment groups. In one of the two, the data on the groups suggest an appreciable selection bias—in favor of the treatment group. In the other, the data do not show the characteristics of the two groups. Since some of the quasi-experimental group matchings appear to be closer than the experimental matchings, and since their sample sizes are considerably larger, it is not clear whether the two experimental studies carry any advantages over the quasi-experiments.

Most of the studies pursued three measurement goals: recidivism in-program or post-release, job status, and benefit-cost ratios. In the cluster study, comparison group selection and some parts of the data collection procedure broke down and the study will not be discussed further here. In the other six studies, there was a generally uniform reporting of reduced recidivism on the part of successful program participants, improvement in job status during the program, and benefits from diversion (jail and prison cost reductions) that more than equaled the cost of the project. In two of the six studies the improvements in recidivism were not statistically significant. In one of the projects, an independent evaluation by a second research team found results contradictory to the first findings on recidivism. No check was made of the cost-benefit findings.

Three comments may be made in summary. First, the quasi-experimental and experimental studies were in general agreement that pretrial diversion programs reduced recidivism rates among successful participants, improved the job statuses of the participants, and produced benefits in excess of project costs. Second, the consistency of the findings was such that a policy-maker would be inclined to accept as a "reasonable likelihood" that pretrial diversion usually pays off on all criteria. Third, the studies were all at least moderately crude in design and execution so that the careful reader comes away from the evaluative reports feeling some lack of credibility in practically all the reports. There was frequent evidence of selection bias in favor of the program participants, doubtful assumptions about extent of projected benefits, lack of precise information about follow-up times, and a tendency to ignore important

segments of possible benefits when calculating benefit/cost ratios.

One reason for the poor quality of the research designs and executions was the fact that evaluation tended often to be "tacked on," with a part-time consultant brought in after the start of the project to "conduct an evaluation." Whatever the full set of reasons, a fair rating of the "credibility" of the seven studies might consist of no "high," four "moderate," and three "low" judgments.

The fact that the studies tended to be generally moderate to low in credibility suggests that the evaluation of pretrial diversion programs still has some way to go before there can be confidence that a good first-round measurement has been achieved. Beyond the still-needed first round of rigorous evaluations there also lies an indeterminate number of desired studies whose purpose would be to develop or refine whatever pretrial diversion procedures appear to be promising.

Despite the poor quality of the research, the practice of pretrial diversion has been widely operationalized, first at the sites where the marginal research was performed, and also at sites where the procedure has been accepted either on its "reputation" or on faith. The haste to adopt the procedure carries a moral for the researchers who feel that administrators are reluctant to utilize research, or that they ignore research findings. In actuality, administrators will innovate, in the absence of research, on the basis of poor research, or with good research. One of the administrator's problems is that good research, usefully timed, is almost non-existent. How much the administrator himself is responsible for this non-existence is a perplexing question.

As a final comment on the quality of the research, it is pertinent to add that the best of the several studies, if a fair judgment can be made, was a quasi-experiment. The two controlled experimental studies fell somewhere in the middle of the range of credibility. In view of the fluid and chaotic nature of the arrest-and-adjudicate processes, there is little reason to believe that controlled experiments will have fewer validity problems in such a setting than quasi-experiments. There may have been a valuable lesson in the mingling of the two types of research method in the pretrial diversion studies. This gives a better chance to observe both methods in operation on similar problem, and it may help clarify some of the issues relating to the capabilities and shortcomings of the two methods.

E. Discussion

Although the quasi-experiment is little used in corrections, it is apparently more commonly used than the true or controlled experiment. This greater frequency of use may increase relatively in the next several years. The convenience, flexibility, speed of application, immunity to the "denial of treatment" charge, and other characteristics give the quasi-experiment several advantages over the true experiment. The latter is superior in precision of measurement and in documentation of possible causal effects, but these are less important to the administrator than to the researcher.

Some of the disfavor with which the quasi-experiment is regarded probably arises from our tendency to focus on form rather than on results: the controlled experiment is a more perfect and more prestigious device. Some disfavor arises from the fact that quasi-experiments frequently are applied in slip-shod manner. The circumstances of their use often make them look worse than need be. Given equal care in design and implementation, there is no reason why the quasi-experiment cannot perform significant tasks in correctional evaluation, carrying out many assignments now thought possible only by use of the controlled experiment.

Within the broad array of evaluative methods in corrections, it is not easy to forecast the most likely role of the quasi-experiment. It seems assured of some growth in relative importance, but where it will rank in relation to the older and the newer non-experimental methods (from case studies to simulation) is not clear and also perhaps not critical at this time.

REFERENCES AND NOTES

1. Caporaso, for example, identifies five quasi-experimental designs: a) the nonequivalent control group design, b) the field experiment, c) the cross-lag panel correlation design, d) interrupted time-series design, and e) control series or multiple time-series design. See James A. Caporaso and Leslie L. Roos, Jr., *Quasi-Experimental Approaches: Testing Theory and Evaluating Policy*. Evanston: Northwestern University Press, 1973, p. 12.
2. Francine Berkowitz, *Evaluation of Crime Control Programs in California: A Review*. Sacramento: California Council on Criminal Justice, April 1973; D.C. Department of Correction, *A Summary of Completed Research: 1967–70*, January 1971, and *A Summary of Completed Research:* 1970–71, July 1971. These reviews suggest that from two to five percent of correctional studies are controlled experiments and that from five to 15 percent are quasi-experiments at the present time.
3. Charles J. Eichman, *Impact of the Gideon Decision upon Crime and Sentencing in Florida: A Study of Recidivism and Sociocultural Change*. Tallahassee: Florida Division of Corrections, Research Monograph No. 2, December 1965.
4. Stuart Adams, "A cost approach to the assessment of gang rehabilitation techniques," *Journal of Research in Crime and Delinquency*, 4 (January 1967), 166–82.
5. Colleen Barros, William Caldwell and Stuart Adams, *Effectiveness of the Lorton Prison College Project: Third Interim Report*. Washington, D.C.: DCDC, Research Memorandum 72–4, May 1972.
6. Alvin Rudoff, *The Pilot Intensive Counseling Organization Project, Second Technical Report*. Sacramento: California Department of Corrections, January 1959.
7. Stuart Adams, "The PICO project," in Norman Johnston, Leonard Savitz and Marvin E. Wolfgang (eds.), *The Sociology of Punishment and Correction*, Second Edition, New York: Wiley, 1970, pp. 548–61.
8. Malcolm Klein, *Street Gangs and Street Workers*. Englewood Cliffs: Prentice-Hall, 1971.
9. Final reports have been released on three of these projects: Leon G. Leiberg, Roberta Rovner-Pieczenik and John F. Holahan, *Project Crossroads*, 3 vols. Washington, D.C.: National Committee for Children and Youth, 1971; Peter S. Venezia, *Pre-Trial Release to Supportive Service of "High Risk" Defendants: Third-Year Evaluation of the Des Moines Community Corrections Project*. Davis, Calif.: National Council on Crime and Delinquency, 1973; Edward de Grazia, *Pre-Trial Diversion of Accused Offenders to Community Mental Health Treatment Programs*. Washington, D.C.: Georgetown University School of Medicine, Department of Psychiatry, 1972. One or more interim reports have been distributed on the following projects: Manhattan Court Employment Project of the Vera Institute of Justice, New York City; Project Operation Midway, Nassau County Probation Department, New York; Pre-Trial Intervention Project, Metropolitan Dade County, Florida; and the Pre-Trial Intervention Project, An Evaluation of the Nine-Site Court Diversion Program of the Manpower Administration, U.S. Department of Labor.

CHAPTER 11. CONTROLLED EXPERIMENTS

The controlled experiment may be losing ground to the quasi-experiment as the "work horse" of rigorous correctional evaluation. It is also being criticized by the "exploratory holistic" school of evaluators as unproductive.[1] This does not necessarily mean that this research design is "through" as an aid to information-testing and decision-making in corrections. It does mean that it is time to reconsider its role and potential in correctional evaluation. We need to review its characteristics, accomplishments, strengths and weaknesses, and its possible future.

A. Background

The controlled experiment, which assesses a treatment experience by measuring differences in response by "treatment" and "control" groups who came randomly out of a pool of treatment-eligibles, has been used extensively in corrections since the early fifties. The Department of Corrections in California attracted much attention with its SIPU projects, 1953-64, in which many thousands of parolees were involved in a four or five-stage experiment on the effects of caseload size, parolee-agent matching, base-expectancy differences, regional differences, and other variables. It used the same design in more elaborate form in the PICO Project, 1955–61, in which 1,600 older youth prisoners were involved in a test of the efficacy of individual interview therapy on "amenable" and "nonamenable" personality types.

Equally as massive but less well publicized were the Department's experiments with group counseling in a prison, with parole caseloads varied by parolee service needs, with treatment and supervision of narcotic offenders, and with a number of offender management processes of other kinds.

The California Youth Authority also conducted experiments on the effects of parole caseload size, randomizing several hundreds of youths into regular and small caseloads over a period of about three years in one of the larger counties. It conducted the Community Treatment Project, in which, during a 13-year span, several hundred youths were classified under an interpersonal-maturity typology, matched with appropriate parole agents, assigned to differential treatment in the community, and compared on outcomes with control cases who had followed the usual course through the CYA training schools. Several other experiments, large and small, assessing various types of institutional treatments, were conducted between 1958 and 1974.

The Los Angeles County Probation Department, responsible for the supervision and local detention of many scores of thousands of probationers, also conducted a number of controlled experiments during the fifties and sixties. One of the first (TOPS: To Improve Probation Services) dealt with the issue of probation caseload size. Others dealt with institutional treatment of girls and with the effectiveness of street work with juvenile gangs—the latter a university-staffed experimental project. One of the most recent, RODEO (Reduction of Delinquency through Expansion of Opportunity), examined regular probation, probation with indigenous aides as part of the probation supervision team, and county camp placement as treatment alternatives.

Although no counts have been made, it is a reasonable conjecture that the controlled experiments by these three state and local correctional agencies include half or more of the body of controlled experimentation that has occurred in corrections during the past twenty years. If this is not true in numbers of projects, it is certainly true in numbers of subjects involved.

The outcomes of these numerous experiments involving many thousands of adult and youth prisoners, parolees and probationers are difficult to summarize concisely. There have been both operational and scientific gains, but undoubtedly in smaller measure than was originally hoped. Some assessments of the entire array of experiments would, by implication at least, judge them to be "inefficacious," or "without decisive effect in reducing . . . recidivism." A more objective judgment might see the experiments as paying off in some manner in perhaps ten to twenty percent of the total number attempted.

B. Assessing the Impact of Experiments

Since it is desirable to know something about the impact of particular kinds of evaluative methods when their use is planned, it is long past time when the results of controlled experimentation should have been examined systematically and critically. Before proceeding further, it will be useful to make an on-the-spot assessment of some of the better known experiments (or intended experiments) on operations in the agencies just mentioned. This assessment will be subjective and perhaps controversial, but it will move toward filling an important need.

The following list of "experiments" includes the better known from the paragraphs above, with additions from the general literature to broaden the perspective. The experiments are rated "Low", "Mod" (Moderate), and "High" on contributions to agency operations and to scientific knowledge. They are rated similarly on credibility, which denotes quality of research design, adequacy of the design to the specified problem and the context of the study, and general impressions of reliability and validity of the findings.

Experiment[1]	*Contributions* Opns.	Knowl.	*Credibility*
Community Treatment Project.	Mod	High	High
C-Unit: Community in Prison*.	Low	Low	High
CYA Parole Caseload Project.	Low	Low	High
PICO: Individual Psychotherapy.	Low	High	High
Group Counseling in Prison*	Low	Low	Low
Provo: Guided Group Interaction*.	Low	Mod	High
PWUP: Parole Work Unit Project.	High	Low	Mod
RODEO: Camp vs RODEO vs Probation.	Mod	Low	Mod
Silverlake: Therapeutic Milieu*.	Low	Low	High
SIPU: Parole Caseload Study	Low	Low	Mod
Street Gangs and Street Workers*.	Low	Low	Low
TOPS: Probation Caseload Study.	Mod	Low	Mod
Youth Center Research Project.	Mod	High	High

*University-based studies

This assessment is rough and intuitive, with ground rules only partially defined. It is presented primarily to emphasize a problem (that we know all too little about the effectiveness of our evaluative models) and to motivate discussion of the results of controlled experimentation in correctional evaluation. The discussion will be brief and tentative since with different ground rules and other judges, different ratings would be arrived at for some of the experiments. Furthermore, space does not permit review of such matters as accumulative or indirect contributions, variations in size and duration of experiments, the possibility that low levels of contribution may be very important over time, nonutilization as a barrier to operational contribution, and so on.

With these qualifications, it appears that the contribution of the controlled experiment to correctional evaluation and correctional decision making is less than might be anticipated, given the high value placed on this design. How much this contribution has been lessened by faulty design, execution, or utilization is not clear, although evidence of each of these kinds of problems is present in some of the studies.

It is of interest that the Community Treatment Project and the Youth Center Research Project stand highest in overall rating. Experiments with youth programs seem more carefully executed and generally more productive than those with adult programs. It is also of interest that experiments in local agencies have a higher proportion of operational consequences of at least moderate level. And, finally, the state agency experiments rate higher in contributions to operations and knowledge and lower in credibility than the university-based experiments.

It may be added that credibility appears generally to outrun contributions to knowledge or operations, suggesting that "looking well" is an easier achievement in research than "doing well." And it is perhaps worthy of note that the credibility ratings of the thirteen studies listed here are considerably higher than the credibility ratings of the five quasi-experiments on pretrial diversion that were discussed in the preceding chapter.

These somewhat conjectural assertions lead to two further points. First, it is not yet clear that the controlled experiment is an especially productive means of improving either correctional practice or correctional science. It makes contributions to both but at a relatively high cost in time and resources. It would be well to look for other means of obtaining at least some of the information that is now being derived so laboriously and with such low payoffs from controlled experiments. One pos-

sibility, obviously, is the quasi-experiment. If taken seriously and used with care, this method can reduce appreciable the time and resource costs in some areas of correctional evaluation. Another possibility is operations research, and still another is simulation, both taken up in later chapters.

Second, it is possible that the controlled experiment can be made more useful in corrections. This might come through fuller knowledge of its characteristics and capabilities, and by analysis of ways in which its effects can be optimized. It is relevant that six of the nine experiments shown above that involve youth make a contribution to either operations or knowledge. Only one of the experiments with adults makes a contribution, and this contribution may derive more from operational wisdom than from research findings.

There are undoubtedly some useful clues here. Analyses of these kinds may help in clarifying research priorities and research strategies. They may also turn up needed guides as to staffing, timing, structuring, and utilizing experiments.

We turn now to some aspects of the implementation of the controlled experiment and some of the problems that are often encountered in implementation.

C. Implementation of Controlled Experiments

The controlled experiment measures the effect of a process by applying it to one group of subjects while withholding it from an identical group and observing what happens. It proceeds on the assumption that if there is an effect, it will appear in the behaviors of the treatment group but not the controls.

The design of the experiment may be two-way as in SIPU, or four-way as in PICO, or 18-way as in CTP. The more elaborate designs permit more numerous inferences about treatment impact and about interactions between treatment and the characteristics of the subjects. In the PICO Project, for example, it was possible to draw conclusions about 1) the validity of the amenability-nonamenability classification scheme, 2) the magnitude of the treatment impact, and 3) reasons for differences in performance of the treated amenables, treated nonamenables, and the two varieties of controls. Some of these conclusions were relatively certain and some were relatively uncertain.[3]

The major steps in a controlled experiment are the definition of objectives, formulation of treatment plans, definition of treatment eligibles, implementation and testing of a randomization procedure, application and monitoring of treatment, monitoring and collection of data, analysis of the data, and the drawing of inferences.

Most of these features are common to several types of evaluation, particularly the quasi-experiment. Comments follow on the one feature that is different—the implementation and testing of the randomization procedure.

1) *Randomizing into Experimental and Control Statuses.* The persons who are eligible to enter an experiment may differ from the total population in definable ways, as in CTP (which excluded "dangerous" and "unacceptable" youths), or they may exclude no one, as in some parole caseload studies, which regarded all parolees as eligibles.

Once the eligibility criteria have been specified and procedures for screening have been set up, a method for randomizing the eligibles to treatment and control statuses is needed. One method is to list the eligibles in alphabetical order and assign every other one to experimental or treatment status; the rest automatically become controls. Another method is to work with a chronological or identification number file, taking every other one in order. If an even division into treatment and control is not wanted—as when the subjects must go preponderantly into treatment status because there are barely enough eligibles to fill the treatment slots—every third or fourth in line may be designated control and three-fourths or two-thirds will go into treatment.

Taking every other one or every fourth one is "interval" sampling. It is a simple, mechanical procedure for creating an "identical" group whose performance will later be a baseline for measuring effect. Sometimes interval sampling is unwise because it becomes apparent who is going into treatment and control in the future and there are some staff members who cannot resist putting a control designee who "really needs treatment" into the treatment group. Many randomization procedures have been upset by "soft-hearted" staff members who are in positions where they can manipulate the assignments.

This susceptibility of experiments to being seriously affected or even invalidated by either accidental or deliberate biasing of the assignment procedure creates a problem. To counter the problem, the researcher will need to employ a variety of monitoring, testing and adjustment procedures.

2. *Coping with Biased Assignment.* One procedure for coping with the problem of biased "randomization" is to prevent the problem as much as possible. This can be done in part by the use of non-obvious selection methods, such as a table of random numbers. Another procedure is to keep a constant watch on the assignment process so that it may be adjusted in midstream if necessary. Still another is to make a post-assignment review of the characteristics of the accumulated treateds and controls to see whether they are in fact randomly distributed. If the percentages of each group with a given characteristic—for example, history of drug involvement—are close, randomization has apparently worked.

If appreciable differences occur in distributions of key characteristics, significance tests may be run; also, the directions of differences may be noted to see whether all the differences favor one group. In the Des Moines Community Correction Project, 1970–73, a controlled experimental evaluation was conducted, involving 281 experimentals and 36 controls. The two-year report showed the profiles of the two groups on 37 characteristics. On 4 characteristics, the controls were significantly "worse" (that is, varied in the direction that appeared associated with higher risk of criminality); on 14 others they were worse but not significantly so, and on 5 they were better. On 14 characteristics there were no appreciable differences between the two groups.[4] There was apparently a systematic bias in favor of the treatment group in this study, and the outcomes (which showed the treatment group as being "helped" significantly by the pretrial release program) were probably influenced to some degree by the bias. As a consequence, both the measures of effect and inferences about the nature of the effect are made conjectural by the selection bias.

In some cases where selection bias has been introduced by the implementation of an experiment, adjustments to counter the biasing effect have been considered or made. The practical effect of a single biasing factor can sometimes be estimated arithmetically, and if it is small, it can be ignored. The PICO Project report showed a significant difference between treated nonamenables and control nonamenables on "offense type." It was estimated that adjusting outcomes for this difference would increase the performance differential trial between the two groups by 1.3 percent, which was considered small enough to ignore.[5]

The need to adjust for selection bias should occur only rarely in controlled experimental studies; it would be expected more frequently in quasi-experimental studies. In both cases, however, the possibility of selection bias should be routinely checked.

D. Problems of Experimentation

When experiments move out of the laboratory into the field, particularly the field of social action, they run into numerous difficulties. The problem of selection bias has already been noted. Campbell has listed a number of others, some internal to the experiment, others external.[6] Of his nine internal problems, categorized as "threats to internal validity," the correctional evaluator needs to be especially concerned with three: selection, history (external events), and data instability. The other internal threats (maturation, regression, testing effect, instrumentation, experimental group mortality, and selection-maturation interaction) appear to be of less concern in the typical correctional experiment.

Other problems of concern to the evaluator, not listed by Campbell, are masking effects, contamination of data, criterion problems, and "erosion" of treatment effects. Some of these problems deserve careful attention lest they invalidate the experiment or seriously confuse the interpretation of the results. They need to be prevented, accounted for in the experimental design, or adjusted for in the analysis and interpretation of the data.

1. *Masking.* Experimental treatments may have opposite effects upon different kinds of subjects. If the different subjects are included in the same experimental group, their performance changes may cancel each other in the summing up, with the result that no differences can be perceived between the performance scores of experimentals and controls.

Masking effects were inferred in the results of SIPU I and II (1953–55) and attempts were made to deal with these by various classification or typing procedures in SIPU III and IV (1956–64). One procedure was the use of base expectancy (parole success probability) scores. Another was to type parolees by whether their problems appeared to be "internal" or "external" and parole agents by whether they were "directive" or "supportive." None of these exploratory solutions appeared effective to an appreciable degree.

PICO (1955–61) dealt with anticipated masking by starting with a four-way controlled experimental design based on judged "amenability" to the proposed individual psychotherapy. This procedure was

at first nullified by an external event (change in the leadership of the clinic that made the amenability judgments), but a later reanalysis of the data uncovered the "remasking" effect. When the data were properly analyzed, the amenable-non-amenable classification proved highly successful in uncovering differential responses to treatment and in showing that individual psychotherapy with amenable older Youth Authority wards was significantly and practically effective. This series of "unmasking," "remasking," and "re-unmasking" events in the PICO Project yielded valuable insights into both the differential effects of treatment on diverse general types of offenders and the difficulty that may be encountered in establishing these effects.

The Community Treatment Project (1961–74) carried the unmasking process still further. Applications of typologies were made to both clients and parole agents, thus advancing beyond PICO in comprehensiveness and beyond SIPU in the theoretical sophistication of the typologies employed. These advances by CTP resulted in both operational and scientific contributions well beyond those achieved by either SIPU or PICO.

Other projects have shown the utility of classifying or typing procedures in evaluative studies, although not to the extent evident in CTP. One informative though quite elementary application was made by the D. C. Department of Corrections in its study of the effectiveness of work release (p. 86, Chapter 13).

2. *Contamination of Data.* An important kind of data "contamination" occurs when experiences of control groups do not follow plans. If subjects in the control group become exposed to program experiences that may conceivably influence their post-program behavior, even though it is not the same type of program that is being investigated in the study, the measurement process may be invalidated by a distortion of the comparison baseline. In the Los Angeles Group Guidance Project evaluation, a partially treated gang performed more like the treated gang than like the control gang. Had the partially treated gang been considered merely another control gang, much of the difference between treated and "untreated" gangs would have been obscured.

3. *"Erosion" of Treatment Effect.* A very common though little-studied occurrence after treatment is the gradual (or sometimes abrupt) disappearance of a performance superiority shown by the experimental group in early months after treatment. Reference has already been made to this effect in a living unit study and a behavior modification study (p. 57). Numerous other examples could be cited after a careful search of the literature. Still more examples could be documented if more careful followups were made of experimental cohort performance patterns in the weeks and months immediately after termination of treatment. In some cases, the onset of the erosion effect is so sudden that it may be missed or its significance be understated if the first followup point after treatment termination is set at a year or more.

Why some experiments are followed by an obliteration of an apparently early superiority in performance capability while others are not is as yet a mystery. It is easy to say that "the street is taking over," but this is only a gross, common-sense hypothesis. A better hypothesis would suggest possible explanations for those instances in which the street does not take over. We would then be able to build into experiments "sensors" for the difference between a PICO, where individual psychotherapy created a performance differential that was increasing at three years out and a Case Project where behavior modification created a differential that was obliterated by three years out.

Until this difference is better understood, it will be difficult to know whether a specific treatment program is "inefficacious" or the deviance system in the community has strongly countered an effective treatment process. The practical effects may be the same, but the implications for theory and for long-range research planning are quite different.

4. *Criterion Problems.* Experimentals and controls may show little difference on some criteria and significant differences on others. There has been relatively little study of the criterion problem in correctional experiments, partly because many experiments have used only one performance criterion. The stimulus to thought that comes from contradictory criterion indications cannot arise in such cases. Thus far, there has been too little recognition of the importance of having more than one performance criterion in evaluating correctional programs.

At least three kinds of criterion problems can be visualized: use of a single criterion, divergent readings on commonly used criteria, and conflicting readings between a familiar and an unfamiliar criterion.

a. *Use of a single criterion,* particularly arrests, to measure performance is a risky procedure. In some cases it serves modest objectives; in others it may invalidate or at least reduce the credibility of the study. On occasion the risk is justified if some early indication is needed or if arrests are the only indicator that is practically available. If a single criterion is used—particularly arrests—it is prudent to attempt at least partial validation by using another criterion on a sampling basis, if possible, or observation and judgment when quantitative measures are lacking.

b. *Divergent indications by familiar criteria* call for careful interpretation. If the divergence is one where arrests and convictions disagree, for example, the indication by the less valid and less reliable criterion, arrest, may be discounted. When the two criteria show zero or even negative correlation, additional data should be sought to provide clues to the unusual relationship.

When prison college parolees and comparison parolees showed similar arrest rates but divergent conviction rates, there was both reason and material for speculation. Perhaps the police were affronted by the concept of a parolee-collegian and "busted" this new category of parolee on slight pretext. On the other hand, the discrepancy may have arisen with judges, who felt that a parolee who was a college student had a special claim to consideration, so that for equal offenses the collegians may have drawn fewer convictions.

These conjectures could have been resolved by additional data that probed the experiences of the college participants and comparison group members at arrest and in court, but circumstances did not permit another study at the time.

c. *Contradictory indications by familiar and unfamiliar criteria* sometimes appear when recidivism and cost data are brought together on the same treatments. Occasionally the recidivism and cost data agree, but this may reflect nothing more than the fact that the latter were calculated from the recidivism data. When these two bodies of data come from different sources and produce divergent evidence, they require interpretation.

One example of such divergence occurred in the Silverlake Experiment, where treatment of juveniles in a "therapeutic milieu" brought no appreciable changes in recidivism but reduced treatment costs by one-half.[7] The authors accepted the cost reduction indication as valid. Another example was the Los Angeles County Group Guidance Program evaluation. One of two evaluations of the program found that street worker intervention increased arrest rates among members of treated gangs.[8] The other study found that the recidivism data presented a complex pattern, apparently favorable to the treated gangs, but the cost data were clear-cut and decidedly in favor of the treated gangs.[9] The study that used arrests as a criterion disregarded the cost data and concluded that street worker programs were ineffective or even dysfunctional in the Group Guidance format.

5. *Problem of Significance.* A paradoxical problem with the controlled experiment is that it is apparently most useful or at least most applicable in structured situations—testing Program A against Program B, for example. In the broad framework of correctional structures and processes, these experiments usually appear to be of minor significance.

This characteristic of the experiment implies that researchers who stress the importance of this evaluation model (e.g., Campbell, Suchman, Weiss) misperceive its role in the decision-making process in social agencies. The major issues tend to be resolved and major actions tend to be taken on the basis of configurations of information and interests, not on the findings of a crucial experiment. The information component in the process seems much more likely to come from non-experimental than from experimental models.

Whether the experimental model will in the future find a more influential role in decision-making than it now occupies is unclear. If corrections moves from fluidity to a more structured, laboratory-like state, with experiment-minded administrators and experiment-promoting chiefs of research, the controlled experiment may rise in importance. If present trends toward change, toward increasing use of uncertain information, toward integration of process and structure in ever widening systems continue, the controlled experiment may decline in importance.

6. *Problem of Relevance.* If the controlled experiment is at its best in carefully structured situations, it is also at its best when ample time is available. SIPU went through a number of phases over a period of eleven years, seeking to throw light on the issues in caseload size. Eleven years were not enough, and the decisions on caseload size reduction went ahead with little help from SIPU. Caseload sizes went steadily downward, largely because of pressures from professional standards, although possibly to some extent from a belief that SIPU would eventually prove that small caseloads paid off in lower recidivism rates. We must conclude that the results of SIPU were less relevant

to the administrators' objectives than were the professional standards on caseload size.

PICO went through only one implementation phase, although it experienced two analytical phases. Significant scientific and operational information came out of the second analytical phase. Yet PICO brought no operational changes of consequence. Institutional programming continued as before, ignoring the PICO findings, although there was occasional discussion of the "significant" results by executive staff. At the time of the uncovery of the significant information—six years after the start of the experiment—PICO was no longer relevant to the principal concerns of the Department of Corrections, it would appear.

The Community Treatment Project went through three phases that covered a span of 13 years. It yielded very important information, both in operational and in basic science terms. However, this project terminates in 1974 on a declining note—overshadowed by such amorphous and nonexperimental entities as probation subsidy, youth service bureaus, and so on. For various reasons, the highly successful CTP seems much less relevant in 1974 than it seemed in 1961.

Conspicuously less relevant was the study of group counseling in a mens' prison—*Prison Treatment and Parole Survival*.[10] Started in 1962, at about the time that group counseling in California prisons had reached its peak, the experiment resulted in a final report in 1971, approximately nine years after the information might have been useful. Since the project yielded no useful findings—partly because of an unproductive design, partly for other reasons—it was in a poor position to impact departmental policy, especially since interest in group counseling had now been declining for several years.

The reason for the apparent irrelevance of many controlled experimental projects at the time of their conclusion is probably a complex of factors. There is the possibility that in a rapidly changing field, an experiment that takes five, ten, or fifteen years for execution falls far behind the central issues of the field as structures and processes evolve. There is also the possibility that the minor problems of policy that experiments are most suited to address are likely to be of low relevance most of the time, particularly after an initial burst of researcher-generated interest recedes. For these and probably other reasons, the controlled experiment has problems of both relevance and significance. In devising research strategies, these are matters of considerable importance for both the agency administrator and the chief of research.

E. Discussion

The controlled experiment has been actively employed in corrections over the past twenty years, at first quite confidently and hopefully, then more cautiously and with growing reservations. A review of results would appear to indicate that this evaluation design has had less impact on the field of corrections than looser and "softer" research designs. The experiment is perhaps more a researcher's instrument than an administrator's aid to management.

Whether the controlled experiment is destined for a period of decline as systems methods—which appear more suited to the growing complexity and scope of criminal justice evaluation problems—rise in importance is a pertinent issue. If this is indeed to be the case, the controlled experiment may lose ground on two fronts—to systems methods on the wider horizon, and to the quasi-experiment on the narrower arena of rigorous measurement. In the latter area, the controlled experiment is the researcher's favorite, but the quasi-experiment may fit the needs and styles of administrative decision-making somewhat better.

The ultimate role of the controlled experiment will be determined in part by the extent to which corrections becomes a rapidly changing subsystem within the criminal justice system. The latter must eventually become the primary frame of reference in the evaluation of processes that deal with offenders. Since corrections is a relatively minor part of the larger system, broader considerations must sooner or later dominate evaluative research in corrections.

Administrator styles and administrator-researcher interaction may also play a part in determining the future of the controlled experiment. If a growing proportion of administrators adopt the traditional research point of view, there may be a resurgence of interest in making evaluative research in corrections a series of well-chosen experimental tests of key hypotheses concerning correctional process and structure.

There will continue to be areas within corrections where the controlled experiment will be considered the preferred evaluation model. This might occur where there are relatively stable structures and procedures and where highly complex designs leading to causal inferences are desired. The latter two conditions do not rule out use of the quasi-experiment, but they encourage a preference for the true experiment.

The possibility that the controlled experiment has been used to poor advantage over the past twenty years raises a number of questions. What is the best role for the experimental method in corrections? Should it emphasize small studies such as RODEO at the local level, or large studies such as SIPU at the state level? Should it seek precision in measurement, or should it stress theoretical issues and causal inference? What has been the relative impact of the experiment as compared with non-experimental studies? Of the several uses thus far made of the experiment, which have had the most important operational consequences?

There is a tendency on the part of some researchers and administrators to assert that evaluative research should make more use of the controlled experimental design. The wisdom of that assertion is not self-evident. It appears to be more important to ask what have been the operational results of experimentation in corrections over the past twenty years and what next steps appear to be indicated by a review of those results.

REFERENCES AND NOTES

1. Michael Marien, "Daniel Bell and the end of normal science," *The Futurist*, 7 (December 1973), 262–68.
2. Following are more complete references to the thirteen experiments: Marguerite Q. Warren, *The Community Treatment Project*, Sacramento: California Youth Authority, various reports, 1961–1974; Elliot Studt, Sheldon L. Messinger and Thomas P. Wilson, *C-Unit: Search for Community in Prison*. New York: Russell Sage, 1968: Bertram M. Johnson, *Parole Research Project: Evaluation of Reduced Caseloads*. Sacramento: California Youth Authority, Research Report No. 27, October 1962; Stuart Adams, *Effectiveness of Interview Therapy with Older Youth Authority Wards: Interim Evaluation of the PICO Project*. Sacramento: California Youth Authority, Research Report No. 20, January 1961; Gene Kassebaum, David A. Ward and Daniel M. Wilner, *Prison Treatment and Parole Survival*. New York: Wiley, 1971; LaMar T. Empey and Maynard L. Erickson, *The Provo Experiment*. Boston: Heath, 1972; E. Farley Hunter, *Reduction of Delinquency Through Expansion of Opportunity*. Los Angeles: L. A. County probation Department, Research Report No. 33, June 1968; Walter R. Burkhart, "The parole work unit programme; an evaluation report," British Journal of Criminology, April 1969, 125–47; LaMar T. Empey and Steven G. Lubeck, *The Silverlake Experiment*. Chicago: Aldine, 1971; Joan Havel, *Special Intensive Parole Unit: Phase IV*. Sacramento: California Department of Corrections, Research Division, Research Report No. 13, September 1965; Malcolm Klein, *Street Gangs and Street Workers*, Englewood Cliffs: Prentice-Hall, 1971; Edith M. Tufts, "Effectiveness of Reduced Caseloads," in *The Treatment of Delinquents in the Community: Variations in Treatment Approaches: Monograph No. 1*. California State Board of Corrections, July 1960, pp. 15–21; Carl F. Jesness, *Youth Center Research Project*. Sacramento: American Justice Institute, 1972 (2 vols.)
3. Stuart Adams, "The PICO project," in Norman Johnston, Leonard Savitz and Marvin E. Wolfgang (eds.), *The Sociology of Punishment and Correction*. Second Edition, New York: Wiley, 1970, pp. 548–61.
4. Peter S. Venezia, *Des Moines Community Corrections Project: Pre-Trial Release to Supportive Services of "High Risk" Defendants*. Davis, Calif.: National Council on Crime and Delinquency, Evaluation Report No. 2, February 1972, pp. 32–4.
5. Stuart Adams, op. cit., p. 557.
6. Donald T. Campbell, "Reforms as experiments," in James A. Caporaso and Leslie L. Roos, Jr. (eds.), *Quasi-Experimental Approaches*, Evanston: Northwestern University Press, 1973, pp. 187–225.
7. LaMar T. Empey and Steven G. Lubeck, *The Silverlake Experiment*. Chicago: Aldine, 1971, pp. 309–10.
8. Malcolm W. Klein, *Street Gangs and Street Workers*. Englewood Cliffs: Prentice-Hall, 1971, p. 139.
9. Stuart Adams, "A cost approach to the assessment of gang rehabilitation techniques," *Journal of Research in Crime and Delinquency*, 4 (January 1967), 166–82, p. 181.
10. Gene Kassebaum, David A. Ward and Daniel M. Wilner, *Prison Treatment and Parole Survival*. New York: Wiley, 1971.

CHAPTER 12. COST ANALYSIS AND COST-BENEFIT ANALYSIS

We have now seen that there are several ways to measure correctional effectiveness. One method compares situations with standards and specifications; another compares behaviors with behaviors; still another compares agency actions with other agency actions.

One additional method for measuring whether a process or program is preferable to another is to introduce monetary values as a basis for comparison. These values may be costs or returns or both.

There are several reasons for introducing the monetary criterion into correctional evaluation. On a general or philosophical level we have the fact that many offenders who are institutionalized (the most expensive form of correctional procedure) are widely regarded by knowledgeable persons as not in need of incarceration. No good is served by the process. This means that corrections is being needlessly inefficient, often to great extreme. In essence, it is wasting scarce resources.

Furthermore, there is wide belief that institutionalization of many and perhaps most offenders is not only needless but also dysfunctional. It not only does not help; it harms the offender, or his family, or the community, or all three. It reduces the offender's economic status and potential directly, and punishes his family and community indirectly.

Finally, only a small percentage of crimes are cleared by arrest and conviction of the guilty offender. If incarceration is punishment for an offense, it is a wasteful, dysfunctional and inequitable way of administering punishment. The inequity is frequently defended on the grounds of deterrence—of the offender himself and also other potential offenders. However, the evidence that deterrence actually results from most convictions or most incarcerations is not conclusive.

Consequently, corrections as now conducted appears wasteful, dysfunctional and inequitable. It may be described as an ill-advised use of resources, and the best way of understanding its lack of advisedness is to study it in terms of resources expended and returns received. Such a study is good not only for understanding but also for action. Policy makers sometimes find it easier to make decisions on the basis of economic loss and gain than on more abstruse considerations.

There are also technical or tactical reasons for using the monetary criterion. One is that it often discloses positive outcomes even when the behavioral criterion indicates no gain. In the Silverlake Experiment there was no improvement in recidivism on the part of the treatment group, but there was a decided economic advantage for the treatment program. Another is that this criterion can summarize meaningfully a mass of behavioral data that is statistically unmanageable as behavioral data. The monetary criterion provides a common denominator that translates behaviors into economic consequences and permits easier summation and analysis. Finally, the monetary criterion permits valid choices between programs when the choice would be invalid on the basis of recidivism rates alone or program costs per capita alone.

The monetary criterion is thus more powerful and more versatile than the behavioral criterion. It may also be generally more useful. It speaks the language of the policy maker and the budget keeper without losing sight of the offender behaviors that underlie the problem at hand. Because of these qualities, the monetary criterion is likely to grow in importance in corrections. We will consider some of its applications, looking particularly at the categories of 1) cost analysis, 2) cost comparisons, and 3) cost and benefit comparisons.

A. Cost Analyses

Cost analyses as considered here are the operating costs of attempting specified correctional objectives such as "rehabilitating a juvenile gang," or "correcting 25 youthful offenders." These studies trace the correctional or criminal justice actions or services involved in dealing with specific offender management tasks or processes. Each action or service is "costed" by applying business office or auditor figures to each unit of action and service and totaling the costs. This may be done for a particular program experience by one or more offenders or for a series of experiences in several agencies over a period of years. Two examples follow:

1. *The Cost of Correcting a Juvenile Gang.* In 1964 the research office of the Los Angeles County Probation Department became interested in the magnitude of the criminal justice costs that were generated by one of the delinquent juvenile gangs that was known to the Department. The Department's Group Guidance Program, a street worker activity that was trying to control and change a few of the County's estimated 300 delinquent gangs, performed "area studies" and maintained rosters on the core members of the more troublesome gangs. The research office decided to trace and cost the official histories of core members in one of the gangs on which Group Guidance had a membership list.

Arrests, court appearances, detentions, probation terms, county camp stays, jail or juvenile hall stays, Youth Authority commitments and parole terms were identified as the basic data. Information on these actions and services was assembled by tracing the core individuals through the police, probation, and Youth Authority records. Unit costs were obtained from the business offices of the agencies involved or from the county auditor.

The followup disclosed a total expenditure of about $200,000 for the 24 core members of the gang over a six-year period. This amounted to about $33,000 per year for the 24 boys or about $1,400 per boy-year. The biggest component of cost ($61,491) was borne by the California Youth Authority. The next largest component ($51,099) was borne by the Probation Camp system of the Los Angeles County Probation Department.[1]

By itself, this case study was primarily descriptive. In conjunction with similar studies on properly selected gangs, it could serve as part of an evaluative study of considerable significance. For example, when compared with data on a "treated" gang that had similar characteristics, it could provide evidence on the economic consequences of street work.

2. *Cost of Correcting Youthful Offenders.* A similar study was undertaken in the D. C. Department of Corrections in 1968. A 25-youth sample of the 1968 "graduates" of the Department's Youth Center was taken and their criminal justice careers were traced and costed. The rationale here was similar to that in the Los Angeles gang costs study. Taken alone the study would be merely descriptive; used later in comparative analysis, it would be a means to an evaluative study.

The 25 DCDC youths were traced through the police, courts, welfare, and corrections departments to identify all criminal justice actions and services recorded in each case history. Unit costs were obtained from business offices where possible, calculated from budgetary and work-load data where necessary, or, as in one instance, adapted from the table of costs in the Los Angeles County report on the cost of a juvenile gang.

The summed costs for the 25 youths, whose criminal histories extended back nine years on the average, were about $842,000. This worked out to approximately $31,000 per youth, or about $3,444 per youth-year.[2] These youths were older than the Los Angeles gang members; they were farther along in criminal histories, had been maintained in more expensive facilities, and the costs were current as of a later date than those of the gang members; hence, the higher year-cost per case.

B. Cost Comparisons

Cost comparisons bring together two or more cost analyses and thus provide a basis for an evaluative judgment. The judgment may take several forms, including "The cost to the community of managing these offenders is less by Method A than by Method B." The comparison shows the magnitudes and differences in costs of the two or more methods.

1. *The Saginaw Project (1957–61).* This project grew out of a belief that an increased use of probation would reduce the need for commitment to state prison and would thus save prisoner maintenance costs and county welfare costs for the families of prisoners.[3] Saginaw County (Michigan) added three probation officers to its existing staff (at that time, three officers, thus doubling its staff size). It then kept records as to dispositions of cases that came before the court after expansion of services. Probation, jail and prison dispositions during a three-year period were compared with similar dispositions during an equivalent period before the expansion of probation.

The analysis of the project was by staff members of the University of Michigan School of Social Work.[4] The results for the three-year period (compared with an equivalent three-year "before" period) were as follows:

- Reduction in prison costs $296,560
- Reduction in parole costs 49,280

• Avoidance of welfare costs	78,594
Total savings	$424,434

The reductions in costs were presumed savings made on 88 dispositions that were estimated to have avoided prison as a result of the augmentation of probation staff. The staffing cost was approximately $100,000 over the three years; the alternative costs without the staffing would have been about $525,000, hence, the $424,434 savings.

Since this was a before-after study that was cruder in design than a quasi-experiment, the reliability of the estimates is a matter for speculation. It should be noted also that jail costs, which increased as a result of the reduction in prison terms, were not explicitly accounted for in the costing process. The rationale given was that some of the jail terms were very short, and persons who finished jail terms often did not receive probation supervision afterward, thus not incurring probation costs.

The cost comparison—$100,000 with probation increase, $525,000 without probation increase—does not go into the matter of benefits such as earnings during the time the probationer might have been in prison without the program. Thus the analysis is a cost comparison rather than a cost-benefit analysis, strictly speaking.

However, if averted costs are defined as benefits, a first-year benefit/cost ratio of about 5 to 1 may be estimated for the project. This is unusually high for correctional projects, but it is a plausible figure because the primary function of the project was the elimination of expensive institutionalization.

2. *The PICO Project.* The PICO Project ran for six years (1955–61) and processed about 1,600 older Youth Authority wards through a four-way controlled experiment to test the efficacy of individual psychotherapy with such wards.[5] During the first two years of the project, an amenability-nonamenability classification divided incoming wards into roughly equal halves. The two types of controls received the regular institutional program of academic education and vocational training. The experimental amenables and experimental nonamenables received the regular program plus twice weekly individual therapy, and some group therapy. One therapist carried a caseload of 25 clients and worked with them an average of nine months.

The experimental results, as disclosed in a 33-month post-release followup, showed reduced recidivism by the treated amenables and *increased* recidivism by the treated nonamenables. In other words, the nonamenables who were randomized into experimental status performed worse upon release than their controls. This apparently indicated that the nonamenables were affected adversely, or disorganized, or harmed by the experiment.

Over the 33 months of followup that occurred in the study, the experimental amenables averaged 2.1 months back in California prisons, the control amenables 4.8 months, the control nonamenables 4.8 months, and the experimental nonamenables 5.5 months.

The cost of the additional treatment, beyond the regular institutional programs, can be estimated from the therapists' caseloads and salaries. With caseloads of 25 that were carried for an average of nine months, a therapist could treat 33.3 cases in a year. If he were treating "amenable" cases, he would save his clients 33.3 × 2.7 or 89.91 months of prison time over the first 33 months after release. If he were treating nonamenables cases, on the other hand, he would increase their prison time about 23 months over the amount expected. He would thus be generating a loss for his clients and for the correctional system.

The cost of a therapist-year of treatment and the savings or losses over a 33-month period for three types of caseloads can be summarized as follows:

Type of Caseload	*Therapist's Salary (Assumed)*	*33-month Savings*	*Losses*
Amenables	$15,000	$20,964	-------
Mixed	15,000	11,640	-------
Nonamenables	15,000	-------	$24,324

The savings of $20,964 were arrived at by estimating the cost of 89.91 months of prison at $400 per month and subtracting the $15,000 therapist's salary from the estimated value, $35,964. Another way of stating this is that the prison can either spend $15,000 per year on a therapist or an estimated 2.4 times that amount because it has not provided a therapist for its amenable cases. As for the nonamenable cases, the prison administrator would be very well-advised to keep such cases out of individual psychotherapy.

It should be noted that this presentation is rather elementary, disregarding the matters of discount rates and marginal costs. These are points for later discussion. It will also be noted that this kind of cost analysis does not include benefits from increased earnings after release or other kinds of benefits apart from prison time avoided. A full accounting of costs and benefits would have required a broader range of data then were collected for the PICO Project.

It may be concluded that the PICO Project demonstrated rather definitively that individual psychotherapy in prison, with older youth (or young adults), has a significant effect upon the "amenables" among such youth, and that the operationalization of the project would have been advantageous to the Department of Corrections in a monetary sense.

As matters developed, the Department failed to discover the effectiveness of the project because of a masking problem that appeared in the experiment after the second year. The results were reported to be nil, and the experiment was essentially written off as "inefficacious." In 1960 there was a reanalysis of some of the data, the masking effect was discovered, and the significant findings were reported out. However, the opportunity appeared to have been lost, and the individual therapy program was never operationalized. Why this occurred in a project that was planned and executed by in-house practitioners and that yielded outstanding results in performance modification as well as attractive economic gains should interest all proponents of research utilization.

3. *The Group Guidance Project.* In 1965 the Group Guidance program of the Los Angeles County Probation Department was seriously challenged by the Police Department of Los Angeles, which claimed that the program encouraged rather than reduced delinquency among gang members. The program was first suspended then resumed pending an evaluation. One evaluation was presently underway in the Youth Studies Center of the University of Southern California under a foundation grant. The YSC evaluation was initially planned as an experiment, with control and experimental gangs and areas; in its final form, it might best be described as an experiment in which the gang under Group Guidance served as their own controls.[6]

A second evaluation was undertaken by the Probation Department's research office, on an emergency basis, using a quasi-experimental design, and emphasizing costs as a criterion. For this evaluation, three roughly comparable gangs were identified from the Department's area studies. One gang had had full Group Guidance service several years previously, another had had partial service, and the third had had no service. The study was wholly retrospective, and could be concluded as rapidly as the data could be collected from the records and analyzed.

The data were abstracted from police, probation, and Youth Authority records. Criminal histories and criminal career costs of the core members of the three gangs were reconstructed. The histories in each case were followed for a span of six years, which was estimated to cover the gang career of the typical core member. The six-year span was chosen so as to be split at midpoint by the date of inauguration of treatment for the two treated gangs. This provided "before" and "after" segments of gang careers, before treatment and after treatment, with the control gang careers running parallel for the six-year period.

The cost of criminal justice management of the average member of each of the three gangs, before and after the treatment inuaguration, is shown in the following tabulation:

Average Criminal Career Costs before and During Treatment

Gang	*3 Years Before*	*3 Years During*
Untreated	$2,934	$4,576
Partially Treated	3,695	2,601
Fully Treated	3,944	2,345

Although the initial desire of the researchers had been to find three gangs that were quite similar in all respects except for having been involved in supervision by Group Guidance workers, this proved impossible. Service went usually to the most troublesome gangs. On the basis of the pre-service criminal career costs, one would nominate the third gang as the one most likely to come to the attention of Group Guidance, and that is in fact what happened.

Observing that the untreated gang members incurred about $1,600 more in criminal justice costs in the second three years than in the first, while the fully treated gang incurred about $1,600 *less,* we can infer a reduction of about $3,200 in criminal justice costs per gang member in the after period as a result of the application of Group Guidance service. For the 43 core members of the fully treated gang, the exact amount comes to $139,263 over three years, or a reduction of about $45,000 per year. This reduction was achieved by the application of one-half the time of a Group Guidance worker to the fully treated gang.

The $45,000 reduction in criminal justice costs must be seen as the consequence of spending one-half the salary of a group guidance worker. Assuming the worker's salary to be about $15,000, we have a treatment cost of $7,500 and a saving of $37,500 for the year. Thus, the County could weigh the desirability of investing $7,500 per year for street worker service, or paying $45,000 for the criminal justice costs of gang activity that the worker might avert.[7]

It is of some interest here to contrast the cost analysis with the behavioral analysis of the three gang performances. If performances are compared in terms of arrests, juvenile hall detentions, probation supervision, camp stays, Youth Authority stays and Youth Authority parole, the data distribute over as few as 12 and as many as 36 cells, depending on how the performance categories are grouped. Data of this kind are virtually impossible to manage for 100 cases under the conventional rules of statistical inference. Judgmentally, they appear to show "superior" performance and higher levels of "freedom" for the untreated gang in the before period and a reverse of these conditions in the after period. Monetarily, they show a marked increase in the efficiency of gang control as a result of modest investments in street workers.

Since these results conflict with the findings of the Youth Studies Center research, a word on this matter is pertinent. Either or both of the studies may be in error. Also, both may be correct. The two studies used different bodies of data, from different sets of gangs, recorded at different times in the history of Los Angeles. In each instance, the quality of the data, the data collection and analytical procedures, and the impact of historical or external events could have invalidated one study without invalidating the other. On the other hand, one or more of these factors may have worked to the serious detriment of both studies. Replication of one or both studies could help resolve some of the questions of credibility that have been raised by the conflicting findings.

Both studies could be correct since neither had major external validity. Only one street worker was involved in the Probation Department study, and he was recognized as a "genius" in handling gang members by even the Police Department, who wished to see the street worker program disbanded. Since there was wide variability in worker capability, two narrowly focused studies could correctly report different results by incorporating, on the one hand, effective street workers, and on the other, ineffective workers or dysfunctional study procedures. What actually occurred in the two studies remains to be ascertained, although it is of interest that the YSC study concluded that its project design had inadvertently increased the level of gang delinquency in the groups studied.[8]

As for the Probation Department study, the fact that its street worker was an unusually capable individual raises questions as to how widely the results may be generalized. Low generalizability to other street workers or to other communities would reduce the significance of the monetary findings. However, the data suggest provisionally that there is at least a moderate claim to both internal and external validity in the study.

4. *The Parole Work Unit Program.* This program began in 1965 as a test of small parole caseloads, adjusted in size to meet parolee needs for supervision or support, against the larger, traditional caseloads. The results for the first six months appeared unfavorable to the Work Unit principle, but changes in decision-making procedures and use of community resources shifted the results of the project. In the small caseloads, rates of return to prison gradually declined. The rate of return at two years after release dropped to about 21 percent in 1973 as compared with an average rate of 44 percent in the four years (1961–64) before the project began.

Paradoxically, the large caseloads showed about the same drop in rate of return as the small caseloads, suggesting that the significant factor may have been neither caseload size nor "supervision need" but the manner in which parole agent decisions were made.

The drop in rate of reimprisonment meant a reduction of an estimated 5,827 man-years of prison time, equivalent to savings of $23,000,000, during 1966–73. The savings per man-year were $4,601, or the difference between the cost of incarceration ($5,246 per man-year) and the cost of parole ($645 per man-year).[9]

Since this saving was accomplished by 1,513 years of parole-agent effort (an average of 190 parole agents per year over eight years), we can estimate cost savings per agent-year at $15,200.

5. *The Silverlake Experiment.* This experiment was a four-year inquiry into the effects of a "therapeutic milieu" on delinquent boys who ordinarily would have been placed in a private residential facility. Assignment to Silverlake house, in Los Angeles, and to Boys Republic, outside the city, was random. Lengths of stay were about six months in Silverlake and about 13 months in Boys Republic.[10]

Evaluation of the project rested on comparisons of experimentals and controls on a) relative frequency of arrest for new offenses during the first year back in the community, b) seriousness of new offenses committed, and c) relationship of arrests in the before period to those in the equivalent after period.

The two groups did not differ significantly in frequency of arrests in the post-release period, nor were there differences in before-and-after patterns of arrest. Experimentals appeared to become involved in slightly less serious offenses in the post-release period as compared with the controls.

The major finding in the experiment was that the community saved about $2,000 per delinquent boy by processing him through Silverlake rather than through Boys Republic. Silverlake is thus a notable example of a project in which there is a noteworthy monetary gain even though there is no appearent improvement in the behaviors of the treated subjects.

6. *Saginaw to Silverlake: Cost Reductions per Therapist-Year of Effort.* It is interesting to compare the cost savings per therapist-year of effort in the five projects just examined. The meaning of the comparison is not wholly clear, but it leads to some useful speculations, and perhaps in time to productive inquiries.

The cost savings are based on the data given in the foregoing discussion and do not cover all possible benefits from the projects. It will also be noted that three kinds of occurrences were evident in the projects: behavioral changes, personality changes, and decision changes. The following tabulation lists the five projects, the reported cost savings per therapist year, and the kinds of changes that took place in the projects.

Project	*Approximate Cost Savings per Therapist-Year*	*Kinds of Change*
Street Worker (Juvenile Gang).	$75,000	Behavioral
Saginaw Project	47,000	Decision
PWUP: Parole Work Unit Program.	15,200	Decision
PICO Project: Mixed Cases.	12,000	Personality
Silverlake Project	8,000	Decision

To make clear the distinction between decision changes and personality or behavioral changes, it may be noted that the Saginaw Project featured decisions by the court to place some men on probation who would in earlier years have been sent to prison. There is no evidence that the savings in any way required a personality change or fundamental behavioral change on the part of the offenders. In the PICO Project, the nine months of intensive psychotherapy presumably resulted in personality changes in at least some of the experimental cases—favorable changes in the amenables and unfavorable changes in the nonamenables. These changes, one might infer, led to behavioral differences that were reflected in the different rates of return to prison.

In the Silverlake Experiment, since there appeared to be no appreciable behavioral changes in the experimental subjects, the project's main effect was cost savings from allocation of subjects to the more economical Silverlake program.

Behavioral changes in the street gang members in the Group Guidance study may be assumed from the apparent reduction in involvement with the police, courts, and corrections. Personality changes may have occurred, but they were not as directly sought as in the case of the PICO Project, and they would have to be inferred from the improved behaviors of the gang members.

C. Cost-Benefit Comparisons

A third type of program evaluation in which the monetary criterion is important is that in which benefits as well as costs are examined. This type of analysis is possible when there are benefits other than cost reductions and data on the benefits are available or can be reliably estimated.

1. *Cost Analysis of a Work Release Program.* The D.C. Department of Corrections began a work release program in 1966, basing it first in a countryside prison, then moving it to a jail in D.C., and finally lodging it in a scattering of halfway houses in the inner city.[11] The principal operating costs were for staffing the work-release unit; the principal benefits, the earnings of the work releasees. Because work releasees were maintained under lower security than other prisoners, staffing costs were lower.

During Phase I (the prison phase), the cost of maintaining an average of 54 felons and 44 misdemeanants was $104,209. Compared with regular prisoners, this was a saving of $59,548 to the Department. During Phase II, the program was housed in the Jail, with an increased but still relatively small staff. For this period there was a saving of $109,342.

In Phase I the work releasees returned $94,006 to the government in payment for lodgings and taxes. This came out of total earnings of $387,376. They also paid $74,104 to dependents and to creditors in the community and retained $219,265 for their own use in the program and after release.

The average return to the government from these earnings was $769 per felon work-release-year and

$580 per misdemeanant work-release-year. Returns to government and community (dependents and creditors) together totaled $1,363 for felons and $1,071 for misdemeanants per participant year. With the addition of departmental cost reductions, total benefits from work release came to $2,315 per felon-work-release-year and $2,023 per misdemeanant work-release-year. These figures do not include benefits to the economy or to the individual from monies retained by the work releasee.

During Phase II, the work releasees earned a total of $370,524, of which $125,559 went to the government, $59,873 went to creditors and dependents, and $185,091 was retained for personal use. The average return to the government alone per man-year in the program was $789 in Phase II. The average return to government and community together was $1,166 per man-year. With reduction in departmental maintenance costs added, the benefits from work-release to persons other than the participants came to $2,363 per program-man-year. The cost to society of maintaining an inmate in the Department of Corrections thus appeared to be reduced about one-half by placing him on work release rather than in an institution.

These calculations omitted any benefits that might have accrued in the future because of increased earnings or reduced recidivism after graduation from work release. Additional data would be required to ascertain whether such benefits did in fact occur because of work release.

Future Benefits. The possibility of future benefits was not examined in two of the examples previously cited—PICO, or the Group Guidance Project. In the Silverlake Experiment, the absence of differences in performance between experimentals and controls suggests that there may be no benefits beyond the reduction in costs in Silverlake.

There are programs, however, in which there appears to be a strong likelihood of benefits that extend into the future. A full accounting of the monetary values relating to such projects requires some consideration of future benefits. In the Group Guidance Project, interviews with ex-gang members yielded information about value systems and modes of existence that suggested enduring continuation of project benefits. The same tendency might be inferred from PICO, but on different grounds: the treated amenables showed rapid improvement over their controls in avoidance of lockup during the first year out, and a slower but continuing improvement until the end of the 33-month followup period. This implies that here could be at least a modest expansion of the recidivism rate differential for an indefinite period of time.

The wisdom of predicting further gains, or even the maintenance of present gains, beyond the end of a specified period of followup is not clear. Experience shows both types of outcomes in well-known studies. The PICO differential was growing at 33 months; the CASE differential had shrunk to nothing at the end of three years. The Fricot gains were wiped out in three years; the Provo gains were continuing at the last point of followup. The frequency of the CASE and Fricot types of outcome suggest that there is considerable risk in predicting gains for more than a year or two into the future, hence, the estimation of future benefits in corrections needs to be done with caution.

2. *Costs and Benefits of Narcotic Addiction Treatment.* One example of estimation of costs and benefits of treatment is an analysis of the economic consequences of addict treatment in the District of Columbia.[12] The treatment program grew out of research and planning in the D.C. Department of Corrections and so is relevant to the subject matter of this Guide.

The analysis was undertaken in 1970, several months after the DCDC community-based centers for narcotic-involved offenders were taken over by the District and expanded. At that time, using overdose deaths as a clue, it was estimated that there were between 10,000 and 20,000 heroin addicts in the District.

If there were 15,000 addicts, using heroin at an average cost of $40 daily, and assuming that one-fifth of the supply came through street pushing, the total cost of heroin purchased yearly in the District would be $175,000,000. At the same time, the annual cost of crime to victims of larceny (including shoplifting) was estimated at about $160,000,000, using reported offenses and adding by formula for unreported offenses. If addicts were responsible for about one-half the property crimes, the realized monies (after fencing, street sales of merchandise, etc.) would obviously be too low to cover the estimated heroin costs. Consequently, other sources of funds, including welfare, prostitution, and earnings from legitimate employment were used in estimating costs.

Benefits of successful treatment were calculated for the following categories: drug purchase monies freed for other uses; earnings increase; and police, court, corrections and parole costs averted. Assuming 400 successes in one year of treatment of 1,000 cases, the benefits accruing in one year would amount to $5,750,770. The cost of treating the 1,000 cases would be $1,400,000 given the 1969–70 costs.

The benefit/cost ratio ($5,750,770/$1,400,000) for one year would be 4.1 to 1, assuming $40 per day per addict for heroin costs and 50 percent of property crime arising from addicts. If the assumptions were changed to $25 per day for drugs and 33 percent of property crime arising from addicts, the one-year benefit/cost ratio would decrease to 2.7 to 1. Under the latter assumptions, and with specified rates of readdiction of treated addicts in the first seven years and stabilization thereafter, the 11-year benefit/cost ratio would be 8.5 to 1.[13]

These conclusions obviously depend upon assumptions about success in treatment. Knowledge of success under the newer programs is limited, although Gearing reported a 40 percent success rate in one addict treatment program that was similar to the District's program in extensive use of methadone maintenance.[14] If the success rate were to fall to 100 in 1,000, or 10 percent at the end of the first year, the benefit/ratio would be $1,437,680/$1,100,000, or about 1.3 to 1. The 11-year ratio would be appreciably higher.

These values assume a 10 percent discount rate—that is, a "marking" down of future benefits in acknowledgement of the fact that commitment of present monies for future benefits requires an incentive.

3. *Costs and Benefits of Pretrial Diversion.* A diversion demonstration project in the District of Columbia courts was initiated in early 1968 and subjected to evaluation a year later to ascertain whether it was effective and efficient.[15] Effectiveness was measured by means of a quasi-experiment, which compared participants and "controls" on arrests, employment, and earnings one year after leaving the program.[16]

Efficiency was measured by comparing the economic benefits of the program, present and future, with the costs of the program.[17] Benefits were estimated for the following: savings through diversion from court appearances, trials and incarceration; savings through future recidivism reduction; and savings through earnings enhancement. Recidivism and earnings benefits were projected about five years ahead of the termination of followup.

The results, with future benefits discounted at 10 percent, were as follows:

Benefits from diversion	$104,994.77
Benefits from earnings	170,729.00
Benefits from recidivism	198,448.07
Total benefits	474,171.84
Total costs	223,256.00

Benefit/cost ratio: 2.1 to 1

The cost estimates were made with research funds excluded and 25 percent of the project administrator costs considered as inapplicable (because they pertained to a youth phase of the project, not included in the present analysis). The cost per month per adult came to $194.38; the project cost per adult participant amounted to $506.52.

Education benefits had been considered for inclusion in the benefit estimates but were finally judged to be too small and also difficult to estimate. Other potential benefits not included in the estimates, and which might be inferred from a projected reduction of recidivism, are reductions in welfare costs, reductions in private crime prevention equipment and manpower, reductions in community disruption, and so on.

The earnings benefits from Crossroads were assumed to continue over a five-year period, diminishing to equality between participants and hypothetical controls at that time. Benefits from recidivism averted were calculated by use of assumptions about future crimes by the participants and controls and about the police, court, and prison costs associated with those crime rates. Diversion benefits were based on assumptions about the costs of police time, court costs, and incarceration costs averted through success in completing the program and bypassing adjudication.

The Crossroads cost-benefit analysis rests on a complex set of assumptions in a number of areas, including the future performance of hypothetical comparison group members. As a result, many observers may be inclined to accord the conclusions of the analysis rather low credibility. This is in many ways a defensible judgment. It should be noted, however, that the 2.1 to 1 benefit/cost ratio derived for Crossroads is relatively modest compared with the 6 to 1 ratio that may be derived for Miguel Duran's work with the Spartan Gang in Los Angeles or the 5.3 to 1 for the Saginaw Project. In each of these cases specifiable benefits have been excluded from the estimates and all the ratios may be safely conservative. The possibility that they may all be shown to be highly credible should come as no surprise in an activity as inefficient as present-day corrections or criminal justice.

4. *Cost-Benefits in Prison College Programs (1973).* In a recent evaluation of five NewGate and

five non-NewGate prison college programs, recidivism, social impact, and benefit/cost ratios were used as methods of ascertaining effectiveness and inefficiency.[18] Recidivism was found to be no lower among program participants than among comparison group members. Social impact was inferred to be positive but the conclusions were on wholly qualitative bases. Benefit/cost ratios were computed by focusing on taxes paid by recent participants and non-participants, projecting the differential for twenty years, and comparing the twenty-year tax differential with the costs of the college programs to the community.

The assumption was made that that the increments in federal income taxes paid annually by persons finishing one, two, three or four years of college over those paid by persons with one or less years of college would provide the best base for estimating benefits from the programs. Calculating the annual increment for the mix of college years completed by 100 parolees from each of the programs, multiplying by 20 (years) and comparing with the cost of educating 100 men in each of the programs disclosed the amount of gain or loss at 20 years.

The results of the comparison are shown in the tabulation on the page following. Data were lacking on two of the ten programs hence only eight comparisons were made. Instead of discounting benefits at 10 or 15 percent, the tabulation shows the gains in undiscounted values and also in terms of the compound interest rate that would generate the gain in 20 years.

Prison College Program	*Cost per 100 Prisoners*	*20-yr Tax Increment*	*Gain or Loss*	*Rate of Gain*
Ashland*	365,123	50,590	−314,533	−9.4%
Minnesota*	326,742	480,760	154,918	1.9%
New Mexico*	86,156	247,880	161,724	5.4%
Oregon*	216,566	495,920	279,354	4.2%
Pennsylvania*	401,451	640,120	238,669	2.3%
Lompoc	21,641	276,740	255,099	13.6%
Illinois	33,945	376,520	342,575	12.8%
Texas	53,325	62,810	9,485	0.8%

*NewGate programs

This approach to cost-benefit analysis probably underestimates seriously the benefits that derive from the college programs. However, the basic data required for other approaches are lacking, so this is perhaps a reasonable strategy under the circumstances. The DCDC evaluation of its own prison college program showed recidivism rates definitely favoring the program participants and benefits from recidivism reduction alone approximately equal to costs.[19] With additional benefits from increased earnings and lower claims against welfare and health services, the DCDC benefit/cost ratio should be higher than most of those shown in the tabulation.

D. Discussion

In a society where most social action programs are seen as impacts upon resources as well as upon persons, cost-benefit analysis is likely to become increasingly important in agency evaluations. The use of this form of analysis gives the agency administrator a more precise and more convincing language for the support of productive programs. It gives the evaluator a versatile and powerful technique that greatly increases his ability to assess processes and programs.

Cost-benefit analysis is at once open-ended, exploratory, and demanding. It is open-ended because the benefits that stem from rational choices of dispositions and from behavioral or personality changes in offenders are numerous. The evaluator may draw on few or many of these possibilities in planning his evaluation strategy. Open-endedness marks the technique both across possibilities and through time.

The cost-benefit technique is exploratory because of its newness and the dearth of exemplary cases and recognized guidelines. Its exploratory nature makes it well suited to application in a social action field that is in a state of transition. It also makes it especially useful to evaluators who are resourceful and uncommitted to evaluation's more traditional methods.

Cost-benefit analysis is demanding because it requires new learning, provides few clear-cut rules, and often pushes for heavy reliance upon assumptions and relatively free judgments in coming to conclusions.

The technique is most easily applied in conjunction with experimental or quasi-experimental research designs and with costs and benefits calculated only in the period of follow-up. It becomes more difficult to apply when projections of costs and benefits are made and when the comparative costs and benefits have to be determined by estimation, particularly when the estimates require complex systems of assumptions.

REFERENCES AND NOTES

1. Borden Olive, Roger E. Rice and Stuart Adams, *The Cost of a Juvenile Gang: A Nine-Year Follow-up of Law Enforcement, Court, and Correctional Expenditures*. Los Angeles: L.A. County Probation Department, Research Report No. 16, October 1964.
2. Barbara Cantor and Stuart Adams, *The Cost of Correcting Youthful Offenders*. Washington, D.C.: Department of Corrections, Research Report No. 6, September 1968.
3. Michigan Crime and Delinquency Council, *Saving People and Money: The Saginaw Project*. East Lansing, Michigan, January 1963.
4. A. C. Ball, R. C. Fletcher and E. Cranefield, *The Saginaw Probation Demonstration Project*. Ann Arbor: Michigan School of Social Work, 1963.
5. Stuart Adams, *Effectiveness of Interview Therapy with Older Youth Authority Wards*. Sacramento: California Youth Authority, Research Report No. 20, January 1961.
6. Malcolm W. Klein, *Street Gangs and Street Workers*. Englewood Cliffs: Prentice-Hall, 1971, p. 233.
7. Stuart Adams, "A cost approach to the assessment of gang rehabilitation techniques," *Journal of Research in Crime and Delinquency*, 4 (January 1967), 166–82, p. 181.
8. Malcolm W. Klein, *op. cit.*, p. 137.
9. Parole and Community Services Division, *The 1973 Report on the Work Unit Parole Program*. Sacramento: California Department of Corrections, March 1974.
10. LaMar T. Empey and Steven G. Lubeck, *The Silverlake Experiment*. Chicago: Aldine, 1971.
11. Virginia A. McArthur, Barbara Cantor and Sara Glendinning, *A Cost Analysis of the District of Columbia Work Release Program*. Washington, D.C.: D.C. Department of Corrections, Research Report No. 24, June 1970.
12. John F. Holahan, *The Economics of Drug Addiction and Control in Washington, D.C.: A Model for Estimation of Costs and Benefits of Rehabilitation*. Washington, D.C.: D.C. Department of Corrections, November 1970.
13. Ibid, p. 71.
14. Frances R. Gearing, "Successes and Failures in Methadone maintenance treatment of heroin addiction in New York City." Presented at the Third National Conference on Methadone Treatment, November 14, 1970.
15. Leon G. Leiberg, *Project Crossroads: Final Report*, Washington, D.C.: National Committee for Children and Youth, 1971.
16. Roberta Rovner-Pieczenik, *Project Crossroads as Pre-Trial Intervention: A Program Evaluation*. Washington, D.C.: National Committee for Children and Youth, 1970.
17. John F. Holahan, *A Benefit-Cost Analysis of Project Crossroads*. Washington, D.C.: National Committee for Children and Youth, 1970.
18. Keith Baker, John Irwin, Steven Haberfield, Marjorie Seashore and Donald Leonard, *Summary Report: Newgate and Other Prison College Education Programs*. Washington, D.C.: Office of Economic Opportunity, November 1973.
19. Colleen Barros, William Caldwell and Stuart Adams, *Effectiveness of the Lorton Prison College Project: Third Interim Report*. Washington, D.C.: D.C. Department of Corrections, Research Memorandum 72–4, May 1972.

CHAPTER 13. OPERATIONS RESEARCH AND SYSTEMS ANALYSIS

Project evaluation, program evaluation and process evaluation give a piecemeal view of how things are going in a correctional agency. These narrow perspectives leave executive staff ill-prepared for many aspects of the decision-making process. A more global view of what is going on is desirable, and to achieve this global view a different evaluation strategy is required. We need approaches and techniques such as operations research and systems analysis.

These two terms are often regarded as synonymous. The two approaches in fact overlap considerably in method, and there is some justification for considering them the same. However, there are some clear distinctions between the two, and for the purposes of this Guide, it is important to preserve those distinctions.

A. Definitions

Operations research focuses on the description and analysis of an ongoing system. Its purpose is to "optimize" or to make the best use of processes, people, materials and resources already in existence and at hand.[1]

Systems analysis, to the extent that it differs from operations research, is concerned with alternative means of achieving objectives. It uses real and hypothesized assessments of performance, costs, and risks to aid the administrator in deciding how a system should be structured and operated.[2]

If one wished, he could incorporate operations research within systems analysis, since both use the same techniques, and both may employ mathematical equations and computers. However, the concept of operations research as stated here is more descriptive of the needs of self-contained organizations that are concerned primarily with upgrading their internal performance. Systems analysis is more relevant to the needs of changing organizations in a transitional situation where comparison among systems is important.

Operations research is an appropriate evaluation stance for an agency that is seeking to routinize its evaluation as proposed by Glaser,[3] while systems analysis may be a more suitable strategy for an organization whose philosophical orientation is toward new ways of functioning. The two approaches are illustrated to some extent by the research strategies of the California Department of Corrections and the California Youth Authority, respectively.

While both of these approaches are frequently defined in terms of advanced skills and methods such as mathematical equations, cost-benefit analysis, computer applications, and elaborate statistical analyses, each also can be carried on in simple forms. Each has its conceptual and arithmetical examples as well as its computerized and algebraic examples.

B. Operations Research: Elementary Forms

In its broadest sense, operations research is concerned with the measurement of input, process and outcome and with understanding their interrelationships for purposes of control and improvement. For the correctional administrator, the primary input is a heterogeneous mixture of offenders, varying over time in offense type, age, socioeconomic and ethnic backgrounds, previous involvement in crime, and attitudes toward authority. The input also includes physical facilities, staff, equipment, and supplies. The processes include work, job training, education, counseling, incapacitation, depersonalization, and punishment. The outcomes, by intention, are law observance, social conformity or adjustment, economic productivity, self-esteem, and capacity for personal growth.

Operations research has the objective of measuring and analyzing these quantities, qualities and events to provide bases for decisions about continuation or change of the various aspects of agency structure or procedures.

Such operations research as has been done in correctional systems in the past two or three decades has focused primarily upon outcomes. Both agency staff and the general public show constant concern about whether the correctional output is rehabilitated. Much discussion of recidivism rates has been heard, in public places, in legislative halls, and in the executive offices of correctional agencies, and this particular aspect of measurement has dominated research in corrections.

1. *California Department of Corrections.* One SCA that has made a persistent and massive attempt to achieve an operational research capability is the California Department of Corrections. Since 1945 it has reported some aspects of this attempt in its publication series, *California Prisoners*. The Foreword of a recent edition of the series conveys a sense of the effort:

> "The Department of Corrections population of 28,642 persons on December 31, 1968 was principally felons and persons in the civil narcotic program but also included approximately 1,250 Youth Authority Wards.
>
> "The Department has industrial programs to teach work habits, vocational training to prepare a person for gainful employment, academic instruction, occupational therapy, parole or outpatient casework, community correctional centers or halfway houses and work and training furlough programs. All planning and programs are to assist men and women to change their behavioral patterns and attitudes so that they will not offend or use narcotics again.
>
> "This edition of California Prisoners presents a comprehensive statistical picture of the State prison and parole programs for the calendar year 1968. Such reports have been published since 1945. The tables, text and charts cover a variety of administrative measures. This book has resulted from an ever-improving system of statistical record keeping and reporting.
>
> "California plans to continue to refine and expand the administrative statistical system in order better to measure the results of the correctional work in this State and to develop more effective methods for the control and rehabilitation of offenders." [4]

The areas of concentration of this 150-page volume can better be understood by a glance at the table of contents. The headings are as follows:

- Institution population and movement
- Felons newly received from court
- Characteristics of the population in prison
- Felons released from prison
- Felon parole population and movement
- Characteristics of felons on parole
- Parole suspension
- Reinstatement
- Felons discharged from parole
- Felons returned to California prison
- The work furlough program and the community correctional centers

California Prisoners describes offender characteristics and outcomes. To obtain information on processes, one would have to turn to the many research reports the Department's research division has produced in the past fifteen years.

This effort and the reports that describe it are only a small part of the information and analysis that are required for a comprehensive operations research activity. To some extent this reflects the fact that no organization can do all the research it might conceivably need. Priorities are involved. Furthermore, some aspects of process are not present because they are taboo in one sense or another. The punishment and depersonalizing processes are not included in evaluation possibly because they are not "positive" elements in rehabilitation, but also because they are hard to quantify or operationalize, and perhaps also because they are not yet accepted as proper subjects for evaluation.

What effect the omission of such processes from evaluation has upon the results of evaluation is not clear. The effects of punishment and incapacitation probably vary by offender personality and career type, and some types probably are either influenced positively or are not affected by these aspects of process. Others, it may be presumed, are adversely affected, and on a cost-benefit basis at least, are poor prospects for improvement precisely because of incarceration. Whether the overall effect on the prison population is positive or negative, we can only surmise. There is a body of opinion that assertedly rests on empirical data and concludes that the net effect is unfavorable. Whether this is demonstrable in a rigorous sense, and whether the difference is large enough to be practically important will require further analysis to determine.

It is clear that the California Department of Corrections has not yet reached the stage where its entire operation is described by mathematical equations and can be fully computerized. Nevertheless, there have been a number of studies involving the Department or its data system that constitute operations research in a relatively advanced sense. One was a "space age" report on the prevention and control of crime and delinquency in California.[5] Another was an analysis of the interaction of prison, parole and offender characteristics in Department

of Corrections releasees.[6] Yet another was a study of changes in prison admissions in California and their implications for the correctional program.[7] Still another simulated the California criminal justice system on a computer to make corrections cost projections.[8] Finally, there is an ongoing study in which a simulation model has been developed to aid in projecting the California prison population.[9] These studies fall largely under what we prefer to call systems analysis and simulation and will be taken up in later sections of the Guide.

2. *The District of Columbia Department of Corrections.* While the California Department of Corrections approached the task of operations research through the development of an elaborate data system first and a research capability later, the D.C. Department of Corrections reversed this process. In 1966 DCDC received an OLEA grant to establish a statistics and evaluation unit. This would enable it to replace its manual record system with something more adequate and begin to evaluate its programs. In 1967 the department hired three researchers, two of whom were at the entry level, and a systems analyst who had the task of building a data system. In early 1968 the analyst was replaced by an electronics data processing administrator.

During the next five years the Department increased its research and data processing staffs and moved ahead with both evaluative research and a computerized management information system. The research consisted mainly of small-sample studies, aimed primarily at measuring outcomes of the total release population and of subpopulations by program type and releasee type. A few studies focused on costs and benefits; some analyzed intake and made population projections.

The research strategy was in large part a commonsense variety of operations research, seeking to relate behavioral (return to system) outcomes and cost outcomes to process. Following are selected items from the studies in these areas:

a. *Post-Release Performance of Reformatory Releasee.*[10] A three-year follow-up of 432 offenders released in 1965 from the major institutional unit of the Department showed 25 percent returned to the system for terms of 30 days or more by 36 months after release. At 12 months after release, the return rate on the same criterion was 7 percent. The 36-month rate was apparently distorted upward by the inner-city riots of early 1968.

b. *Community Performance of Three Categories of Institutional Releases.*[11] When the 432 releasees of 1965 were separated into the release categories of parolee, conditional releasee and expiree (or dischargee), the return-to-system rates (for 30-day or longer terms) at 36 months were 11 percent, 26 percent, and 38 percent, respectively. At 12 months the rates were 3 percent, 7 percent, and 12 percent, respectively.

c. *In-Program and Post-Release Performance of Work Releasees.*[12] A total of 281 felon and misdemeanant work releasees who entered the program during the first 16 months after its inauguration in April 1966 were traced through the program and for 12 months after release from the program or from re-release after in-program failure. Of the 156 felony offenders, 32 percent failed during an 11-week (average) in-program stay; and, of the 156, 26 percent were returned to the D. C. jail for detention until release on bond or until trial during the first 12 months after release from work release or re-release after in-program failure.

The 125 misdemeanant releasees showed an in-program failure rate of 29 percent and a 12-month return to jail rate of 24 percent for the program parolees and re-released program failures combined. The 24 percent rate compares with 15 percent detention at 12 months for reformatory releasees.

d. *Quasi-Experimental Evaluation of Work Release.*[13] After two proposals by research staff for an experimental evaluation of work release had been disapproved by operations staff on grounds of difficulty and inconvenience, a quasi-experimental evaluation of work release was carried out in 1970. This study compared the post-release performance of in-program successes and direct parole releasees from the institution from which the work releasees were drawn.

The work release participants and the comparison group members were matched on several characteristics that were judged to be related to post-release performance. A total of 120 work releasees (WR) and 119 institutional releases (INST) were compared.

To make the comparison more informative, the two groups were subdivided into four groups each by classifying by felony and misdemeanant offense and by presence or absence of a problem (with drugs, primarily, but including alcohol abuse and emotional difficulties). The resulting classifications and post-release failure rates (return to the system within 8 months by parole revocation or by sentence of 30 days or more) were as follows:

	Comparative Failure Rate		% Gain	Superior Release Modality
	WR	*INST*		
Non-Problem Felons	8.6%	21.7%	13.1%	Work Release
Non-Problem Misdmnts	2.7%	14.3%	11.6%	Work Release
Problem Felons	33.3%	47.4%	14.1%	Work Release
Problem Misdemeanants	41.7%	39.1%	2.6%	Inst. Release
Total	18.4%	30.2%	11.8%	Work Release

e. *Performance of Institutional Releasees by Release Category.*[14] In 1971 the DCDC research unit turned again to a study of post-release performance of release types, this time separating releasees into five types as compared with three in the 1968 study. The five types included the three release-from-institution categories and two release-via-work-release categories—the successes and in-program failures. The work-release in-program failures are shown in the tabulation below by their in-program failure rates. The other four groups are shown by their 12-month post-release performance rates (return to system for 30 days or more).

Release Group	*Percent Returned to System within 12 Mo.*
Expirees (dischargees)	45%
Work-release (in-program failures)	39% (3 mo.)
Conditional releases	26%
Institutional parolees	19%
Work-release graduates	17%

The work-release in-program failure rate of 39 percent was calculated against the number of admissions to work release. By 1971 this rate was several percentage points higher than in the 1968 study, possibly because of a rise in drug-related problems among the work releasees in the later study. Similarly, the parolees of 1971 had a one-year return-to-system rate of 19 percent, several points higher than the 11 percent rate of the parolees of 1965—a rise undoubtedly related to the drug epidemic.

The failure rate for work-release graduates was calculated against the number of residents who completed the work-release program. The disparate failure differentials for work releasees and parolees in the 1970 and 1971 studies presumably arose in part because in the latter study the two groups were not matched on performance-related characteristics.

f. *A Prediction Device for Selection for Work Release.*[15] The study of work releasees in 1968 suggested that a sizable proportion of individuals admitted to work release were not likely to be aided by the program, even though they appeared to need help of some kinds. A number appeared, if not traumatized, at least disorganized by the work release experience.

The 1970 study indicated that if in-program failures were combined with post-release failures, the failure rate of releasees who start into the community through work release is much higher than that of comparable institutional parolees. These results show an apparent need for better methods of selecting persons for work release. The pertinent criteria are *need* for the program and *amenability* to constructive influence by the work-release experience.

To help identify types of individuals who seemed the best candidates in the "likely-to-succeed" sense, DCDC, with the assistance of an LEAA grant, secured the services of a private research organization to develop a selection instrument for such candidates. Using about 900 former work-release assignees as a construction sample, the researchers developed an instrument that predicted in-program and post-release success.

The construction technique was known as "direct search in mathematical space." It yielded a selection scale with 17 items for in-program success and a scale of 16 items for post-release success. The in-program scale was validated on a test sample of 200 cases and showed extremely high predictive efficiency—approximately twice that previously reported for instruments based upon discriminant analysis and configuration analysis techniques.[16]

g. *Performance of Successive Youth Center Cohorts.*[17] As in the case of adult offenders, comparisons were made of post-release success of Youth Center cases, whose ages ranged from 18 to 26 and who were released either directly to parole or through a community treatment center. The releases occurred during 1967–69. For analytical purposes, they were grouped by six-month intervals.

The direct parole releasees showed better performance in the community, apparently because the community center cases were higher risks, assigned to the center because of staff judgments that they needed more "support" than direct parole could give them. Of the five release cohorts, the earlier ones performed worse than later cohorts. This was appar-

ently because of unfavorable conditions in the community around the time of release of the first two cohorts—the street disorders of 1968. The actual and projected 36-month return-to-system rates were as follows:

Released	*36-Month Return to System Rate*	
Cohort I (Early 1967)	56%	Actual
Cohort II (Late 1967)	63%	Projected
Cohort III (Early 1968)	35%	Projected
Cohort IV (Late 1968)	33%	Projected
Cohort V (Early 1969)	29%	Projected

In addition to these seven studies, four others that were "operations research" in nature have already been discussed: the cost of correcting youthful offenders, a cost analysis of the DCDC work release program, the costs and benefits of narcotic addiction treatment, and trends in intake of narcotic-involved offenders. Finally, the operations research included a number of population characteristics, movement and projection studies.

These studies had as primary emphasis a focus on outcomes as measured by the criterion of return to the system. Secondary emphases were the relation of outcome to process such as program experience or manner of release, the effectiveness of programs in terms of cost criteria or cost and benefits criteria, and population types and trends.

Taken together, these studies illustrate commonsense approaches to operations research, using conceptual-arithmetical rather than computerized-algebraic methods. There were two exceptions to the general method: the development of the prediction instrument for work-release, and the cost-benefit analysis of narcotic addiction treatment. In each of these cases, the contractor or staff member made use of computers and mathematical models to process the data for the study.

Of these fifteen or twenty operations studies, the narcotic-offender intake analysis was the one that had the most direct and obvious impact on executive staff decision-making. The recommendation that community-based treatment for narcotic addicts be established was acted on promptly and with good success. The population-projection studies and the related recommendations for space management were of considerable importance for departmental decision-making. The cost-analysis of work release, which indicated cost-effectiveness for the program even though recidivism rates were higher than for direct parole, was instrumental in developing support for expansion of the work-release program.

Several of these studies had impact both within and outside the department, even though many of them did not receive distribution beyond the criminal justice and governmental agencies of the District. Such developments, out of the first few years of operations research activities of a new planning and research unit in a correctional agency, suggest that even elementary forms of operations research can play an important part in planning and operations in a modern department of corrections.

Two further points concerning this operations research experience are worthy of comment. First, the development of an unusually efficient prediction instrument in DCDC should have led to utilization of the instrument in the department. The failure of this to occur raises a number of questions: What were all the factors involved in the breakdown of what would initially have been presumed to be a clear pathway to utilization? Would a case study of this particular instance of nonutilization be of value in clarifying the general problem of nonutilization? Was the shift from an experimental to a more traditional administrative philosophy in the department in 1972–73 a major factor in the disregard of the prediction study results?

Second, the active role of research in DCDC during the five-year period 1968–72 and the major impacts achieved by research on the department and some interacting agencies is of interest in view of the fact that the research program was non-experimental in emphasis and operations-research oriented. The amount of impact was considerably greater than might have been predicted were one to start from the pessimistic position that some reviewers of correctional evaluation have taken in recent years. Furthermore, the studies that achieved the heaviest impacts were time-series studies, cost studies, and quasi-experiments, in diminishing order. The department's one controlled experimental study, now in its fourth year, has as yet had no marked impact on general departmental policy, although it may have aided the survival of the operating program to which the experimental component is attached. These observations suggest that non-experimental studies of the contemporary variety—operations research and systems analysis—like the older types of non-experimental studies, fit the needs of correctional decision-making better and promise higher levels of impact than the traditional "strong design" studies with elaborate experimental or statistical controls.

C. Systems Analysis

A shorthand definition of operations research is

ently because of unfavorable conditions in the community around the time of release of the first two cohorts—the street disorders of 1968. The actual and projected 36-month return-to-system rates were as follows:

Released	*36-Month Return to System Rate*	
Cohort I (Early 1967)	56%	Actual
Cohort II (Late 1967)	63%	Projected
Cohort III (Early 1968)	35%	Projected
Cohort IV (Late 1968)	33%	Projected
Cohort V (Early 1969)	29%	Projected

In addition to these seven studies, four others that were "operations research" in nature have already been discussed: the cost of correcting youthful offenders, a cost analysis of the DCDC work release program, the costs and benefits of narcotic addiction treatment, and trends in intake of narcotic-involved offenders. Finally, the operations research included a number of population characteristics, movement and projection studies.

These studies had as primary emphasis a focus on outcomes as measured by the criterion of return to the system. Secondary emphases were the relation of outcome to process such as program experience or manner of release, the effectiveness of programs in terms of cost criteria or cost and benefits criteria, and population types and trends.

Taken together, these studies illustrate common-sense approaches to operations research, using conceptual-arithmetical rather than computerized-algebraic methods. There were two exceptions to the general method: the development of the prediction instrument for work-release, and the cost-benefit analysis of narcotic addiction treatment. In each of these cases, the contractor or staff member made use of computers and mathematical models to process the data for the study.

Of these fifteen or twenty operations studies, the narcotic-offender intake analysis was the one that had the most direct and obvious impact on executive staff decision-making. The recommendation that community-based treatment for narcotic addicts be established was acted on promptly and with good success. The population-projection studies and the related recommendations for space management were of considerable importance for departmental decision-making. The cost-analysis of work release, which indicated cost-effectiveness for the program even though recidivism rates were higher than for direct parole, was instrumental in developing support for expansion of the work-release program.

Several of these studies had impact both within and outside the department, even though many of them did not receive distribution beyond the criminal justice and governmental agencies of the District. Such developments, out of the first few years of operations research activities of a new planning and research unit in a correctional agency, suggest that even elementary forms of operations research can play an important part in planning and operations in a modern department of corrections.

Two further points concerning this operations research experience are worthy of comment. First, the development of an unusually efficient prediction instrument in DCDC should have led to utilization of the instrument in the department. The failure of this to occur raises a number of questions: What were all the factors involved in the breakdown of what would initially have been presumed to be a clear pathway to utilization? Would a case study of this particular instance of nonutilization be of value in clarifying the general problem of nonutilization? Was the shift from an experimental to a more traditional administrative philosophy in the department in 1972–73 a major factor in the disregard of the prediction study results?

Second, the active role of research in DCDC during the five-year period 1968–72 and the major impacts achieved by research on the department and some interacting agencies is of interest in view of the fact that the research program was non-experimental in emphasis and operations-research oriented. The amount of impact was considerably greater than might have been predicted were one to start from the pessimistic position that some reviewers of correctional evaluation have taken in recent years. Furthermore, the studies that achieved the heaviest impacts were time-series studies, cost studies, and quasi-experiments, in diminishing order. The department's one controlled experimental study, now in its fourth year, has as yet had no marked impact on general departmental policy, although it may have aided the survival of the operating program to which the experimental component is attached. These observations suggest that non-experimental studies of the contemporary variety—operations research and systems analysis—like the older types of non-experimental studies, fit the needs of correctional decision-making better and promise higher levels of impact than the traditional "strong design" studies with elaborate experimental or statistical controls.

C. Systems Analysis

A shorthand definition of operations research is that it is a comprehensive set of procedures for optimizing the results of an ongoing system. Similarly, a shorthand definition of systems analysis is that it is a comprehensive set of procedures for optimizing systems. The distinction here is that systems analysis is more concerned with selecting among alternative system elements or alternative systems.

This distinction is critically important in contemporary corrections and criminal justice because of the high probability that most of the crime and delinquency problem is at bottom a problem of systems rather than of people.

If the problem is in the system, or in a hierarchy of systems starting with the broad family-neighborhood-school-and-society configuration and ranging down to corrections as the last system in the person-shaping continuum, it is largely futile to look for its solution in the person—the arrested, convicted and imprisoned offender. This is perhaps what the data on treatment in institutions—particularly institutions for adults—are telling us. Treatment in institutions for adults may be almost totally ineffective; it may even be counter-productive. The only plausible uses of institutions may be to incapacitate or punish, and these too may be counter-productive ideas that have too long gone untested.

If the trouble is largely in our hierarchy of systems, we will need to do more systems analysis and less treatment analysis. But what is systems analysis, and how is it applied to corrections and criminal justice?

We have already defined systems analysis as a process of selecting among structural or procedural alternatives to further achievement of objectives, taking effects, costs, and risks into account. Enthoven has added a few qualifiers to this general definition.[18] He sees it as a "reasoned approach" to decision-making, "accurately described as 'common sense.'" In addition to being used by decision-makers, it must be "fed with ideas by a broad interdisciplinary research program." It is "nothing more than quantitative or enlightened common sense aided by modern analytical methods." It does not require the use of computers, although if the data load becomes sizable, computers may be very useful. Furthermore, this approach is especially oriented to selecting a best course of action among several alternatives.

As a final comment, systems analysis has its elementary conceptual-arithmetic as well as its elaborate computerized-algebraic forms. As in the case of operations research, which can begin with elementary procedures and be augmented by occasional computerized analyses, systems analysis can begin with "systems thinking" and common-sense analysis, adding more sophisticated efforts and products as it goes.

1. *Present Uses in Corrections.* "Systems thinking" has been apparent at various times and places in corrections over the past two decades. It has shown itself in the introduction of cost-benefit analysis as a decision-making aid, in the development of information systems that support management decision-making as well as research, in actions and analyses that deal with subsystem interrelations (e.g., the effects of probation subsidies on county probation, state corrections, and public safety), and in the growing insistence that much of the ineffectiveness of criminal justice stems from its being a "non-system" rather than a functional system.[19]

More elaborate systems analysis has also been evident, mostly in efforts of exploratory kinds. These include a number of studies that have been carried out in corrections since 1965, primarily by private research organizations.

2. *Systems Thinking About Probation Subsidy.* One instance of sustained systems thinking (and related action) is evident in the contents of a report on the California probation subsidy. The report was recently submitted to the California legislature by the California Youth Authority.[20]

The probation subsidy program is described as having four goals:

- More even administration of justice in California,
- Reduction of commitments to State correctional institutions,
- Increased protection to the public,
- Rehabilitation of offenders.

The first two of these goals have been achieved during the eight years that the subsidy program has been in effect, the report states. With respect to a more equitable administration of justice, the report notes that prior to the start of the program, county commitments to state institutions varied widely, from 22 per 100,000 population in the lowest county to 119 in the highest. In the last full fiscal year (1972–73) the variation was from 8 to 66 per 100,000.

Reduction of commitments to state institutions was definitely achieved. The mean commitment rate of the 47 participating counties (out of 58 total) dropped from 68 to 34 in 1972–73. Since the beginning of the program in 1965, the commitment

total to state institutions had been reduced by about 25,000 convicted offenders. In the last year of this period, the reduction included 3,431 juveniles who would have gone to the Youth Authority and 2,018 adults who would have gone to the Department of Corrections.

The subsidy program does not appear to have had a marked effect toward increasing protection of citizens. The major reason is that the special community programs developed under the subsidy act involve only a small proportion of the offenders in the community. However, the report pointed out that there had been a decline in the number of juvenile arrests in California since 1969, dropping from 389,394 in that year to 353,232 in 1972. This was considered inconclusive evidence by the authors of the report.

The results regarding the fourth goal—rehabilitation of offenders—were also considered inconclusive. Preliminary data indicate that there has been little change between 1960 and 1972 in the success rate of probationers. The authors prefer to interpret this favorably; it indicates that there has been no decline in the overall rehabilitative effectiveness of statewide probation programs as a result of probation subsidies.[21]

The example shows us a "system," or rather a subsystem, consisting of adult and juvenile county probation, adult and juvenile state institutions, involving the county courts, making decisions about culpability and disposition, and placing individuals under supervision or in confinement. The subsystem feeds back rates of intake, of adjudication, of conviction, of disposition (to state or county programs), and these rates can be examined on a one-time basis to see whether the subsystem is achieving its objectives. The subsystem could also be monitored continuously, given a suitable management information set-up. The monitoring could provide periodic information on "how the system is doing."—how equitably, how safely, and (with cost data) how efficiently it is operating.

At the present time, the "analysis" of this system is partly common sense, partly simple statistical methods. If it were desired, the operation of the system could be described by a number of equations and its functioning could be simulated on a computer. The initial results would probably be of academic interest only. In several years, with suitably heavy investments in data collecting and processing and in successive stages of testing and analysis, practical decision-making and operating results might be derived from such a sophisticated system.

It is not clear whether "going sophisticated" is a useful thing to do in evaluating the present California probation-institutions configuration. If rapid change is to continue, the first efforts may be largely wasted. If change slackens but the corrections configuration is much less adequate than can be conceived in another few years, the development of a costly evaluation model may lock administrators into a design that will be hard to abandon and difficult to adapt. It may be advantageous, therefore, to rely for a time on common sense systems analyses before moving into the more elaborate forms on a state-wide basis.

3. *Elaborate Systems Thinking.* Although it may be premature to expect major impacts from sophisticated systems analyses in corrections, it will be useful to comment briefly on at least one example of such analysis. This is a study that was done several years ago to throw light on how California parolee recidivism rates are affected by length of prison stay and the releasing policies of the California Adult Authority.[22]

The study was done by a consulting firm, which based its analysis on data relating to about 41,000 adult male felons released to parole or discharged during the years 1960–68. For each offender, personal-social history, previous criminal history, the correctional programs to which he had been exposed, and his experiences with the criminal justice system in the years following release to parole were included.

The central matter of interest was the possibility of a relationship between time served in prison and the probability that the ex-prisoner would return to the criminal justice system for a future offense. Since this relationship was susceptible to definite influence by a number of factors, statistical controls were applied to selected factors in the analysis. These included such variables as intelligence, education, age, race, previous criminal record, and type of crime for which sentenced. Mathematical models or equations were then constructed to describe quantitatively the relationship between time served and probability of return.

The analysis used several statistical techniques, including discriminant analysis, response surface techniques, and multiple regression formulations. The analysis was facilitated by use of a large-scale computer to which the Department of Corrections had access.

The results of the analysis were presented to the Select Committee on Criminal Justice of the California State Assembly, where they might perhaps be useful in the drafting of legislation relating to corrections or to paroling procedures. The indications are that legislator viewpoints were influenced but no specific legislation has been drafted to incorporate the findings of the study.

4. *Significance of the Examples.* These two examples illustrate two approaches to systems analysis—one, the conceptual and arithmetic; another, the computerized and algebraic. Thus far, it is clear, the practical consequences from the simple approach have been far more important than those from the complex approach.

These two examples are not presented to demonstrate the comparative merits of the two levels of systems analysis. The principal point to be made here is that systems analysis does not have to wait for mathematicians and computers. There is possibly another point. The benefits of using the simple approach may be considerable, and the correctional administrator and the evaluator both need to be alert to possibilities for productive systems thinking and systems analysis at whatever level their staff qualifications and information sources permit. They should not hesitate to use a simple conceptual, arithmetical approach to understanding and evaluating systems if they lack data, mathematicians and computers for the more elaborate approach. In the short run, they are likely to progress faster by using the former. In the long run, however, there may be great advantage to working toward the more sophisticated forms of systems analysis. Keeping alert to potential instances of its use, as well as the procedures, conditions, and results of its use, may prepare an agency for successful implementation when the occasion presents itself.

D. Discussion

Operations research at both the common-sense level and the computer-and-algebra level will play increasingly important roles in correctional evaluation, planning and management. Administrators and evaluators need some conception of what these research approaches are, what their uses might be, and how implementation can be achieved.

For most SCAs, involvement in the more sophisticated forms of operations research and systems analysis does not appear desirable at the present time. For most such agencies, however, the use of the simpler forms of both techniques is a necessary element of good management. With continued application of the simpler forms, there may come increasing readiness for implementation of the more elaborate forms when the time is appropriate.

Larger systems, especially those with substantial backgrounds in the application of research, modern information systems, and effective working relationships with private consulting firms that have shown competence in correctional areas, should move ahead. They may find trial applications of the more elaborate forms of operations research and systems analysis an appropriate step at this stage in the development of evaluation.

REFERENCES AND NOTES

1. A. D. Hall, *A Methodology for Systems Engineering.* Princeton; Van Nostrand, 1962, p. 18.
2. Fred Hoffman, "Systems analysis for national security decisions," in J. Morley English (ed.), *Cost Effectiveness: The Economic Evaluation of Engineered Systems.* New York: Wiley, 1968, 214–41, p. 215.
3. Daniel Glaser, *Routinizing Evaluation.* Rockville, Md.: National Institute of Mental Health, 1973.
4. Vida Ryan, *California Prisoners: 1968.* Sacramento: California Department of Corrections, June 1969, p. v.
5. Space General Corporation, *Final Report: Prevention and Control of Crime and Delinquency.* El Monte, California, 1965. Prepared for the Youth and Adult Corrections Agency of the State of California.
6. Public Systems Inc., "Study of the California Adult Authority," in *A Proposal to Construct a Computer Simulation Model to Project Workload in the D. C. Department of Corrections.* San Jose, California, 10 June 1970, Appendix IV–4.
7. Steve E. Kolodney, Paul L. Patterson, Douglas Daetz and Robert L. Mark, *A Study of the Characteristics and Recidivism Experience of California Prisoners.* San Jose, Calif.: Public Systems Inc., 15 February 1970.
8. Steve E. Kolodney and Douglas Daetz, *Corrections Cost Projections: A Simulation of the California Criminal Justice System.* Mountain View, Calif.: Sociosystems Laboratory, Sylvania Electronic Systems, January 1969.
9. Steve E. Kolodney and Vida Ryan, "A Computer model for corrections population projections," in *Proceedings: 102nd Congress of Corrections.* College Park, Md.: American Correctional Association, 1972, pp. 34–41.
10. Stuart Adams, Wanda S. Heaton and John D. Spevacek, *Post-Release Performance of 432 Reformatory Releasees.* Washington, D. C.: D. C. Department of Corrections, Research Report No. 11, February 1969.
11. Wanda S. Heaton and Stuart Adams, *Community Performance of Three Categories of Institutional Releasees.* Washington, D. C.: DCDC, Research Report No. 15, June 1969.
12. Stuart Adams and Joseph B. Dellinger, *In-Program*

and Post-Release Performance of Work-Release Inmates: A Preliminary Assessment of the Work-Release Program. Washington, D. C.: DCDC, Research Report No. 13, March 1969.

13. Judith Hecht, *A Comparison of the Community Performance of Community Correctional Center and Institutional Releasees: Some Preliminary Findings.* Washington, D. C.: DCDC, Research Release, March 1971.
14. Stuart Adams, William Caldwell, Barbara Allen and Colleen Barros, *Twelve-Month Performance of Correctional Complex Releasees, by Category of Release.* Washington, D. C.: DCDC, Research Memorandum 7110, December 1971.
15. Stuart Adams, Wendell Plair and John D. Spevacek, *The Use of a Predictive Scoring System to Improve Selection for the DCDC Work Release Program.* Washington, D. C.: DCDC, Research Memorandum 71–7, November 1971.
16. Ibid, p. 10. Mean Cost Ratings of .41, .19 and .19 were calculated for the DCDC instrument and for two California instruments (multiple regression analysis and configural analysis), respectively, by Dean V. Babst of the New York State Narcotic Addiction Control Commission. For use of the MCR, see James A. Inciardi, Dean V. Babst and Mary Koval, "Computing mean cost ratings (MCR)," *Journal of Research in Crime and Delinquency,* 10 (January 1973), 22–28.
17. Stuart Adams, John D. Spevacek, Clarence Allen and Kathleen O'Brien, *Trends in Performance in the Community: Youth Center and Community Treatment Center Releasees.* Washington, D. C.: DCDC, Research Report No. 36, November 1970.
18. Alain C. Enthoven, "The systems analysis approach," in Harley H. Hinrichs and Graeme M. Taylor (eds.), *Program Budgeting and Benefit-Cost Analysis,* Pacific Palisades, Calif.: Goodyear, 1969, pp. 159–68.
19. Patrick Murphy, in *Evaluation of Law Enforcement Assistance Administration Programs: A Conference Summary.* Washington, D. C.: National Academy of Public Administration and the U. S. General Accounting Office, February 22–23, 1973, p. 24.
20. Summarized under the heading, "New Subsidy Report Released," in *California Correctional News,* 28 (January 1974), p. 12.
21. The delay of eight years in accomplishing a general evaluation of the California probation subsidy program illustrates at once the difficulty of planning and carrying out an assessment of a broad-scale criminal justice program and the importance of explorations into the use of systems analysis procedures for making such assessments.
22. Steve E. Kolodney, Paul L. Patterson, Douglas Daetz and Robert L. Mark, *op. cit.*

CHAPTER 14. SIMULATION

Broy recently pointed out that perhaps 10 percent of the nation's 500 largest corporations use some form of computer simulation to aid their executives in making decisions.[1] By 1980, he added, "Any concern that wants to remain in the *Fortune* 500 will have to use computer simulation just to survive."

Departments of correction may have no difficulty in merely surviving over the next ten years since they must continue functioning, willy-nilly. The pertinent question is the quality of the survival. Can SCAs move up to a reasonable level of performance and keep pace with demands by the public, by other elements of the criminal justice system, and by prisoners without extensive use of simulation in program and policy development?

A. Definitions

Simulation is a process of creating representations of, and acting out, the functioning of systems or their subsystems. Its purpose is to anticipate, evaluate, and improve control over the real systems that are being simulated.

War games are simulations of war; mock maneuvers test mobility, materiel, communications, and problem solving ability. Population projection in an SCA is a simulation of actual population trends in the future of the agency. *World Dynamics* portrays what will happen to the planet Earth's population, resources and environment over the next several decades given specified trends in population growth, resource consumption and environmental pollution.[2]

Simulation may be a commonsense process, as when an agency administrator constructs a departmental population "model" by connecting several points on a graph, then "operates" the model by extending the line for two or three years into the future to show where the population is likely to go. It may be a logical process, in which agency staff members with a body of facts and a number of assumptions about the department and about various environmental factors (laws, crime rates, budgets, etc.) arrive at conclusions as to what will occur under specific combinations of events. It may be a computerized process when data on the department, mathematical representations of departmental processes, and hypothetical statements of future conditions are "run" through a computer to learn what outcomes will result under alternative conditions.

B. Simulation in Corrections and Criminal Justice: 1964–74

Several simulations of correctional and criminal justice functioning have been made in the past ten years. None of them has had a profound impact on practice, but the potential for impact is there. The potential will increase as learning continues. Heads of SCAs and SPAs may anticipate various kinds of simulations in their fields in the near future.

The mid-1960s saw an "explosion" of simulation projects in corrections and criminal justice. Following are brief sketches of some of the projects that were initiated at that time and in the years intervening.

1. *Prevention and Control of Crime and Delinquency in California.* In 1964–65, Space General Corporation, under contract with the State of California, developed models of offender populations and typical offender careers and criminal justice career costs, examined alternative correctional structures and processes, and made recommendations for long-range approaches to dealing with delinquency and crime.

The broad scope, generality, and costs of the recommended actions limited the immediate usefulness of the project. However, as a basic attempt at systems analysis and simulation of a state correctional system, it has apparently had some effect on correctional and criminal justice thought.

2. *SIMBAD: Simulation as a Basis for Social Agents' Decisions.* In 1966, McEachern and Taylor of the University of Southern California began work on a model for a simulation system for use as a prognostic tool in probation decisions.[3] It was based on intensive studies of juvenile probationer experiences and outcomes in eight California counties. The model, stored in a large computer, provided

probation officers with two kinds of assistance in decisions on new juvenile cases: a) which of several alternative dispositions for the case will produce the "best odds" for success under specified sets of circumstances; and b) what is the most likely disposition of the case, given past experience and traditions in the department?

The model was successfully developed and put into limited operation by 1969. At that point funds for a follow-on phase failed to materialize, and an apparently effective simulation model was terminated.

3. *Corrections Cost Projections (1969).* Kolodney and Daetz developed a simulation model to investigate the comparative costs of operating the California correctional system under any of several alternative sentencing policies.[4] The model was designed to allocate sentenced offenders to jail, probation, prison, and parole according to the various sentencing possibilities. Unit costs in the several program alternatives were known. By projecting costs per offender by duration and type of program, total costs of operating each of the programs could be estimated. From these estimates, the system costs could be obtained.

Since the total system costs for a year would be affected by variations in sentencing policy, the effect of applying more "rational" sentencing policies was examined. One of the recommended policies was one based on sentencing practices in some of the more "progressive" California counties. If implemented state-wide, this policy would result in considerable savings, the simulation model demonstrated.

The report noted that these savings could be returned to the counties as a "performance" subsidy, based upon "performance as measured against sentence policy, and which could be used to upgrade and expand local programs within the system." This suggestion is essentially similar to the probation subsidy plan, adopted in 1965 to strengthen county probation and to reduce the numbers of persons committed to state correctional institutions.

4. *Corrections Population Projections (1972).* Kolodney and Ryan have been working since about 1970 on a computer model for projecting the institutional and parole populations of the California Department of Corrections.[5] Development and use of the model requires operations on three population movement types:

- Input to the system (from Superior Court)
- Movement within the system
- Output from the system (parole, discharge, etc.)

Future inputs can be calculated by applying estimated new commitment rates to the yearly California population data. Movement within the system is represented by a "best" weighted population movement matrix, arrived at by continuous analysis of past inter-program movements. A final aspect of the model is a set of "flow equations." With all these elements in order, the model is ready to be exercised.

In the projection of populations for the coming year, the present population serves as a base. Estimated new commitments from Superior Court are added, then operated on by the population movement matrix. The result is a projection for the coming year, which in turn will become the base for repeating the process next year. These steps can be repeated to obtain projections as far ahead as ten years.

When the model was used for the projection of the last year's population, the results showed that the model was "exceptionally accurate." In projecting the three major outputs—prisoners, active parolees, and inactive parolees—the maximum error was 4.1 percent. The authors conclude that the model brings together actual historical data, the expertise of corrections personnel, a simulation of the corrections system, sound mathematical techniques and knowledgeable analysts to produce "projections that an administration can use to support decision-making."

5. *Computer-Assisted Diagnosis and Evaluation (1973).* Vondracek and others, at the Pennsylvania State University, have developed in Project CARES (Computer-Assisted Regional Evaluation System) a computed-based model of decision making reminiscent of SIMBAD.[6] CARES was designed to aid juvenile courts and probation departments in Pennsylvania to evaluate their wards, comprehend the wards' problems, and arrive at the best decision on how to deal with the juveniles' difficulties.

The system, projected for completion in mid-1974, is ambitious in concept and versatile in its proposed outputs. It will provide the probation officer with a) a "natural language case summary" that is concise, integrated, and useful; b) a "syndrone analysis and problem summary," which uncovers developmental disorders and summarizes potential problems in wards; c) "treatment or disposition recommendations" suited to the needs of the juvenile and compatible with available resources; and d) "resources" information from the resources data bank that will aid in the making of realistic as well as constructive recommendations for disposition or treatment.

Since CARES is still in the "conceptual" stage, it is difficult to foresee its operating characteristics and the practical consequences of its application. Its installation costs will be considerable, and its ability to justify its operating costs in practical terms may be problematic. It may possibly escape the fate of SIMBAD, which failed to receive new funding in 1969 for reasons not readily apparent to proponents of the system. In any event, either as an exploratory concept or as an implemented system, CARES should contribute to at least a minor extent to the advancement of systems thinking and perhaps even to systems modeling in juvenile corrections.

6. *Simulation of the D. C. Court System (1966).* Navarro and Taylor reported in 1967 on the simulation of the processing of felony defendants through the District Court of the District of Columbia.[7] The project had the purpose of pretesting alternative approaches to court processing of cases with the ultimate aim of solving the chronic and aggravated problem of delay in case movement.

A simulation model was defined by the researchers as "a representation of the system and its operations which can be used to determine the effect of changes in the system." When used with the court system, it should be useful in analyzing alternative methods for accelerating the movement of defendants through the system—a much-needed change in view of the backlogs in the court.

Development and application of the simulation model proceeded through the following steps:

- *Description of the system:* This included the organization and structure of the court and its procedures for the handling of felony defendants;
- *Data Collection:* Data on felony defendants were collected and analyzed for distribution of defendant time in the system, principal events over time, delays, and causes of the delays;
- *Developing the simulation model:* The model was schematic or graphic rather than algebraic. It showed numbers (defendants and days), percentages and averages at points and intervals in the model;
- *Operating the model:* The model was "manipulated" or "exercised" to observe what effects specified organizational and procedural changes had on the rates at which defendants moved through the system and on the manner in which the available resources (staff, space, records) were used.

The simulation model addressed itself primarily to the problem of delays in defendant processing; it disregarded problems that might be associated with the content of the substantive law relating to the courts of defendants; it also did not concern itself with management procedures. A model that attempted to deal with delay as a possible function of these two factors would presumably have been much more complex; alternatively, three different models might have been required to handle these three aspects of court functioning simultaneously.

By operating the court model, staff arrived at recommendations for procedural and organizational changes in the court and at estimates of time savings that could be expected from the changes. They also made recommendations regarding the maintenance and processing of information pertaining to the defendants.

7. *Developing a Computer Simulation Model of the Juvenile Justice System (1974).* The National Center for Juvenile Justice recently announced plans for the development of a computer simulation model of a juvenile justice system.[8] The model is expected to provide a framework and procedures for assessing the impact of new juvenile justice programs while they are in the demonstration stage. It will also permit evaluation of programs already in the operating stage.

The model is still in the early conceptualizing phase, and its development and implementation depend on the availability of funds. If funding occurs, the model will go through three phases: conceptualization, field trial, and operations. It will be developed first for a metropolitan juvenile justice system and then, if circumstances warrant, be extended to a region.

C. Single and Multiple Models

Most of the preceding examples have described an approach to a limited aspect of decision-making within a system: decisions about dispositions of juvenile probationers, projecting institutional and parole populations, deciding what resources to use in dealing with the problems of a particular juvenile offender.

One of the models, pertaining to delinquency and crime in California, was all-encompassing. However, this model was highly general and proved to have no immediate practical impact on criminal justice in that state. Its effect on the structure of systems or on decision-making will come, if at all, by the roundabout route that is traced by some basic science products.

These observations raise the question: How many simulation models are required for the description, evaluation and improvement of a department of corrections or a criminal justice system? Will one model, possibly quite complex, suffice? Or must there be a number of models, each addressed to an important subsystem within the overall system?

Broy provides an illuminating description of how simulation modeling has worked out in one public utility.[9] The president of the company observed that all utilities face a number of problems such as 1) lengthy lead times for construction of new facilities, 2) rising costs for fuel and construction, 3) regulatory commissions, 4) environmental considerations, and 5) increasing frequency of fluctuation in money-market rates.

These complex factors plus an unprecedented demand for energy caused this particular utility to turn to computer simulation to chart its growth plans. It developed the following simulation planning models:

- *Rate and Revenue Model:* This allows the company to ask the "what-if" question concerning the timing and amount of rate increases;
- *Operating and Maintenance Cost Model:* This model projects fuel costs and operating expenses for the utility;
- *Capital Budgeting Model:* The function of this model is to monitor plans and execution of construction projects, examining alternatives at appropriate points;
- *Finance Model:* This model is concerned with the examination of alternative methods of financing utility operations and construction.

These four models operate in conjunction with a *consolidation model,* which takes information from the four and develops reports, documents, estimates and other generalized materials for income estimates, tax estimates, and so on.

This array of information, according to the president of the utility, enables the company to react more quickly, and presumably more effectively, to questions raised by changes in interest rates or gas costs, or to ideas involving construction or financing.

Blumstein and Larson indicate in somewhat less specific terms that modeling in criminal justice also proceeds through a number of sub-models rather than one general model.[10] They attempt to show how systems analysis is applied to the study of crime and criminal justice. The approach is through the modeling of the total administration-of-justice system, which is made up of the police, the courts, and corrections. In dealing with the total system, a generalized description was presented, and then one sector—the area of "police apprehension"—was brought into focus for the purpose of developing a "patrol detection model."

Blumstein and Larson point out that the systems approach to the analysis of complex social systems avoids the description of system operations in terms of explicitly individual human behavior. Rather, it uses "macroscopic models and empirically derived data to express the behavior implicitly." This new approach to the study of agency functioning offers a relatively simple means of testing the consequences of alternative actions in the operation of social systems.

D. Discussion

Simulation models are centuries old in some forms but only about twenty years old in computer-based forms. They have begun to impact significantly upon modern industry, and there is ample evidence that they have begun to find application in the field of corrections and criminal justice. Their use in the latter area is now about ten years old.

Present indications are that simulation may become one of the most important evaluation models in the correctional and criminal justice systems. It is the model that appears most likely to cope effectively with the many problems of assessment in complex and rapidly changing systems. Because of its focus on operating problems, it is compatible with the needs and interests of administrative decision makers. Its global perspective will be extremely useful in tying together for analytical purposes the components of criminal justice, which thus far have managed to remain a "non-system." Its capacity for differentiating efficient from inefficient alternatives gives it an edge in promoting innovation and making constructive breaks with past practice.

Although simulation has its commonsense aspects, its main effects will come from computer-based analyses that rest upon adequate bodies of data pertaining to the operations that are being simulated. This implies that effective use of simulation will require good information systems, expertise in mathematics as well as operational knowledge, and access to computers. Since these trends are now well-established in some correctional agencies, it may be anticipated that simulation will play an increasingly significant role in the evaluation and development of correctional programs and systems. Parallel developments may be anticipated in the courts and police systems and, ultimately, in all three of these entities taken together as the criminal justice system.

REFERENCES AND NOTES

1. Anthony Broy, "Computers aid in decisions," *New York Times,* 14 November 1971.
2. Jay W. Forrester, *World Dynamics.* Cambridge: Wright-Allen, 1971.
3. A. W. McEachern, Edward M. Taylor, J. Robert Newman and Ann E. Ashford, The juvenile probation system: simulation for research and decision-making," *American Behavioral Scientist,* 11 (January-February 1968), 1–45.
   Also, Public Systems Research Institute, *SIMBAD: A System for Computer-Aided Probation Decision Making.* Los Angeles: University of Southern California, June 1969.
4. Steve E. Kolodney and Douglas Daetz, *Corrections Cost Projections: A Simulation of the California Criminal Justice System.* Mountain View, Calif.: Sociosystems Laboratory, Sylvania Electronic Systems, January 1969.
5. Steve E. Kolodney and Vida Ryan, "A computer model for corrections population projections," in *Proceedings: 102nd Congress of Corrections.* College Park, Md.: American Correctional Association, 1972, pp. 34–41.
6. Fred W. Vondracek, Hugh B. Urban and William H. Parsonage, "The Computer-Assisted Regional Evaluation System (CARES)." Presented at the annual meeting of the American Society of Criminology, New York, November 2–6, 1973.
7. Joseph A. Navarro and Jean G. Taylor, "Data analysis and simulation of court system in the District of Columbia for the processing of felony defendants," in *Task Force Report: Science and Technology.* Washington, D. C.: President's Commission on Law Enforcement and Administration of Justice, 1967, pp. 199–215.
8. Hunter Hurst, "A computer simulation model of a juvenile Justice system." Reported under "Announcements," *Criminal Justice Newsletter,* 5 (11 March 1974), p. 8.
9. Anthony Broy, *op. cit.*
10. Alfred Blumstein and Richard C. Larson, "A systems approach to the study of crime and criminal justice," in Philip M. Morse (ed.), *Operations Research for Public Systems.* Cambridge: Massachusetts Institute of Technology Press, 1967, pp. 159–80.

PART FOUR.
LOOKING AHEAD

CHAPTER 15. EVALUATION OF EVALUATIONS

If evaluation of correctional programs is to be reasonably effective, continuing assessment of the evaluative effort is necessary. Evaluation is a changing, learning process, sometimes characterized by great inefficiency and errors in direction. To guard against and to correct these tendencies, frequent review and redirection are essential.

Self-assessment is especially desirable at the present time because correctional evaluation is in a period of unusual difficulty, brought on by pressures for hurried action. The drive to apply unprecedented amounts of funds to action programs, with evaluation by inexperienced persons, has produced vast bodies of what Berkowitz terms "junk" research. The current pressures for evaluation coincide with a new era in the history of applied science, when new methods are beginning to appear. These, if effectively applied, promise major improvements in the evaluative process.

There are, then, two reasons for heightened interest in self-assessment. First, there is the possibility of reducing the gross waste in correctional evaluation. Second, there is the hope of discovering methods, strategies, and priorities of research that are especially appropriate in contemporary corrections and criminal justice.

But how can evaluation of evaluations be organized, guided, and be brought to useful conclusions? How does one determine what methods, subject-matter areas, research personnel and research structures are desirable?

One method is by examination of the results that are produced. Another is by study of the internal characteristics of the research procedures. Practical wisdom suggests that evaluation of research in terms of its actual and intended outcomes is the generally most satisfactory method. However, it is also important to examine the other alternative—the design and quality of the research effort. With appropriate designs and good quality, the probability of useful results should be enhanced.

A. Assessment by Results

There is a strong tradition in research to judge the quality of the effort by the canons of scientific method, by compliance with rules, forms and principles. Some attention is paid to results, but since the results of research are often far into the future, the form of effort rather than the results of effort has been the most convenient framework for assessing research.

In the practical world, this might be called the "managerial" approach to assessment. However, the practical world has found it more satisfactory to judge efforts by their results rather than by their adherence to specifications or to principles. In evaluative research, it is possible to use the former criterion more consistently because such research is oriented toward results in the immediate present. Consequently, the assessment of evaluative research by its results should be considered an important step in the total evaluative effort.

According to this view, the design of the research is less significant than its consequences. If "weak" designs bring more relevant and more important consequences than "strong" designs, weak is to be preferred to strong, even though we traditionally accord more prestige to strong designs. If systems analysis deals more powerfully with the wide assortment of evaluative problems in criminal justice than the controlled experiment, systems analysis deserves a hearing even though many scientists are not sure it is respectable.

In assessing research by its consequences, we can pay attention to magnitude of impact, immediacy of results, and implied or remote consequences.

1. *Magnitude of Impact.* The impact of research has both measurable and intangible aspects. Some are immeasurable because they are delayed indefinitely into the future; some are vague and diffused, hard to quantify. Consequently, the measurement of impact is a mixed process, partly judgmental, partly quantifiable, and often largely ambiguous. Neverthe-

less, there is a real need to measure impact, and some beginnings have been made.

When we are told that 25,000 adjudicated offenders have been retained in their communities in California rather than going into state institutions, this is evidence of the impact of the studies that led to the probation subsidy. When we are told that several hundred California youth remained in their communities because of the Community Treatment Project, it may be concluded that the probation subsidy studies had an immediate impact of greater magnitude than the CTP studies. These relative figures do not measure all the impacts of these two bodies of evaluative research, but they illustrate the possibility of talking quantitatively about the impacts of research.

2. *Immediacy of Impact.* Some researches yield practically useful information in a matter of weeks or months while others take years. The D. C. Department of Corrections study of narcotic-involved offenders led within six months to the establishment of community-based centers for the treatment of addicted offenders and within twelve months to the establishment of a District-wide Narcotics Treatment Administration.[1] The Youth Studies Center experimental analysis of the effectiveness of Group Guidance began in 1961 and the final report was published in 1966. By then Group Guidance had been transferred out of the Probation Department and given other roles in Los Angeles community processes.[2]

Gang activity is still a problem of high priority in the criminal justice field, although less so perhaps than narcotic addiction. The Group Guidance experiment has not yet had an effect on gang management practice. Is it because the time for decision on gang management in Los Angeles was 1964, not 1966? Is it because of the design of the YSC experiment, which ". . . inadvertently led to greater gang delinquency rather than a reduction in gang delinquency"?[3] Is it because of the credibility level of the study, which had serious problems with both internal and external validity?

If these are all relevant questions, immediacy of impact relates to many factors and its control may require attention to a wide variety of matters. However, whether or not there is an impact, and whether it is early or late, is usually easy to determine, and hence it is a means of assessing the results of evaluation.

3. *Implied Consequences.* Some important consequences of evaluative research tend to be delayed, and the assessment of evaluations needs to account somehow for this fact. This is perhaps the most judgmental of the aspects of evaluation that we are considering here, the hardest to document, and presumably the hardest to control.

One aspect of the delay is the time required to effect utilization of an obviously important finding or concept. The eight-year delay between the first recommendation on probation subsidy in California and the enactment of the law is an example. Another aspect is the need to translate what is basic knowledge rather than programmable material into a usable form. The importance of an offender typology began to be sensed in the SIPU study of parole caseload size in the early fifties, and it was demonstrated to be important in PICO in the late fifties. However, it was not until the more complex Community Treatment Project was undertaken in the early sixties that an offender typology became clearly useful as an operational device.

Some aspects of evaluation are knowledge-oriented rather than action-oriented. This means that there may be an application at some time in the future. How much weight to give this aspect of a proposed evaluation is extremely problematic. It is difficult to anticipate basic knowledge, but it is also difficult to recognize it once it has emerged. Perhaps all that we can say is that it should not be ignored. As Utterback pointed out, the major factor in utilization is the "market place," but scientific discovery also has a role in ultimate utilization.[4]

B. Assessment by Proposed Objectives

Assessment by operational results applies a "hard criterion" to evaluative research. A softer criterion is used when research is assessed in terms of achievement of its own objectives.

Evaluative studies specify their objectives as a first step. They propose to examine whether a program has reduced crime significantly (or perhaps by 10 percent over last year), test a practically useful hypothesis, determine whether a test score has increased significantly, and so on. Many or all these objectives may be reached, in which case the studies are successful in their own terms.

Success in meeting own objectives is sometimes but not always the same as being operationally useful. Research objectives generally derive from agency objectives, but in actuality successful research is not followed automatically by agency utilization. The independence of these two elements means that research lacking in operational impact may still be found to be successful in terms of fulfilling its own objectives. This may be of little interest to an agency administrator, even though there has been some addition to knowledge.

Assessment in terms of own objectives has been used by some SPAs that have tried to keep pace with the flood of evaluative reports coming in from their grant recipients. One example of such assessment was a review of evaluative reports by the California Council on Criminal Justice.[5] The review examined a sample of 38 out of approximately 400 project reports to ascertain 1) whether or not stated objectives were attained, and 2) kinds of methodological deficiencies that were evident in the design of the research. The review recognized three types of design in the studies: pre-experimental, quasi-experimental, and experimental.

In view of the intensive effort made by CCCJ to grapple with the matter of assessing its evaluative studies, it will be instructive to present materials abstracted from that review. These include a list of projects by program area, title and project number; a tabulation of objectives attained and failed along with methodological deficiencies noted; and a summary of the major types of objectives attained or failed and the types of methodological deficiencies noted.

Chart 1 lists the projects, and Chart 2 presents a count of objectives attained or failed and deficiencies noted. Chart 3 identifies types of objectives attained or failed and types of deficiencies noted. The data in Chart 3 are condensed from a larger chart that cross-tabulated project identities and types of objectives, outcomes, and types of deficiencies.

In Chart 2, projects 0601 and 0900 are multiple projects or clusters, and they are treated as eight projects or six projects, respectively, when averages or ratios are calculated.

This manner of assessment of evaluations can serve several purposes. It makes possible comparisons between the general strategies, research designs and research outcomes achieved by different SPAs. It provides one example of a procedure for assessment of evaluative research in terms of its stated objectives. And, finally, it is a preliminary exploration of the "second level" of efficiency of evaluative research; that is, the rate at which specified goals are realized and methodological weaknesses are avoided.

The next step in the CCCJ assessment would presumably be a move up to the first level, where evaluative projects are judged by their operational impacts. One useful inquiry at that level would be to ascertain what types and proportions of projects that achieve their objectives also have definite effects upon operations. Another useful inquiry would be the relationship between magnitude of effect and characteristics of the project and its results. Still another area of interest would be the characteristics of the agency and community settings in which an effect was achieved.

C. Design and Quality of Research Procedures

A third level of assessment of evaluative research—twice removed from the level of operational impact—is that of the design and quality of the evaluation model. At this level, focus is heavier upon rational than upon empirical considerations. It rests upon the belief, supported to some extent by experience, that certain kinds of designs are more likely in the long run to bring results. This is a defensible position, especially when the results specified are dependent upon precise comparison or causal inference. To the extent that the latter are important specifications, they will cause higher values to be placed on experimental and elaborate statistical designs.

As was true of assessment by own objectives, assessment by type of design yields rankings that correlate only moderately if at all with the results of assessment by operational impact. In many respects, "type of design" is the softest of the three criteria mentioned thus far; there are several barriers between design and achievement of project objectives, and still more before operational impact is assured.

D. Credibility of Evaluative Research

Credibility of research is an intangible quality that derives not from single factors but from a configuration. It is rarely used in assessment except in an informal sense, although it may eventually prove useful and hence worthy of application. It offers a means of differentiating between projects or evaluations that are similar in many characteristics but different in level of "believability."

Chart 1. PROJECTS FOR PREVENTION AND CONTROL OF JUVENILE DELINQUENCY

**	0283	Juvenile Diversion Through Probation Counseling
**	0289	Juvenile Diversion Through Drug Abuse Counseling
**	0265	Delinquency Control Through Parental Education
**	0563	A Juvenile Day Care Center: Alternative to Incarceration
*	0259	A Residential Treatment Center for Youth
*	0293	A Multi-Service Counseling Center for Delinquency Prevention
	0591	A Community-Based Delinquency Prevention Project
	0601	Eight Youth Service Bureaus for Prevention & Control of Delinquency
		TRAINING FOR CRIMINAL JUSTICE PERSONNEL
	0448	A Regional Criminal Justice Training Center
		PROJECTS TO REDUCE ORGANIZED CRIME
	0564	An Organized Crime Intelligence Unit
		PROJECTS TO IMPROVE POLICE SERVICES AND OPERATIONS
	0608	Restructuring a Police Department
		PROJECTS TO REDUCE A SPECIFIC CRIME
	0900	Six Crime-Specific Burglary Projects
		SERVICES FOR EX-OFFENDERS
*	0481	A Halfway House for Ex-Offenders
	0677	Coordinated Services for Ex-Offenders
		ALCOHOLIC TREATMENT PROJECTS
	0221	A Halfway House for Treatment of Alcoholics
*	0505	Counseling for Probationers with Drinking Problems
		PROJECTS TO REDUCE THE DRUG PROBLEM
	0848	Reducing Drug Traffic at the Mexican Border
	0749	A County-Wide Narcotic Education and Enforcement Project
	1140	A Countywide Narcotic Enforcement Bureau
*	0541	A Comprehensive Drug Treatment Project
	0444	A Family-Structured Treatment Project for Drug Abusers
		POST-RELEASE SERVICES FOR YOUTH
	0474	A Project to Improve Youth Parole Services
		PROJECTS TO IMPROVE COMMUNITY RELATIONS
	0443	A Comprehensive Police Community Relations Project
**	0405	A Values Instruction School Resource Officer Project
	0158	A School Resource Officer Project in Davis
	0419	A School Resource Officer Project in Ventura

** Controlled experiment
* Quasi-Experiment
Non-experimental project

Chart 2. 38 CRIMINAL JUSTICE PROJECTS GOAL ATTAINMENT AND METHODOLOGICAL DEFICIENCIES

	Project Number	Goals: Attained (+),	Failed (0)	Methodological Deficiencies (−)
**	0283	++++		
**	0289	++++	00	−
**	0265		00	−
**	0563	++	000000	
*	0259	++	000	−
*	0293	++	00	−
	0591	+++		−
	0601 (8 projects)	+	00000	−−−−
	0448	++++	0	−−
	0564	++	00	−
	0608	+	0	−−
	0900 (6 projects)	+	0	−−
*	0481	+++	00	−
	0677	+	0000	
	0221	+++++		−
*	0505	+	0	−
	0848	++	0	−−
	0749	++	00	−−
	1140		0000	−−−
*	0541	++		
	0444	+++		−
	0474		00	−−−
	0443		00000	−
**	0405	+++++		
	0158		000	−−
	0419		000	−−

			Goals Attained per Project	Attainments over Failures	Deficiencies per Project
**	Experimental Projects	(N = 5)	3.0	1.50	.4
*	Quasi-Experimental	(N = 5)	2.0	1.25	.8
	Non-Experimental	(N = 28)	1.3	.55	2.6

Chart 3. MOST COMMONLY ATTAINED OBJECTIVES AND MOST COMMON METHODOLOGICAL DEFICIENCIES

Most Commonly Attained Objectives

1. Decreased recidivism
2. Increased diversion
3. Improved cost-effectiveness

Most Common Methodological Deficiencies

1. Lack of control group
2. Lack of standards for measuring impact
3. Insufficient data

The concept may serve best with experimental or quasi-experimental studies. One of the marks of credible research is, in Campbell's terms, high internal validity; another is high external validity.[6]

Other aspects of credibility relate to features of design that anticipate problems of data interpretation. To deal with population heterogeneity, there is a typology; to measure in-community performance, several criteria are used; to detect cost or benefit differences, a monetary criterion is used; to identify decision shifts, appropriate criteria are used.

Credibility is a qualitative criterion for assessing evaluations that may be difficult to operationalize in the ranking or rating of researches. It should be useful, however, for sensitizing evaluators to aspects of their work that have been only lightly attended to in the past.

E. Discussion

Although the evaluation of evaluations may seem an abstract enterprise, it has quite practical implications. It carries the potential of reducing waste of research resources by differentiating ill-conceived and poorly directed or badly coordinated studies from more likely kinds of evaluations. In addition, it offers opportunities to work for refinement of methods, better identification of proper evaluation targets, more effective links between evaluation and utilization, and higher probability of impact.

As between reducing the trial-and-error quotient of evaluative research and maximizing the utility of research, the primary goal of research assessment is probably the latter. This means trying for success in each stage of a two-stage process: 1) attaining project objectives, and 2) securing utilization or application of results when projects prove effective in demonstrating results.

The problems of stage one are different from those of stage two, and success at the first stage does not guarantee success at the second. The relationship here is perhaps like that between managerial efficiency and outcome effectiveness—there may be low correlation between the two.

Improvement in the utility of research might proceed more successfully if one began with a number of projects that have been accepted as useful and made the basis for operational or policy change, then worked backwards, as it were. The focus of interest in the backward perspective would be the characteristics and conditions of the useful research. What are the areas of application, the methods of evaluation, the kinds of persons or organizations that did the evaluation, the relation of the researchers to planning and decision-making, the project auspices, and the antecedents of the project? Examination of these factors and circumstances might indicate with some degree of confidence how evaluative research with higher probabilities of payoff should be planned and executed.

REFERENCES AND NOTES

1. Stuart Adams, Dewey F. Meadows and Charles Reynolds, *Narcotic-Involved Inmates in the Department of Corrections.* Washington, D. C.: D. C. Department of Corrections, Research Report No. 12, February 1969.
2. Malcolm W. Klein, *Street Gangs and Street Workers.* Englewood Cliffs, N.J.: Prentice-Hall, 1971.
3. Ibid, p. 137.
4. James H. Utterback, "Innovation in industry and the diffusion of technology," Science, 183 (15 February 1974), 620–26.
5. Francine Berkowitz, *Evaluation of Crime Control Programs in California: A Review.* Sacramento: California Council on Criminal Justice, April 1973.
6. Donald T. Campbell, "Reforms as experiments," in James A. Caporaso and Leslie L. Roos, Jr. (eds.), *Quasi-Experimental Approaches.* Evanston: Northwestern University Press, 1973, pp. 187–225.

CHAPTER 16. IMPROVEMENT OF EVALUATION

Evaluative research may be improved by assessing its outcomes and using these results to refine methods and procedures. It may also be improved by processes that are primarily rational—developing basic premises and drawing the more likely conclusions about how research should be organized and conducted. There is still another process that we can employ: taking other people's conclusions about how to improve research and working them into a coherent plan. The latter method is the substance of the present chapter.

We begin by presenting some conclusions about how to improve correctional evaluation, or social program evaluation, and then move on to merge some of these conclusions with the major ideas that have developed out of this Guide. The process is an informal one, not concerned with axioms, postulates and theorems, and formal derivation procedures. It draws from the work of several academicians or scholars who have been writing about evaluation, either in corrections or in social programs generally. The works were chosen for their visibility, not by any systematic selection procedure.

A. Ideas for Improving Evaluation

The several ideas for improving evaluation that follow are grouped under the names of the persons who proposed them. Some of the groups of statements may not include all the ideas the observer has ever stated about the subject; they focus primarily on presentations made in particular papers or documents. Six observers are cited.

1. *Bailey (1966)*. The following ideas are drawn from Bailey's paper on 100 correctional evaluations.[1] In his discussion, Bailey, a former correctional practitioner and now a professor of social work, concludes with some suggestions for the improvement of evaluation:

- There should be increased concern with outcome studies,
- The calibre of studies should be improved,
- There should be more research by professionally trained personnel,
- Research design should show increased rigor and sophistication,
- Use of behavioral science theory should be expanded.

2. *Orlans (1969)*. Orlans, a senior fellow at Brookings Institution, focuses on a number of practical concerns: [2]

- Abler people should be involved in practical research,
- More fragments of knowledge should be pieced together,
- Research objectives should be stated more clearly,
- There should be more independent evaluations of government programs,
- Research should be focused on the more significant problems,
- The utilization of research should be promoted.

3. *Ward (1973)*. Like Orlans, Ward, a professor of sociology, takes a practical view of the needs of evaluation:

- Many agency services should be evaluated community-wide, with cost sharing and agreed-upon objectives,
- Research should be centralized to secure consistency,
- Evaluation should be restricted to a limited number of programs,
- Research funds and staff should be concentrated on established organizations in prison, probation and parole,
- Duplication of efforts between regions or states should be reduced,
- Coordination at state or national levels should be effected,
- Continued replication of studies by for-profit organizations should be discouraged,
- University faculty should be used in evaluations of correctional agencies, thus reducing conflict of interest and encouraging publication of results.
- Evaluation should be recognized as a means of "keeping corrections honest."

4. *Glaser (1973)*. Glaser, correctional evaluator, criminologist and professor of sociology, in a recent paper placed strong emphasis on theory as a means of improving correctional evaluation:[4]

- Theory is the key to good study design and the formulation of proper research questions,
- Theory should be elaborate enough to account for the wide variety in types of offenses,
- Theory should differentiate research subjects not only by type of offense but also by careers, life-styles and personality types,
- Research should be designed to evaluate practice on the basis of theory rather than on policy,
- Theory should have several levels of abstraction, interlinked, from very general theory to specific propositions.

5. *Suchman (1967)*. Suchman, an academician, has written at length about the principles and procedures of social action evaluation. His recommendations are primarily methodological:[5]

- A new "social experimentation" model is needed for social action research,
- Evaluative experiments should contribute to both knowledge and operations,
- The evaluative research model should be in the format of "intervening variable analysis,"
- The longitudinal panel design offers the greatest promise for evaluative research,
- Good evaluation continues over time and permits "before" and "after" measurements.

6. *Weiss (1971)*. Weiss, an academic researcher of social action programs, focuses on the practical problems of research utilization, with some attention to methodological concerns:

- There is need for evaluation staff with dual skills—to discover significant data and to influence decision-making,
- Clients and local groups should be encouraged to accept useful findings and to reject programs with poor outcomes,
- Useful results should be presented to persons (program managers) who need and can use them,
- Reports of research should be timed sensibly, not after a decision dead-line,
- Communications to users should be in non-technical language,
- There should be a program planning and development unit to translate research results into action programs,
- Evaluators should feel free to speculate about alternatives to present programs,
- Evaluation should test ideas for alternative programs in "small-scale experimental projects." [6]

These *six sets of proposals* for the improvement of evaluative research fall into several general areas: theory, method, organization, and utilization, primarily. It is of interest that some observers stress theory, others promote particular methodological models, and some focus primarily on organization. These selective approaches would probably broaden considerably if each of the observers were to be asked to present a comprehensive plan for the improvement of evaluation.

Although this haphazard sampling touches upon a number of aspects of correctional evaluation, it leaves a number of gaps. We proceed now to build on the foregoing proposals, filling in at various points and working toward a general plan for improvement of evaluation in corrections. A comprehensive plan should contain at least the following elements:

- Objectives of evaluative research
- Organization of evaluation
- Coordination of evaluation
- Staffing and funding of evaluation
- Theories of correctional improvement
- Planning of evaluation
- Methods and strategies
- Reporting and dissemination of evaluation
- Utilization of evaluation
- Assessment of the evaluative process
- Development and improvement of evaluation

A systematic effort to improve correctional evaluation on a long-range basis must give consideration to each of the foregoing topics. The importance of each topic will vary with the state of the art and with the situation.

B. A General Plan for the Improvement of Evaluation

The following discussion is necessarily very sketchy. A full treatment of this topic could occupy an entire volume, which is out of the question here. The present approach is that of an outline with commentary.

1. *Objectives of Evaluative Research in Corrections.* Objectives are key elements in productive evaluation. Since research objectives derive from organizational objectives, the goals, and ultimately the productivity of research, can be no more successful than the objectives of corrections are realistic and attainable. Consequently, the improvement of evaluative research in corrections is tied to the eventual development of desirable and achieveable objectives in corrections itself.

This means that evaluators must concern themselves with the nature of correctional objectives. Are they rational or irrational, constructive or destructive, costly or reasonable, achievable or unattainable? Goals that are generally negative have nothing to contribute to productivity in research or in operations. Although it is not the evaluator's role to set the goals for corrections, he may be in a position to provide tested knowledge on alternative objectives, and he may thus improve the productivity of evaluative research by promoting better rather than worse objectives.

2. *The Organization of Evaluation.* Without organization, research tends toward chaos and waste. Current problems in evaluation suggest a need for more definite structuring of direction and control. There needs to be some centralization at the national level, with regional or local activity responding to guidance from the center. This permits much-needed coordination, facilitates pin-pointed application of evaluations at several appropriate sites, and reduces duplication that is wasteful without being informative. Such organization allows for the development of evaluation in accord with a broad and sustained evaluation strategy.

The relationship of correctional agency research units to the central entity might be that of voluntary participant in a larger endeavor. Not all state units need participate in most studies. Those that do not will presumably benefit from findings through report dissemination, workshops and institutes, and technical assistance.

The structuring of correctional, court and police evaluative units into the larger system may occur most reasonably through the state planning agencies. Presumably, in every state, the SPA will become a coordinator and possibly a controller of research within the criminal justice system components. However, not every SPA need become an active participant in major evaluation projects that are of national scope.

Part of the organization of research relates to use of cluster analysis or simultaneous replication of projects to provide comparative data and to permit specified types of variation that require several sites for implementation. This aspect of structure belongs to strategy as well as to organization.

3. *Coordination of Evaluation.* A major concern of corrections and criminal justice is to eliminate needless duplication and reduce waste of research resources. Duplication that is not productive—that is, not deliberately replicative nor intentionally accumulative—needs to be reduced through concerted planning and cooperative monitoring. The remaining research can then be better interpreted, disseminated and utilized.

Within states, some coordination will be the responsibility of SCA administrators or SPA heads, as when related projects are carried out simultaneously at several facilities or sites. Across states, coordination may be handled by consortiums or by a national agency or both.

4. *Staffing and Funding of Evaluation.* Large agencies need their own in-house evaluation staffs, if only to carry out the routine operations research that is becoming essential to any sizeable organization. Some large agencies should be encouraged to build and maintain multi-purpose research staffs, with the intent of engaging in other types of evaluation. This could include assessment of new programs, systems analysis and simulation, inter-agency cluster studies, and rigorously designed experimental studies that have both practical operations and knowledge as goals.

In-house evaluation staffs appear especially desirable for at least two reasons: first, they seem to be performing the best agency evaluations that have been accomplished in recent years; and, second, they appear unusually suited to increase the utilization rate of productive research. A third possible reason for favoring in-house staffing of research is that such units are effective centers for technology transfer.

There is some value to outside evaluation, as Orlans and Ward point out; it provides independent measurement. However, there seems to be reason

to doubt that the university researcher is better qualified to perform agency research and is in a position to bring "honesty" to correctional evaluation.

An important argument for increasing support to in-house evaluation is the desirability of encouraging competent researchers to make careers of agency evaluation. "Moonlighting" researchers from college campuses do not provide the continuity, the identification, the communications ability, nor the accumulated experience that is beginning to show up in the superior performance of some in-house units. To attract competent researchers, criminal justice agencies will need to provide salary levels, role stability and organizational status sufficient to interest and hold qualified persons.

The need for evaluation by outsiders should diminish in the future as the present special funding for innovative projects terminates. It should also diminish as agency evaluation moves to higher levels of technical and professional demands, and the ability of the casually interested outsider to meet these demands decreases.

Some outside sources of evaluative research—the private organization with expertise in systems analysis and simulation, for example—will continue as useful supplements to in-house units. University staff who have begun to develop skills in these areas may also become participants in special projects.

5. *Theories of Correctional Improvement.* The improvement of correctional evaluation depends to some unknown extent upon advances in correctional theory. However, the nature of the theory that is likely to be most productive is open to debate. If the key problem is in personality structure and personal behavior, then behavioral theory will be relevant although of indeterminate importance. If the correctional problem is one of "systems," behavioral theory may be a dead end, and corrections will make little progress until systems theory is able to provide new directions.

Since behavioral and social theory have had a relatively free hand in corrections for the past two decades, with little apparent effect, it may be time for extensive borrowing from systems theory. This may provide the most direct resolution of the issue of the role of theory in correctional progress and in productive correctional evaluation.

6. *Planning of Evaluation.* The planning of correctional and criminal justice evaluation will become more specialized and more difficult as knowledge accumulates and research tasks become more complex. Effective planning will be increasingly critical to the improvement of evaluation.

The planning of evaluations will involve a command of both social science and systems analysis methodologies, with increasing emphasis on the latter. The planner will have to deal with several areas of concern: reduction of trial-and-error or ad hoc research; use of methods that promise results; focus on areas where progress is possible; use of replication, particularly cluster analysis, or simultaneous replication; coordination of research and development; and the use of both short- and long-range strategies.

7. *Methods and Strategies of Evaluation.* The rise of systems methods relative to the traditional social science methods in agency research poses new problems for criminal justice administrators and evaluators. There are new skills to be procured, new learnings to be achieved, decisions on correct applications of methods and models to be made, and other elements of research strategy to be tried and evaluated by the results they bring.

If this area is not the key to improvement in evaluative research it is at least one of the central elements. The evaluator needs to be fully aware of the range of alternatives, the appropriate uses, the risks and potential rewards of the various methods and strategies that are open to him.

8. *Reporting and Dissemination of Evaluation.* The need for accelerated studies in a few locations and transmission of worthwhile findings to hundreds or even thousands of potential users places a premium on effective reporting and dissemination. Technical reporting alone is generally inappropriate. Publication in scientific journals also has its problems, even though it is favored by academicians. Technical reporting is a hindrance to utilization; scientific reporting delays use, diverts attention and funds, leads to neglect of findings of low certainty but high interest to decision-makers, and keeps too many "publish-or-perish" individuals in the agency research field.

Since technical reporting is sometimes inescapable, it should be done not alone but in a dual-reporting scheme where a technical report is joined with an administrative report or summary that provides the gist of practically important findings in non-technical terms.

Publication of operationally significant materials can usually be effected quickly and easily in professional (as opposed to "scientific") journals; those items that are of interest to academicians tend later

to be reprinted from the professional journals in books of readings.

An important aspect of dissemination is the standardization of reporting procedures. Evaluations of particular types of programs in one region can then be compared with similar evaluations from other regions with less risk of misinterpretation. This should avert some instances of extended correspondence or communication with several sources to unravel dissimilarities of technique, language, or results.

Evaluation could be improved if special arrangements were made to assemble planning, progress, and final reports of evaluations in separate sites on the same types of programs or activities and to engage in continuing interpretations of the materials. Interaction between interpreter and sources should be mutually useful; it should also be helpful in making the materials more applicable by future users.

9. *The Utilization of Evaluation.* Utilization of research results is facilitated if researchers are also capable of planning and development. It is facilitated when research is carried out by in-house staff. It is facilitated in agencies where the administrator is "experimental" in attitude and action. It is facilitated when the research product relates to significant policy matters and program needs. It is facilitated when researchers communicate with program managers as well as with agency administrators, use nontechnical language, and present their findings in advance of dead-lines for final decision-making.

If research products arrive at a potential use site by dissemination, they are more likely to be utilized if the material is comprehensible and in relatively complete form. The need to seek clarifying information from the original disseminators is a barrier to use.

10. *Assessment of Evaluations.* Assessment of the outcomes of evaluation is an aid to the improvement of further evaluations. Knowledge of successes and failures permits reduction of unprofitable approaches and capitalization on strategies that bring desirable changes.

Evaluation can be assessed in terms of immediate impact on operations and also in terms of long-range influence through accumulating knowledge and through contributions to theory. It can also be assessed in relation to achievement of internal goals and objectives. Projects that successfully test hypotheses, demonstrate specific results, or establish definite structures of desired kinds are presumably useful even though they do not have dramatic impact upon day-to-day operations.

Evaluation can be assessed in terms of credibility, which examines internal and external validity, adequacy of the design for the problem at hand, successful adaptation to the structure and dynamics of the subject matter or setting, use of appropriate standards of comparison, adequacy of criteria in kinds and numbers, and quality of interpretation of the results of the evaluation.

Finally, evaluation can be assessed in terms of its design. Is its primary characteristic "strength" in the traditional sense, or is it a "weak" design, and if so, what are the implications in this particular setting?

11. *Development and Refinement of Evaluation.* Since evaluation is both a structure and a process, it is capable of the same developmental changes as other structures and processes, given feed-back from use and appropriate modification. This applies both to evaluation as a general process and to evaluation as a particular method or design. Judgments about the general process may suggest shifts in the mix of methods used, as is illustrated in the current exploratory movement toward systems methods. Similarly, judgments about the worth of a particular method may lead to its modification, as when acceptability of and familiarity with the quasi-experiment brings greater expertise in the creation of valid comparison groups.

C. Discussion

While the primary function of evaluation in criminal justice is to improve the functioning of criminal justice, there is at the same time an opportunity and a need to improve evaluation. By noting the manner in which the evaluation process works, the results it produces, judgments can be made about what techniques are effective and to what extent these techniques can be made even more effective.

Improvement of evaluation obviously involves more than refinement of technique. It touches upon the identification of worthy objectives, organization and coordination of effort, staffing and funding, choice of relevant theory, making of suitable short-term and long-term plans, selection of better methods and strategies, effective reporting and dissemination of results, and utilization of results. Each of these elements is a large area in itself.

In the recent past, evaluation in corrections and criminal justice has been an anarchic, trial-and-error process, with little leadership and not much sense of direction. The current pressures for rapid assessment of a vastly expanded criminal justice modernization drive finds evaluative research poorly equipped to perform the evaluation. Hopefully, one of the results of the effort will be a vast improvement in the structure and methods of evaluation itself.

REFERENCES AND NOTES

1. Walter C. Bailey, "An evaluation of 100 studies of correctional outcome," in Norman Johnston, Leonard Savitz and Marvin E. Wolfgang (eds.), *The Sociology of Punishment and Correction.* Second Edition. New York: Wiley, 1970, pp. 733–42.
2. Harold Orlans, *Making Social Research More Useful to Government.* Washington, D. C.: The Brookings Institution, Reprint 155, 1969, pp. 2–6. Reprinted from *Social Science Information,* 7 (December 1968), 151–58.
3. David A. Ward, "Evaluative research for corrections," in Lloyd E. Ohlin (ed.), *Prisoners in America.* Englewood Cliffs: Prentice Hall, 1973, 184–206, pp. 205–06.
4. Daniel Glaser, "The state of the art of criminal justice evaluation." Paper presented at the second annual meeting of the (Calif.) Association for Criminal Justice Research, November 1974.
5. Edward A. Suchman, *Evaluative Research.* New York: Russell Sage Foundation, 1967, pp. 171–76.
6. Carol H. Weiss, *Evaluation Research.* Englewood Cliffs: Prentice-Hall, 1972, Ch. 6.

CHAPTER 17. SUMMARY AND CONCLUSIONS

A review of the recent history of evaluative research in corrections discloses a number of pertinent developments:

1. In several innovative departments of correction, research units were created to aid the management process. These units, using variations on a social science research model, have carried out numerous descriptive, analytical, and evaluative studies of agency processes and outcomes.

2. More recently, many other operating and planning agencies have responded to various pressures to evaluate corrections. Additional research units have been set up or research services have been purchased from outside sources. In virtually all states, correctional evaluative studies have become commonplace.

3. Although a number of experimental studies, using random assignment to experimental and control statuses, have been conducted, correctional evaluations have been primarily non-experimental; i.e., case studies, surveys, before-after studies, and other non-rigorous designs.

4. Assessments of correctional evaluations indicate that only a small percentage of the total appear to demonstrate positive effects in "rehabilitating" offenders. The assessments have led some observers to conclude that no significant treatment efficacy has been shown in any programs evaluated.

5. A careful review of the best known of the assessments supports the conclusion that there is "payoff" or "impact" in at least a small percentage of the total body of studies. The level of payoff may in fact be as high as in some other fields of research and development endeavor, including high-technology industry and medicine.

6. The payoff in correctional evaluation thus far appears to occur more frequently in the form of "system improvement" than in "client improvement" or rehabilitation.

7. The biggest payoffs seem to come from "weak" or non-rigorous designs such as case studies or surveys, which are apparently useful in promoting system change. Lesser payoffs come from strong designs such as controlled experiments, which in several instances have conclusively demonstrated significant behavioral change in offenders.

8. The primary framework of correctional evaluation thus far has been the social science model, which features hypothesis testing, theory building, use of the controlled experiment, and project evaluation. In recent years this model has begun increasingly to share the stage with the industrial and policy science models. The latter feature operations research, cost-benefit studies, systems analysis, and stimulation.

9. The most productive evaluative research in corrections, for both weak and strong designs, appears to have been that of the in-house research unit. Despite their newness, a few of these units have produced evaluations superior in quality and higher in utility than the best products of academic researchers or consultant organizations. This edge in productivity seems to arise from the strategic location of the research units in relation to communications with operations staff, to planning and development, and to continuity of effort.

10. Consultant organizations and academic researchers have had significant roles to play in correctional evaluation in efforts that require special methodological or conceptual expertise. Some analytical tasks and some aspects of operations research or simulation, requiring mathematical languages and computer skills, have been performed for a number of correctional agencies by outside evaluators or researchers.

11. Although accomplishments in system change have apparently overshadowed those in offender change in correctional evaluation studies, notable breakthroughs have been occurring in the latter area also.

12. Efficacy has been demonstrated in programs for rehabilitating juveniles both in the community (e.g., the Provo Experiment and the Community Treatment Project)[1] and in the institution (e.g., PICO and the Youth Center Research Project).[2] There has been less evidence of success in rehabilitating adult offenders, although some system changes (e.g., the Parole Work Unit Program)[3] have demonstrated remarkable reductions in return-to-prison rates. The lesser evidence of success in rehabilitating adult offenders may have several possible explanations, including the lower level of research effort, lower sophistication of research design, lack of suitable typologies of adult offenders, the irrelevance of present treatment programs for adult offenders, and lesser amenability to change in adults.

13. A noteworthy aspect of evaluation in contemporary corrections is the growth of pressure for evaluation, the rising interest in improving the evaluation process, and the emergence of guidelines and models adapted more specifically to the needs of the correctional field.

From the present status of correctional evaluation and the character of the dominant trends in research, together with obvious assumptions about agency and community needs, it is reasonable to draw a number of conclusions:

1. Evaluative research in corrections is a young and inexperienced discipline, groping for direction in an increasingly complex setting. There is need for extensive rethinking and restructuring of the established evaluation methods to meet new demands and to keep pace with changes in the field.

2. Correctional evaluation has taken its early leads from academic and social science models, which are not well suited to the information and decision needs of complex systems in a changing world. It has begun to modify its strategies by borrowing substantially from industrial and policy science evaluation models. This development appears to be useful and needed.

3. Substantively, and for the short term, correctional evaluation needs to cope with problems of research overload, unwitting duplication of studies in different locales, inappropriate models of evaluation, and unplanned or hastily planned research efforts. Both administrators and evaluators have a responsibility for dealing directly with these problems.

4. Traditional evaluation guidelines, which have proved only moderately useful in the shaping of the research effort, need to be reviewed and restated if evaluative research is to meet the demands that are being placed on it. In the reconstruction of the guidelines, the results of evaluation need to be analyzed systematically for clues to the reconstruction. Studies of research impact and research utilization need to be carried out to achieve more effective management of research and to obtain higher rates of return from the evaluative effort.

5. The importance of impact analysis in correctional evaluation is indicated by the discovery that "non-rigorous" research designs are apparently more influential than rigorous designs in major decision-making situations. Although the reasons for this are not fully understood, some hypotheses may be stated: a) these non-rigorous studies fit better the decision making styles and needs of administrators; b) there is greater pressure on corrections for system improvement than for client improvement, and these studies provide adequate rationales for system change; c) in time of rapid change, conditions are not favorable for the use of strong research designs; and d) correctional administrators have not yet supported rigorous designs to the extent required to make them generally effective.

6. The discovery regarding impact from weak studies points to another need in corrections—the clarification of goals. The fact that evaluation impact has shown itself primarily in system change is in part a reflection of the obvious need for system revision in corrections and criminal justice. The importance of system change as an objective of evaluation has been obscured by preoccupation with offender rehabilitation as an objective.

7. Much of the current deprecation of the "inefficacy" of correctional programs and correctional research derives from a perception of improved behavior as the focus of evaluative studies. The productivity or the efficacy of research appears in a better light when both system change and offender change are taken into account as objectives or goals.

8. Still other goals require consideration when guidelines for correctional evaluation are drafted. In the long run, increased importance seems foreshadowed for such criminal justice goals as "restoration of the victim" and "equity in criminal justice involvement and processing." As these and other goals come in for discussion and targeting, the present objectives of punishment, incapacitation, deterrence and rehabilitation will undoubtedly undergo some revaluation.

9. To the extent that causal analysis of behavioral change under alternative conditions remains an important strategy of correctional evaluation, the controlled experiment will continue to be a potentially useful study design. In practice, this design has not been as productive in corrections as its pres-

tige in science would appear to promise. Its productivity might be enhanced if "experimental" administrators and adequately trained evaluators cooperated to insure its proper use in appropriate situations.

10. Use of the controlled experiment should occur under full awareness that preoccupation with this method may interfere with the orderly development of more generally useful evaluation methods and their application to equally important criminal justice goals. Continuous assessment of the results of evaluation will be helpful in maintaining a constructive balance in the use of research designs.

11. A fundamental need of the correctional administrator is to know in detail the extent and results of his agency's efforts and how to strive realistically to improve the level of these results. This calls for operations research, which traces inputs and outcomes and analyzes the relation of outcomes to processes. This evaluation model, which has entered corrections as a supplement to the social science model, may overshadow that model after information systems become widespread. Operations research is perhaps the most rational way of routinizing the evaluation process—a strategy that is coming to be recognized as a necessity for modern correctional agencies.

12. As in the case of operations research, cost-benefit analysis is assuming an increasingly important role in correctional evaluation. On occasion, it shows impressive cost savings or returns even when there is no evidence of behavioral change from the program under evaluation. Judging from as yet limited data, the most dramatic cost-benefit gains are from programs that provide alternatives to institutionalization—street work with gang members, enrichment of probation, and rationalization of parole.

13. For correctional evaluator-planners, emphasis needs to be directed toward definition of corrections as a subsystem within the criminal justice system and also within the wider systems of the larger community. Under this view, the organization of evaluative research within the corrections subsystem will be largely operations research, supported by carefully planned information systems, and aimed at optimizing results within the subsystem. Within the larger systems, evaluative research will be a coordinated effort in conjunction with the state planning agency and other entities, designed to assess alternative concepts of correction within changing criminal justice and community systems.

14. At all subsystem and system levels, there is need for evaluative research to focus increasingly on system change rather than on offender change until there is better balance between these two orientations and better understanding of the roles of individual and system in the origin and maintenance of crime.[4]

15. The likely emergence of new goals in criminal justice and corrections and changing relationships at the criminal justice and community levels places a premium on flexibility in the use and development of evaluative research. Some assurance of flexibility can be gained by continual assessment of the procedures and results of evaluation. This assurance can be enhanced by sharpening administrators' awareness of and skills for meeting their responsibilities for evaluation. It can be further enhanced by raising the qualifications of evaluators through selection, special in-house and workshop training, and by provision of exemplary case materials and guidelines to evaluation.

Some of the foregoing conclusions and recommendations are short-term and some are obviously long-term. No time table can be set for either group. The pace at which research can be conceptualized and implemented will depend on social and political climates, the press of criminal justice and community problems, and the existence of administrators and evaluators who can devise, try and assess new concepts of program and system operation.

The kinds of goals to set for either the immediate or long-term effort cannot easily be specified. There is some tendency at present to name offender rehabilitation as the first priority of corrections. The rationale for this was clearly stated in a recent report by a House of Representatives subcommittee:

> "Because probably two-thirds of all serious crimes are committed by persons who have been in jail, more effective rehabilitation would do more than any other single factor to reduce the crime rate." [5]

However, more is involved in reducing the crime rate than discovering the arithmetic of recidivism. There is admittedly much more to learn about the modification of offender behavior through treatment or enlightened supervision. And we have well-documented instances of juvenile rehabilitation, in both institution and community, and marked reduction in rates at which adult parolees are returned to prison. Nevertheless, we may still ask whether in the long run crime rates will diminish as much from improvements in rehabilitation and reintegration as from the development of better systems of socialization and control in the larger community.

The possibility that the latter course will be the more productive might be inferred from the relative rates of offender change and system change in corrections today. If this conclusion is sound, we must then ask what directions we should take in seeking out and setting up the new systems that will cope more adequately with the crime problem.

There is some competition here between the theorists, the methodologists and the pragmatists. Thus far the latter may have the best claim to furthering correctional progress since most of what has been achieved appears to have derived largely from intuition and judgment. How much farther we can proceed on these bases is now a good question.[6] A reasoned judgment is that the task ahead requires a concerted effort in which pragmatists, methodologists and theorists all collaborate, each contributing where he can.

REFERENCES AND NOTES

1. LaMar T. Empey and Maynard L. Erickson, *The Provo Experiment.* Boston: Heath, 1972; Marguerite Q. Grant, *The Community Treatment Project After Five Years.* Sacramento: California Youth Authority, 1966.
2. Stuart Adams, *Effectiveness of Interview Therapy with Older Youth Authority Wards: An Interim Evaluation of the PICO Project.* Sacramento: California Youth Authority, Research Report No. 21, 1961; Ted B. Palmer, *The Community Treatment Project in Perspective.* Sacramento: The California Youth Authority, 1973.
3. *The 1973 Report on the Work Unit Parole Program.* Sacramento: California Department of Corrections, March 1974; Walter R. Burkhart, "The parole work unit programme: an evaluation report," *British Journal of Criminology,* April 1969, 125–47.
4. For a discussion of the relative ease of solving social problems by changing people and by changing their circumstances, see Amitai Etzioni, "Human beings are not very easy to change after all." *The Society: Saturday Review,* June 3, 1972, 45–47.
   A contrasting view is seen in Herbert Quay, "What corrections can correct and how," *Federal Probation,* 37 (June 1973), 3–5. Quay asserts that "What corrections can correct, given the appropriate circumstances, are those behaviors related to offending." Conceivably, however, corrections can do more, including a) changing itself, b) changing offender behavior, c) changing offender personalities, and d) changing the offender's environment. It becomes a point of interest, then, as to whether in the future corrections may not shift its emphasis somewhat from trying to change "behaviors relating to offending" toward trying to change offender's "circumstances."
5. Cited from a forthcoming report of the House Judiciary Committee, Subcommittee on Courts, Civil Liberty and Administration of Justice, in the article, "Panel says an informed public could demand prison reform." *New York Times,* 7 March 1974, p. 18–C.
6. Not much farther, Leslie T. Wilkins would conclude. He sees the future of progress in crime control as dependent primarily on the theorists. See Wilkins, "Crime and criminal justice at the turn of the century," *Annals of the American Academy of Political and Social Science: The Future Society: Aspects of America in the Year 2000,* 408 (July 1973), 13–29.

BIBLIOGRAPHY I: ARTICLES, PAPERS AND REPORTS

A. Philosophy, Objectives, and Organization of Evaluation

1. Abelson, R. P. "Simulation of social behavior," in G. Lindzey and E. Aronson (eds.), *Handbook of Social Psychology*, (revised edition), Vol. 2. Reading, Mas.: Addison-Wesley, 1968.
2. Anderson, S. "Operations research in public health," in Herbert C. Shulberg, Alan Sheldon and Frank Baker (eds.), *Program Evaluation in the Health Fields*. New York: Behavioral Publications, 1969, pp. 254–67.
3. Angell, Robert C. "The ethical problems of applied sociology," in Paul F. Lazarsfeld, William H. Sewell and Harold L. Wilensky (eds.), *Uses of Sociology*. New York: Basic Books, 1967, pp. 725–40.
4. Bennett, Lawrence A. "Should we change the offender or the system?" *Crime and Delinquency*, 19 (July 1973), 332–42.
5. Bogusław, Robert. "Systems concepts in social systems," in Ralph F. Miles, Jr. (ed.), *Systems Concepts: Lectures on Contemporary Approaches to Systems*. New York: Wiley, 1973, pp. 177–90.
6. Bradley, Harold B. "Designing for change: problems of planned innovation in corrections," *Annals of the American Academy of Political and Social Science: The Future of Corrections*, 381 (January 1969), 89–98.
7. Browning, Rufus P. "Simulation: attempts and possibilities," in Jeanne K. Knutson (ed.), *Handbook of Political Psychology*. San Francisco: Jossey-Bass, 1973, pp. 383–412.
8. Churchman, C. West. "A critique of the systems approach to social organizations," in Ralph F. Miles, Jr (see reference 5, above), pp. 191–205.
9. Denzin, Norman K. "Politics and promises: the values of social science," in Norman K. Denzin (ed.), *The Values of Social Science*. Chicago: Aldine, 1970, pp. 1–21.
10. Dror, Yehezkel. "Applied social science and systems analysis," in Irving Louis Horowitz (ed.), *The Use and Abuse of Social Science*. New Brunswick, N. J.: E. P. Dutton, 1971, pp. 109–132.
11. Etzioni, Amitai. "Two approaches to organizational analysis," in Herbert C. Schulberg et al (see reference 2, above), pp. 101–20.
12. ———. "Human beings are not very easy to change after all," *The Society: Saturday Review*, June 3, 1972, 45–47.
13. ———. "Redirecting research dollars," *Washington Post*, 11 June 1972.
14. Gans, Herbert J. "Where sociologists have failed," in Norman K. Denzin (see reference 9, above), pp 83–86.
15. ———. "Social science for social policy," in Irving Louis Horowitz (see reference 10, above), pp. 13–33.
16. Hyman, Herbert H. and Charles R. Wright. "Evaluating social action programs," in Paul F. Lazarsfeld et all (see reference 3, above), pp. 741–82.
17. Kish, Leslie. "Some statistical problems in research design," in Edward R. Tufte (ed.), *The Quantitative Analysis of Social Problems*. Menlo Park, Calif.: Addison-Wesley, pp. 391–406.
18. Lang, Kurt. "Pitfalls and politics in commissioned policy research," in Irving Louis Horowitz (see reference 10, above), pp. 212–33.
19. McGee, Richard A. "What's past is prologue," *Annals of the American Academy of Political and Social Science: The Future of Corrections*, 381 (January 1969), 1–10.
20. Montilla, M. Robert. "Opportunities for action research in community correction," *Journal of Research in Crime and Delinquency*, 6 (July 1969), 123–44.
21. Morris, Norval, and Frank Zimring. "Deterrence and correction," *Annals of the American Academy of Political and Social Science: The Future of Corrections*, 381 (January 1969), 137–46.
22. Morton, Jack A. "From research to technology," in David Alison (ed.), *The R & D Game*. Cambridge: Massachusetts Institute of Technology Press, 1969, pp. 213–35.
23. Orlans, Harold. "Making social research more useful to government," *Social Science Information*, 7 (December 1968), 151–58.
24. Quay, Herbert C. "What corrections can correct and how," *Federal Probation*, 37 (June 1973), 3–5.
25. Ramo, Simon. "The systems approach," in Ralph F. Miles, Jr. (see reference 5, above), pp. 13–32.
26. Rein, Martin, and S. M. Miller. "Social action on the installment plan," in Norman K. Denzin (see reference 9, above), 97–118.
27. Rossi, Peter H. "Booby traps and pitfalls in the evaluation of social action programs," in *Proceedings of the Social Statistics Section of the American Statistical Association*. Washington, D. C.: American Statistical Association, 1966.
28. ———. "Evaluating social action programs," in Norman K. Denzin (see reference 9, above), pp. 87–96.
29. Schon, Donald A. "The fear of innovation," in David Alison (see reference 22, above), pp. 119–34.
30. Schorr, Alvin L. "Public policy and private interest," in Irving Louis Horowitz (see reference 10, above), pp. 155–69.
31. Schrag, Clarence. "The correctional system: problems and prospects," *Annals of the American Academy of Political and Social Science: The Future of Corrections*, 381 (January 1969), 11–20.
32. Tropman, John E., and Karl H. Gohlke. "Cost/benefit analysis—comprehensive planning in the criminal justice system," *Crime and Delinquency*, 19 (July 1973), 315–22.
33. Wilkins, Leslie T. "Crime and criminal justice at the turn of the century," *Annals of the American Academy of Political and Social Science: The Future Society: Aspects of America in the Year 2000*, 408 (July 1973), 13–29.
34. Wilkins, Leslie T., and Thomas Gitchoff. "Trends and projections in social control systems," *Annals of the American Academy of Political and Social Science: The Future of Corrections*, 381 (January 1969), 125–36.
35. Yarmolinsky, Adams. "The policy researcher: his habitat, care and feeding," in Irving Louis Horowitz (see reference 10, above), 196–211.

B. Techniques and Applications

1. Adams, Stuart. "The PICO Project," in Norman Johnston, Leonard Savitz and Marvin E. Wolfgang (eds.), *The Sociology of Punishment and Correction* (second edition). New York: Wiley, 1970, pp. 548–61.
2. ———. "A cost approach to the assessment of gang rehabilitation techniques," *Journal of Research in Crime and Delinquency*, 4 (January 1967), 166–82. See also Norman Johnston et al (reference 1, above), pp. 657–70.
3. ———. "Some findings from correctional caseload research," *Federal Probation*, 31 (December 1967), 48–57; also, Norman Johnston et al (see reference 1, above), pp. 721–32.
4. ———. "Measurement of effectiveness and efficiency in corrections," in Daniel Glaser (ed.), *Handbook of Criminology*. Chicago: Rand McNally, 1974, Ch. 29.
5. Adams, Stuart, Wendell Plair and John D. Spevacek. *The Use a Predictive Scoring System to Improve Selection for the DCDC Work Release Program*, Washington, D. C.: Department of Corrections, Research Memorandum 71–7, November 1971.
6. Babst, Dean V., Don M. Gottfredson and Kelly B. Ballard, Jr. "Comparison of multiple regression and configuration analysis techniques for developing base expectancy tables," *Journal of Research in Crime and Delinquency*, 5 (January 1968), 72–80.
7. Campbell, Donald T. "Reforms as experiments," *American Psychologist*, 24 (April 1969), 402–29. See also, in Bibliography II, below, Caporaso et al, 1973; Caro, 1971; Weiss, 1972a.
8. Carney, Francis J. "Correctional research and correctional decision-making," *Journal of Research in Crime and Delinquency*, 6 (July 1969), 110–22.
9. Churchill, Lindsey. "Systems analysis confronts crime," in Jack D. Douglas (ed), *Crime and Justice in American Society*, Indianapolis: Bobbs-Merrill, 1971, pp. 269–87.
10. Furash, Edward E. "The problem of technology transfer," in Raymond A. Bauer and Kenneth J. Gergen (eds.), *The Study of Policy Formation*. New York: Free Press, 1968, pp. 281–328.
11. Jesness, Carl F. *The Youth Center Research Project*. Sacramento: American Justice Institute, 1972, 2 vols.
12. Klein, Malcolm W., Solomon Kobrin, A. W. McEachern and Herbert R. Sigurdson. "System rates: an approach to comprehensive criminal justice planning," *Crime and Delinquency*, 4 (October 1971), 355–72.
13. McEachern, Alexander W., and J. Robert Newman. "A system for computer-aided probation decision-making," *Journal of Research in Crime and Delinquency*, 6 (July 1969), 184–98.
14. McGlothlin, William H., Victor C. Tabbush, Carl D. Chambers, and Kay Jamison. *Alternative Approaches to Opiate Addiction Control: Costs, Benefits and Potential*. Los Angeles: UCLA Department of Psychology, 1972. Report to the Bureau of Narcotics and Dangerous Drugs.
15. Palmer, Ted B. *The Community Treatment Project in Perspective*. Sacramento: California Youth Authority, 1973.
16. Robins, Lee N., and Shirley Y. Hill. "Assessing the contributions of family structure, class and peer groups to juvenile delinquency," *Journal of Criminal Law, Criminology and Police Science*, 57 (September 1966), 325–31.
17. Utterback, James H. "Innovation in industry and the diffusion of technology," *Science*, 183 (February 1974), 620–26.
18. Warren, Marguerite Q. *The Community Treatment Project After Five Years*. Sacramento: California Youth Authority, 1966. Reprinted as "The community treatment project," in Norman Johnston et al (see reference 1, above), pp. 671–83.

BIBLIOGRAPHY II. BOOKS AND MONOGRAPHS

1. Ackoff, Russell L. *Scientific Method: Optimizing Applied Research Decisions*. New York: Wiley, 1962.
2. Ackoff, Russell L. and Patrick Rivett. *A Manager's Guide to Operations Research*. New York: Wiley, 1963.
3. American Institutes for Research. *Evaluative Research: Strategies and Methods*. Pittsburg: American Institutes for Research, 1970.
4. Bauer, Raymond A. (ed.). *Social Indicators*. Cambridge: Massachusetts Institute of Technology Press, 1966.
5. Campbell, Donald T. and Julian C. Stanley. *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally, 1966.
6. Caporaso, James A. and Leslie L. Roos, Jr. (eds.). *Quasi-Experimental Approaches: Testing Theory and Evaluating Policy*. Evanston: Northwestern University Press, 1973.
7. Caro, Francis (ed.). *Readings in Evaluation Research*. New York: Russell Sage Foundation, 1971.
8. Chartrand, Robert Lee. *Systems Technology Applied to Social and Community Problems*. New York: Spartan Books, 1971.
9. Churchman, C. West, Russell Ackoff and E. Leonard Arnoff *Introduction to Operations Research*. New York: Wiley, 1957.
10. DeGreene, Kenyon B. *Sociotechnical Systems: Factors in Analysis, Design, and Management*. Englewood Cliffs, N. J.: Prentice-Hall, 1973.
11. Dorfman, Robert (ed.). *Measuring Benefits of Govern-*

ment Investments. Washington, D. C.: Brookings Institution, 1965.

12. Drucker, Peter F. Chapters 29–40, *Management: Tasks, Responsibilities, Practices.* New York: Harper & Row, 1973.
13. English, J. Morley (ed.). *Cost-Effectiveness: The Economic Evaluation of Engineered Systems.* New York: Wiley, 1968.
14. Etzioni, Amitai and Eva Etzioni. *Social Change.* New York: Basic Books, 1964.
15. Fairweather, George W. *Methods for Exprimental Social Innovation.* New York: Wiley, 1967.
16. Ferman, Louis A. (ed.). "Evaluating the War on Poverty," *Annals of the American Academy of Political and Social Science,* 385 (September 1969)9.
17. Fisher, Gene H. *Cost Considerations in Systems Analysis.* New York: American Elsevier, 1971.
18. Forcese, Dennis P. and Stephen Richer (eds.). *Stages of Social Research: Contemporary Perspectives.* Englewood Cliffs, N. J.: Prentice-Hall, 1970.
19. Franklin, Billie J. and Harold W. Osborne. *Research Methods: Issues and Insights.* Belmont, Calif.: Wadsworth, 1971.
20. Freeman, Howard E. and Clarence E. Sherwood. *Social Research and Social Policy.* Englewood Cliffs, N. J.: Prentice-Hall, 1970.
21. Gerstenfeld, Arthur. *Effective Management of Research and Development.* Reading, Mass.: Addison-Wesley, 1970.
22. Ginzberg, Eli and Robert M. Solow (eds.). "The great society: lessons for the future," *The Public Interest,* 34 (Winter 1974).
23. Glaser, Daniel. *Routinizing Evaluation: Getting Feedback on Effectiveness of Crime and Delinquency Programs.* Rockville, Md.: National Institute of Mental Health, 1973.
24. Hatry, Harry P., Richard E. Winnie and Donald M. Fisk. *Practical Program Evaluation for State and Local Governments.* Washington, D. C.: Urban Institute, 1973.
25. Haveman, Robert H. and Julius Margolis (eds.). *Public Expenditures and Policy Analysis.* Chicago: Markham, 1970.
26. Herzog, Elizabeth. *Some Guidelines for Evaluative Research.* Washington, D. C.: U. S. Department of Health, Education, and Welfare, Children's Bureau, 1959.
27. Hitch, Charles J. Decision-Making for Defense. Berkeley: University of California Press, 1967.
28. Hinrichs, Harley H. and Graeme M. Taylor (eds.). *Program Budgeting and Benefit-Cost Analysis.* Pacific Palisades, Calif.: Goodyear, 1969.
29. Institute for Defense Analysis. *Task Force Report: Science and Technology.* Washington, D. C.: President's Commission on Law Enforcement and Administration of Justice, 1967.
30. Isaac, Stephen and William B. Michael. *Handbook in Research and Evaluation.* San Diego: Knapp, 1971.
31. Kendall, M. G. *Cost-Benefit Analysis.* New York: American Elsevier, 1971.
32. Michael, Donald N. *On Learning to Plan—and Planning to Learn.* San Francisco: Jossey-Bass, 1973.
33. Morse, Philip (ed.). *Operations Research for Public Systems.* Cambridge: Massachusetts Institute of Technology Press, 1967.
34. Moursund, Janet P. *Evaluation: An Introduction to Research Design.* Monterey, Calif.: Brooks/Cole, 1973.
35. Mullen, Edward J., James R. Dumpson and Associates. *Evaluation of Social Intervention.* San Francisco: Jossey-Bass, 1972.
36. Orlans, Harold. *Contracting for Knowledge.* San Francisco: Jossey-Bass, 1973.
37. Rivlin, Alice M. *Systematic Thinking for Social Action.* Washington, D. C.: Brookings Institution, 1971.
38. Rossi, Peter H. and Walter Williams (eds.). *Evaluating Social Programs: Theory, Practice, and Politics.* New York: Seminar Press, 1972.
39. Schulberg, Herbert C., Alan Sheldon and Frank Baker (eds.). *Program Evaluation in the Health Fields.* New York: Behavioral Publications, 1969.
40. Suchman, Edward A. *Evaluative Research: Principles and Practice in Public Service and Social Action Programs.* New York: Russell Sage, 1967.
41. Tripodi, Tony, Phillip Fellin and Irwin Epstein. *Social Program Evaluation.* Itasca, Ill.: Peacock, 1971.
42. Tripodi, Tony, Phillip Fellin and Henry J. Meyer. *The Assessment of Social Research.* Itasca, Ill.: Peacock, 1969.
43. Tufte Edward R. (ed.). *The Quantitative Analysis of Social Problems.* Reading, Mass.: Addison-Wesley, 1970.
44. Twain, David, Eleanor Harlow and Donald Merwin. *Research and Human Services: A Guide to Collaboration for Program Development.* New York: Research and Development Center, Jewish Board of Guardians, 1970.
45. Weiss, Carol (ed.). *Evaluating Action Programs: Readings in Social Action and Education.* Boston: Allyn & Bacon, 1972a.
46. ———. *Evaluation Research: Methods of Assessing Program Effectiveness.* Englewood Cliffs, N. J.: Prentice-Hall, 1972b.
47. Wholey, Hoseph S., John W. Scanlon, Hugh G. Duffy, James S. Fukumoto, and Leona M. Vogt. *Federal Evaluation Policy.* Washington, D. C.: Urban Institute, 1971.
48. Wilkins, Leslie T. *Evaluation of Penal Measures.* New York: Random House, 1969.
49. Williams, Walter. *Social Policy Research and Analysis: The Experience in Federal Agencies.* New York: Elsevier, 1971.

BIBLIOGRAPHY III. MISCELLANEOUS

1. California Department of Corrections. *Annual Research Review.* Annually since 1960.
2. California Youth Authority. *Status of Current Research in the California Youth Authority.* Annually since 1960.
3. *Crime and Delinquency Research in Selected European Countries.* Rockville, Md.: National Institute of Mental Health, 1971.
4. District of Columbia Department of Corrections. *Summary of Completed Research: 1967–70,* and *Summary of Completed Research: 1970–71.*
5. Evaluation. Vols. 1 and 2, all Nos., 1973 and 1974.
6. Joint Commission on Correctional Manpower and Training. *Research on Correctional Rehabilitation.* Washington, D. C.: The Joint Commission, 1967.
7. Morris, Albert. *A Correctional Administrator's Guide to the Evaluation of Correctional Problems.* Boston: Massachusetts Correctional Association, Bulletin No. 21, 1971.

AUTHOR INDEX

SUBJECT INDEX

☆ U.S. GOVERNMENT PRINTING OFFICE: 1975 O—564-300

(CUT ALONG THIS LINE)

PRESCRIPTIVE PACKAGE: "Evaluative Research in Corrections:

A Practical Guide"

To help LEAA better evaluate the usefulness of Prescriptive Packages, the reader is requested to complete and return the following questions.

1. What is your general reaction to the Prescriptive Package?
   - ☐ Excellent
   - ☐ Above Average
   - ☐ Average
   - ☐ Poor
   - ☐ Useless

2. Does this package represent best available knowledge and experience?
   - ☐ No better single document available
   - ☐ Excellent, but some changes required (please comment)
   - ☐ Satisfactory, but changes required (please comment)
   - ☐ Does not represent best knowledge or experience (please comment)

   Comments:

3. To what extent do you see the package as being useful in terms of: (check one box on each line)

	Highly Useful	Of Some Use	Not Useful
Modifying existing projects	☐	☐	☐
Training personnel	☐	☐	☐
Administering on-going projects	☐	☐	☐
Providing new or important information	☐	☐	☐
Developing or implementing new projects	☐	☐	☐

4. To what specific use, if any, have you put this particular package?

5. In what ways, if any, could the package be improved? (specify)

A. Structure/Organization

B. Content/Coverage

C. Objectivity

D. Writing Style

6. How did this package come to your attention? (check one or more boxes)
   - ☐ LEAA Mailing of package
   - ☐ Your organization's library
   - ☐ Contact with LEAA staff
   - ☐ LEAA Newsletter
   - ☐ National Criminal Justice Reference Service
   - ☐ Other _____

7. Additional Comments:

8. Check ONE item below which best describes your affiliation with law enforcement or criminal justice. If the item checked has an asterisk (*), please also check the related level, i.e.,

☐ Federal ☐ State	☐ County ☐ Local
☐ Headquarters, LEAA	☐ Police *
☐ LEAA Regional Office	☐ Court *
☐ State Planning Agency	☐ Correctional Agency *
☐ Regional SPA Office	☐ Legislative Agency *
☐ College/University	☐ Other Government Agency *
☐ Commercial Industrial Firm	☐ Professional Associations *
☐ Citizen Group	☐ Crime Prevention Group *

9. Your Name (Optional) _____
Organization or Agency _____
Your Position _____

(FOLD)

U. S. DEPARTMENT OF JUSTICE
LAW ENFORCEMENT ASSISTANCE ADMINISTRATION
WASHINGTON, D.C. 20530

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF JUSTICE
JUS-436

U.S.MAIL

(CUT ALONG THIS LINE)

Director
Office of Technology Transfer
National Institute of Law Enforcement
and Criminal Justice
U.S. Department of Justice
Washington, D.C. 20530

(FOLD)

END