61200

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

Interim Analysis of 200 Evaluations

On Criminal Justice

Richard C. Larson, Project Director

Vicki M. Bier

Edward H. Kaplan

Cheryl Mattingly

Timothy J. Eckels

Nancy Reichman

Leni S. Berliner

Massachusetts Institute of Technology
Operations Research Center
Cambridge, Massachusetts 02139

May, 1979

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

This document presents interim findings from one component of LEAA Grant 78NI-AX-0007, "An Empirical Study of Methods Used in Criminal Justice Evaluations." Overall research on this grant, under direction of Professor Richard C. Larson at the Massachusetts Institute of Technology, seeks to begin to answer the following types of questions:

1. What is the "state of the art" in criminal justice evaluation, as practiced?

2. How do attributes of evaluations (e.g., methodology, use of results) vary with program type and sponsoring agency?

3. Are evaluations as actually carried out sufficiently flexible to allow for necessary changes in the program being evaluated? Does information obtained during the course of an evaluation feed back to modify in any way either evaluation procedures or program conduct?

4. How are alternative hypotheses, theories, or models generated and utilized in evaluations?

5. In what ways do "textbook" evaluation methodologies provide useful guidance and/or fall short when applied in practice? Are there identifiable gaps in current methodologies? Can we begin to fill them?

6. Can we begin to construct a comprehensive procedure for evaluating evaluations themselves? How are evaluators to be held accountable for their product?

This report presents preliminary findings relating to questions 1-4 above (i.e., a review of current evaluation practices in criminal justice). It is based on a sample of 200 criminal justice evaluations that have been performed within the last decade. The sample was selected from approximately 1500 abstracts of studies provided by LEAA's NCJRS (National Criminal Justice Reference Service) via a key word search on "evaluations" conducted in late 1977.[1]

Insofar as possible, it is desirable to evaluate an evaluation

comprehensively. In the parlance of the evaluation literature, a comprehensive evaluation is one which analyzes program inputs, process and outcome. Thus, a comprehensive evaluation of an evaluation would include an analysis of the resources (human and material) and methodologies (design) available to the evaluation effort (inputs), the actual conduct of the evaluation (process) and a listing of decisions influenced by the evaluation (outcome). Detailed entries under each category are given in Table 1. This framework for evaluating an evaluation, relevant to question 6 above, is treated more fully in a forthcoming document.[2] For our present purposes it is important to note that the final report of an evaluation is unlikely to capture most of the information about an evaluation necessary to evaluate it comprehensively. The most important information needed is that regarding the outcome of an evaluation. An evaluation, and sometimes the evaluation report, provide information about a program. This information is only useful to the extent that it influences (informs or modifies) decisions. Thus, the only actual outcomes of an evaluation are those decisions based in part on its findings.

This report summarizes our preliminary analysis of the 200 documents produced by the 200 evaluations in our sample. It is labeled interim because we are currently attempting to contact (via follow-up questionnaires) as many of the evaluators who conducted the evaluations as possible, the key staff personnel whose programs were evaluated, and the contacts or grant monitors in the organizations that funded the evaluations. We have found, not surprisingly, that the documents produced in an evaluation constitute only one output of an evaluation: they often tell little of the inputs to an evaluation, the process by which it was undertaken, or the ultimate impact of the evaluation.

TABLE 1

COMPONENTS OF A COMPREHENSIVE EVALUATION OF EVALUATIONS

Evaluation Inputs: An inventory of resources and methodologies brought
to bear on the evaluation

   (*i*)   Budget of the evaluation (and other material resources available
to the evaluators);

  (*ii*)   Duration of the evaluation;

 (*iii*)   Timing of the evaluation with respect to the program being evaluated;

  (*iv*)   Attributes of evaluation personnel (e.g., training, experience,
'world view');

   (*v*)   Attributes of program personnel (e.g., experience, commitment,
education);

  (*vi*)   Attributes of the program (e.g., goals, substantive area of criminal
justice, client group);

 (*vii*)   Evaluation methodology and design;

(*viii*)   Audience, or "client group" for the evaluation and purposes of
the evaluation.

Evaluation Process: Actual conduct of the evaluation compared with that
planned in the evaluation design

   (*i*)   Types, intensity and frequency of interaction of evaluators with
program staff;

  (*ii*)   Extent to which information acquired during evaluation is fed back
to program staff, perhaps modifying program procedures;

 (*iii*)   Extent to which information acquired during evaluation is used to
modify the allocation of evaluation resources;

  (*iv*)   Turnover in personnel (e.g., evaluators, program staff, client group);

   (*v*)   Response of program staff and client group to the presence of
evaluators;

TABLE 1 (continued)

Evaluation Process (continued)

(vi) Adaptiveness of evaluation design (capacity to respond to changes in the program);

(vii) Testing of hypotheses regarding the program;

(viii) Documentation of findings.

Evaluation Outcomes: A listing of decisions influenced by the evaluation

(i) Decision by funding agency to fund, refund, modify, or cancel program;

(ii) Decision by program staff to modify any of the program procedures;

(iii) Decision by members of client group to alter participation patterns in program;

(iv) Decision by one or more members of the research community to study further the questions/issues raised in the evaluation;

(v) Decision by one or more other funders and/or program personnel (in other jurisdictions) to initiate, modify, or terminate similar programs.

The Sample of 200

The 200 evaluations in our sample, while primarily obtained from a list of 1500 "evaluation" studies cited by NCJRS, constitute a structured rather than a random sample. We used the NCJRS classification of "evaluations" deliberately, in order to point up the weaknesses of that classification system. (Pilot studies, simulations and "evaluability assessments" appear as "evaluations" under the NCJRS system.) Pursuant to our interest in methodology, we deliberately selected approximately 50 percent of the sample from "logistical" programs, that is, programs in which the movement of persons, materiel or other entities was an important element. It was thought that structural models of the system being evaluated would be most relevant for these evaluations; as de facto evaluators of these evaluations, we are interested in the use, abuse, or misuse of such models in the evaluation process. The remaining 50 percent of the sample was drawn primarily from social service type programs, in which counseling or some other type of service is provided to one or more client groups. Examples include: special court-assigned public defenders; certain pre-trial diversion projects; certain police public relations programs; volunteers in probation; referred services for persons requiring assistance for family troubles, alcoholism, drugs, etc. It was thought that delineation of goals, measurement of achievement, and models of process would be considerably more difficult in the social service area than in the logistical area. Hence, evaluation design, conduct, and outcome should be dependent on this distinction.

We also made an attempt to include elements of three LEAA evaluation efforts in our sample.

(1) <u>National Evaluation Program (NEP) Phase I Studies</u>.
    NEP evaluation assessments to date have been commissioned in
    approximately 20 different criminal justice areas. Their
    purpose is to evaluate the information available about a
    particular topic area (e.g., traditional police preventive
    patrol) on a national scale. Thus, they do not constitute
    evaluations of specific local programs. A listing of the
    NEP Phase I evaluations in our sample is given in Table 2.

(2) <u>"Exemplary Projects" and "Exemplary Validation Reports"</u>.
    LEAA has encouraged criminal justice projects throughout the
    country to be nominated for "exemplary" status. This de-
    signation, when approved, implies that replication of the
    program is recommended on a wider scale. The exemplary
    project and validation reports included in our sample are
    given in Table 3.

(3) <u>LEAA Impact Cities Program</u>. In the early 1970's LEAA
    designated eight cities as Anti-Crime Impact Cities,
    funding each at a level of $20 million to design and im-
    plement a major anti-crime program. We have included 8
    evaluations of components of this program in our sample.

In structuring the sample to reflect various methodological concerns,

we were particularly interested in those methodologies which are foci of

current research. These include time series analysis, experimental de-

sign, use of models in evaluation, and use of decision analysis in eval-

uation. Thus, we tended to favor evaluations which purported to use one

or more of these methodologies.

In structuring the sample in areas related to research hypotheses

of our grant, we were particularly concerned with the following issues:

(1) use of evaluations by decision-makers, both long-term and short-term,

and even identification of appropriate decision-makers in the report;

(2) the likely value of information to be acquired from an evaluation;

(3) misuse and abuse of quantitative methods and (4) use of adaptive

evaluation methods to respond to feedback from the field.

TABLE 2

NATIONAL EVALUATION PROJECT PHASE I STUDIES EXAMINED IN THIS STUDY

Banks, J. (1976) Knowledge Assessment: Phase I Evaluation of
  Intensive Special Probation Projects.

Kreihdel, Burton, et al., (1977) Court Information Systems: Phase
  I Report.

National Center for State Courts (1976) Assessment of the Present
  State of Knowledge Concerning Pre-trial Release Programs:
  Phase I Evaluation of Pretrial Release Programs.

Rutherford, Andrew and McDermott, Robert (1976) Juvenile Diversion:
  National Evaluation Program Phase I Summary Report.

Thomas, et al., (1976) Phase I Evaluation of Pretrial Release Programs:
  Assessment of the Present State of Knowledge.

Walker, J.P. (1976) Theory and Practice of Delinquency Prevention in
  the United States: Review, Synthesis and Assessment.

Table 3a

EXEMPLARY PROJECTS INCLUDED IN THIS STUDY

Anon. (1975) Philadelphia Neighborhood Youth Resources Center --
    An Exemplary Project.

Blew, C.H., et al., (1977) Project New Pride, Denver, Colorado

Carlson, K. (1977) One Day/One Trial: Jury System, Wayne County,
    Michigan.

Halper, A. (1975) Administrative Adjudication Bureau (AAB), New York
    State Department of Motor Vehicles.

McDonnell, John J. (1975) Central Police Dispatch (CPD), Muskegon
    County, Michigan.

McGillis, Daniel (1977) Major Offense Bureau, Bronx County District
    Attorney's Office, New York.

Zimmerman, et al., (1976) Criminal Justice Research Assistance
    Project: Final Report.

Table 3b

EXEMPLARY PROJECT VALIDATION REPORTS INCLUDED

IN THIS STUDY

Abt Associates (1976) New Haven (Ct.) Case Incident Regional Reporting System (CIRRS) Exemplary Project Validation Report.

Abt Associates (1975) Neighborhood Assistance Officer Program, Dayton Police Department.

Abt Associates (1975) New York City Police Department Street Crime Unit-- An Exemplary Project.

Abt Associates (1975) Reserve Deputy Sheriff Program, Los Angeles County Sheriff's Department.

Abt Associates (n.d.) Exemplary Project Screening and Validation Report: Community Education on Law and Justice, Chicago, Illinois.

Abt Associates (n.d.) Exemplary Project Screening and Validation Report: Project Teletraining, County of Suffolk Police Department, New York.

Abt Associates (n.d.) Exemplary Project Screening and Validation Report: Southeastern Correctional Management Training Council, Athens, Georgia.

Abt Associates (n.d.) Legal Services for Prisoners, Inc. -- Topeka, Kansas.

Abt Associates (n.d.) Norfolk Fellowship Foundation, Inc: Massachusetts Correctional Institution, Norfolk.

Anon. (1975) Seattle Pre-Sentencing Counseling Program Exemplary Validation Report.

To illustrate the type of information available by which to select the original sample, Figures 1 and 2 present two typical NCJRS document summaries. Approximately 1500 such summaries were reviewed according to the criteria discussed above and graded subjectively on a scale of A to D. When distributional constraints were satisfied, an attempt was made to select the best evaluations (i.e., those given the highest grades). Thus, the sample is deliberately biased towards the better evaluations of criminal justice programs. The complete listing of evaluations in the sample is given in Appendix I.[2] Breakdowns of the sample by subtantive areas and evaluation methodology are given in Figures 3 and 4.

The necessary documents were researched and ordered by Dorothy Green, our project secretary, mainly during the months of April, May and early June, 1978; these are the months in which the original requests were sent out (we encountered considerable difficulty in obtaining some of the NCJRS-listed documents). Requests were based on information culled from the printout and other sources identified by project research assistants.

Checklist Review of Evaluations

During the summer of 1978 three project research assistants acted as "readers" of the evaluation reports, dividing the sample of 200 roughly evenly among them. Ed Kaplan, whose expertise is in mathematical modeling and the logistical area, focused almost exclusively on logistical studies. Vicki Bier, a doctoral student in operations research selected studies from both the logistical and social service area, with particular focus on studies containing statiscally-based methodologies. Cheryl Mattingly, a doctoral student in Urban Studies and Planning and a researcher of process evaluation methodology, read almost exclusively social service type evaluations.

FIGURE 1

NCJRS PRINT-OUT (DOCUMENT 174)

SET/2 DOCUMENTS 1:80(

VARIABLES SUCH AS THE CULTURAL SETTING OF THE PROJECT, HUMAN, FINANCIAL
AND MATERIAL RESOURCE AVAILABILITY, AND WHETHER THE PROJECT IS IN
KEEPING WITH ACCEPTED STANDARDS OF CORRECTIONAL PRACTICE. CORRECTIONAL
PROJECTS ARE DIVIDED INTO THREE CATEGORIES - YOUTH SERVICE BUREAUS,
INSTITUTIONAL CORRECTIONS, AND COMMUNITY CORRECTIONS. THE APPENDIXES
CONTAIN FORMS SUGGESTED FOR COLLECTING AND ANALYZING DATA REQUIRED FOR
JUSTIFICATION AND EVALUATION OF PROJECTS IN EACH CATEGORY. (SNI
ABSTRACT)

DOCUMENT 174

ACCESSION NUMBER:... C9900.00.012328
TITLE: EX-OFFENDERS AS PAROLE OFFICERS - AN EVALUATION OF
                    THE PAROLE OFFICER AIDE PROGRAM IN OHIO
                    PUBLICATION DATE: 73   PAGES: 182
AUTHOR(S):   SCOTT, J E   BENNETT, P A
CORPORATE AUTHOR: OHIO STATE UNIVERSITY
                    COLUMBUS OH 43210
GRANT(S):   2860-00-J3-72 (LEAA)
SPONSORING AGENCY: LAW ENFORCEMENT ASSISTANCE ADMIN
                    WASHINGTON DC 20531
SALES AGENCY: OHIO STATE UNIVERSITY PRESS
                    HITCHCOCK HALL, ROOM 316
                    2070 NEIL AVENUE
                    COLUMBUS OH 43210
ANNOTATION:
   AN EVALUATION OF OHIO'S USE OF EX-OFFENDERS AS PAROLE OFFICER AIDES.
ABSTRACT:
   THIRTEEN AIDES WERE ORIGINALLY HIRED TO PERFORM TASKS SIMILAR TO THOSE
OF A PAROLE OFFICER. THEIR PERFORMANCE IN COMPARISON TO A CONTROL GROUP
OF OHIO PAROLE OFFICERS WAS ASSESSED USING SEVERAL TECHNIQUES - AN
ATTITUDINAL QUESTIONNAIRE, IN-DEPTH INTERVIEWS, FIELD WORKER'S REPORTS,
UNIT SUPERVISORS RATINGS, A SURVEY OF INMATES, AND A SURVEY OF PAROLEES
SUPERVISED BY AIDES AND PAROLE OFFICERS. THE RESULTS FROM THE ANALYSIS
OF THE ATTITUDINAL QUESTIONNAIRE INDICATE AIDES POSSESS THOSE
ORIENTATIONS AND ATTITUDES ASSOCIATED WITH SUCCESSFUL SOCIAL SERVICE
WORKERS MORE OFTEN THAN DO PAROLE OFFICERS. THE IN-DEPTH INTERVIEWS
FOUND AIDES VERY SATISFIED WITH THEIR WORK. THE FIELD WORK REPORTS BY
STUDENTS INDICATED PAROLE OFFICERS SAW MORE PAROLEES PER DAY AND SPENT A
GREATER PERCENTAGE OF THEIR WORK TIME WITH PAROLEES THAN DID AIDES.
HOWEVER, FIELD WORKERS EVALUATED AIDES AS HAVING A BETTER RELATIONSHIP
WITH BOTH PAROLEES AND FELLOW WORKERS. UNIT PAROLE SUPERVISORS RATED
PAROLE OFFICERS AND AIDES ON SEVERAL DIMENSIONS, INDICATING THAT IN MOST
RESPECTS PAROLE OFFICERS WERE SUPERIOR TO AIDES. THE ONLY TWO CRITERIA
ON WHICH SUPERVISORS RATED AIDES SUPERIOR WERE GETTING JOBS FOR PAROLEES
AND BEING WILLING TO 'GO THE EXTRA MILE' TO HELP PAROLEES. INMATES AND

FIGURE 2

NCJRS PRINT-OUT (DOCUMENT 175)

SET/2 DOCUMENTS 1:800

PAROLEES RATED AIDES SUPERIOR TO PAROLE OFFICERS ON EVERY INDICATOR. THE
MAJORITY OF INMATES FELT PAROLEES SUPERVISED BY AN AIDE WOULD BE MORE
LIKELY TO SUCCEED ON PAROLE. AN OVERWHELMING MAJORITY OF INMATES (86
PERCENT) INDICATED THEY WOULD PREFER BEING SUPERVISED ON PAROLE BY AN
AIDE RATHER THAN A PAROLE OFFICER. (AUTHOR ABSTRACT MODIFIED) (SNI
ABSTRACT)

DOCUMENT 175

ACCESSION NUMBER:... 09900.00.012435
TITLE: POLICE-COMMUNITY ACTION - A PROGRAM FOR CHANGE IN
                      POLICE-COMMUNITY BEHAVIOR PATTERNS
                      PUBLICATION DATE: 73   PAGES: 217
AUTHOR(S):   EISENBERG, T   FOSEN, R H
SALES AGENCY: PRACTICING LAW INSTITUTE
                      1133 AVENUE OF THE AMERICAS
                      NEW YORK NY 10036
ANNOTATION:
  DISCUSSION AND EVALUATION OF PROJECT PACE, A TWO-YEAR PROGRAM INVOLVING
  THE OVERALL SCOPE OF POLICE-COMMUNITY RELATIONS WITHIN THE CITY OF SAN
  FRANCISCO.
ABSTRACT:
  THE GOALS OF PROJECT PACE (POLICE AND COMMUNITY ENTERPRISE) WERE
  TWO-FOLD - TO REDUCE ANTAGONISMS AND POLARIZATION WHICH HAD DEVELOPED
  BETWEEN THE POLICE AND THE SAN FRANCISCO COMMUNITY, AND TO CREATE A MORE
  COOPERATIVE ATMOSPHERE FOR CRIME PREVENTION AND CONTROL AND THUS REDUCE
  VARIOUS KINDS OF CRIME. THE INITIAL FOCUS OF THE PROGRAM, WHICH WENT
  INTO OPERATION IN APRIL 1969, WAS ON RELATIONS BETWEEN POLICE AND
  MINORITIES, ESPECIALLY BLACKS. THE FOUR PHASES COMPRISING THE PROGRAM
  INCLUDED ATTITUDE SURVEY AND CURRICULA DEVELOPMENT, POLICE AND RESIDENT
  DISCUSSION SESSIONS, ACTION PROGRAMMING, AND PROGRAM IMPLEMENTATION.
  THIS VOLUME IS ORGANIZED INTO TWO PARTS, THE FIRST OF WHICH PRESENTS
  CHAPTERS DESCRIBING THE PROJECT'S EARLY DESIGN AND DEVELOPMENTAL EFFORTS
  AS WELL AS THE APPROACHES EMPLOYED AND THE RESULTS OBTAINED IN EACH OF
  THE FOUR PROGRAM PHASES. IN PART TWO A NUMBER OF POLICE-COMMUNITY
  RELATIONS PROGRAMS ARE DESCRIBED. THE FINAL CHAPTER DISCUSSES SOME
  PRACTICAL ISSUES AND IMPLICATIONS FOR THOSE CONSIDERING IMPLEMENTING OR
  PRESENTLY ENGAGED IN SUCH COMMUNITY RELATIONS PROGRAMS. THE APPENDIX
  CONTAINS INTERVIEW SCHEDULES, PROJECT PROPOSALS, POLICE AND RESIDENT
  ATTITUDE COMPARISONS, AND EVALUATION FORMS.

DOCUMENT 176

ACCESSION NUMBER:... 09900.00.012441

FIGURE 3

SAMPLE COMPOSITION: BY SUBSTANTIVE AREA
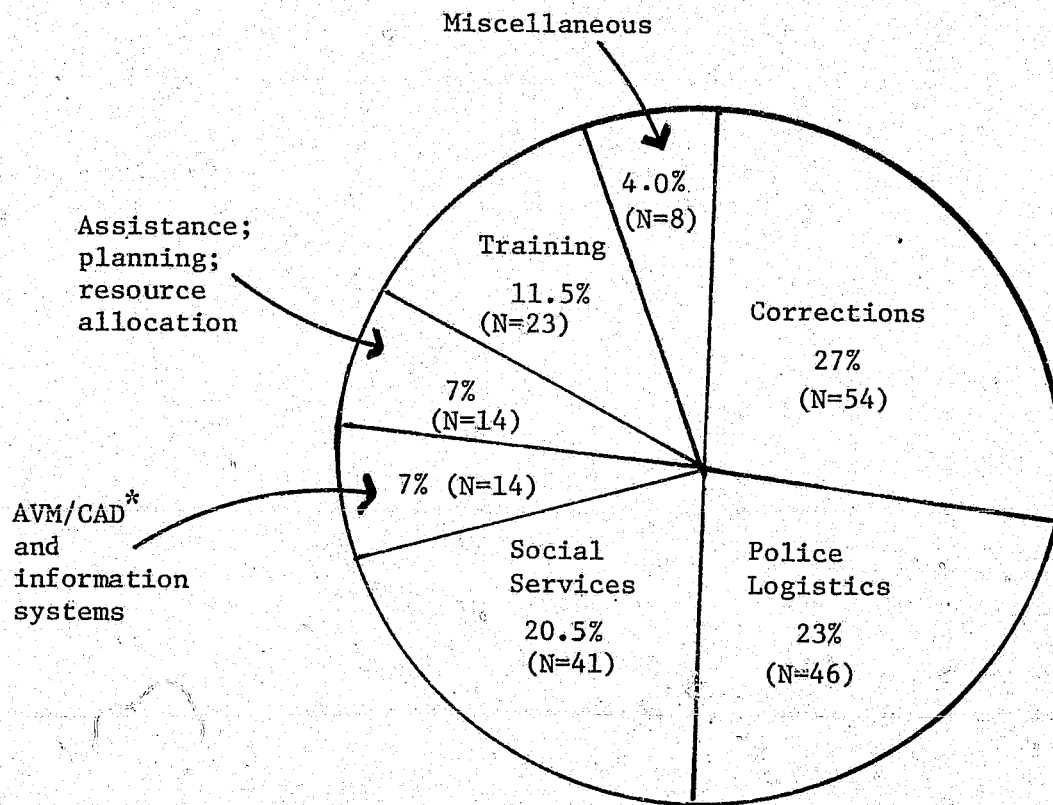
* AVM = Automated Vehicle Monitoring
  CAD = Computer Aided Dispatch

FIGURE 4

SAMPLE COMPOSITION: BY PRIMARY EVALUATION METHODOLOGY

General
outcome
evaluation

15%
(N=29)

Quasi-
experimental
design

21%
(N=42)

Comprehensive
evaluation

15%
(N=25)

(N=17)

Narrative
case study

20%
(N=40)

Experimental
design

8.5%

7.5%

input
evaluation
or audit
(N=15)

7%

4%

Formal models, .5%
(N=1)

Process
evaluation with
performance measures
(N=14)

Statistical
Models, 1.5%   (N=3)

Process evaluation
without performance
measures
(N=8)
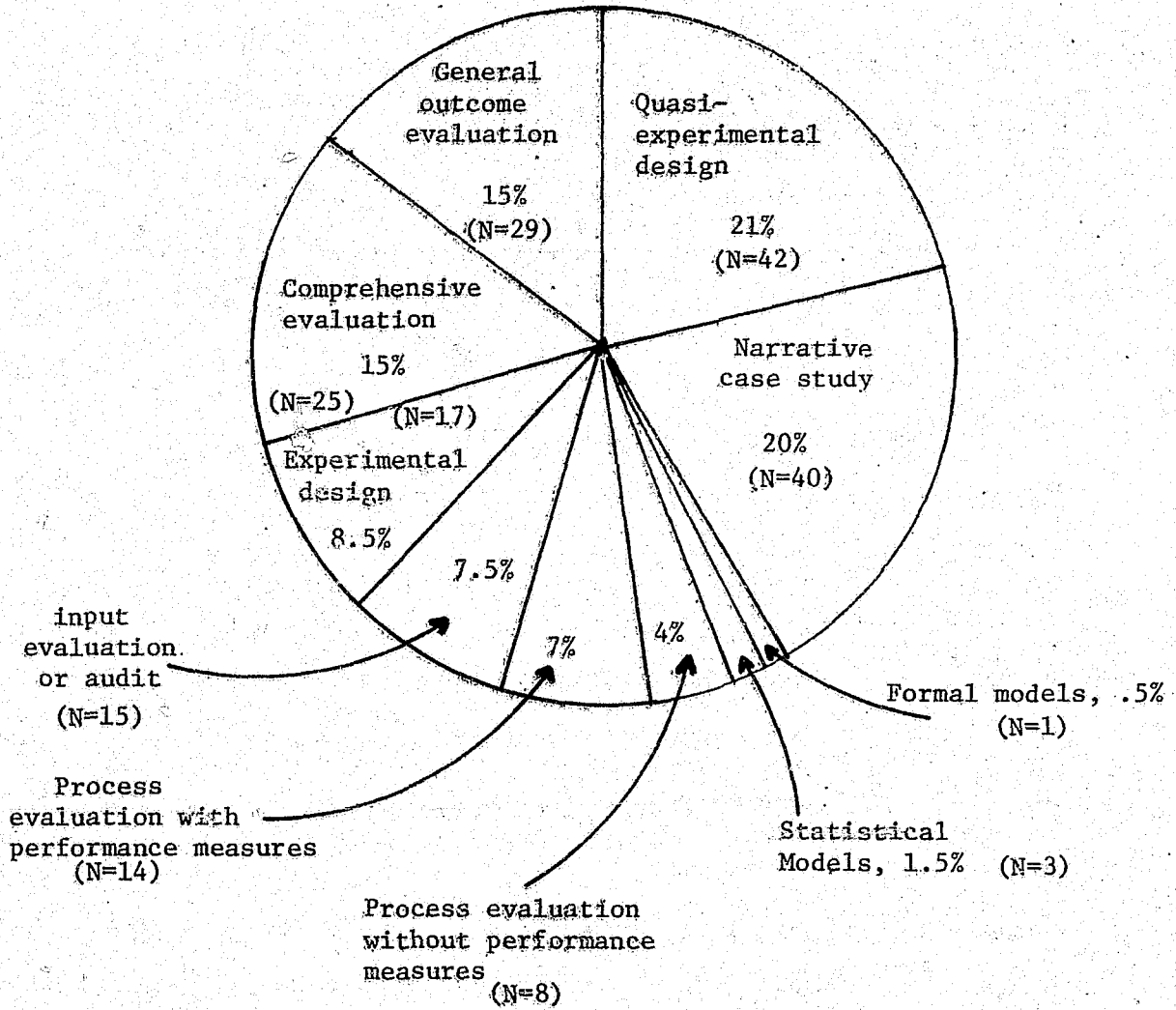
For each study the reader completed a checklist having 31 entries
(see Appendix II). Each study was assigned an identification (ID) number
to aid in the analysis after the checklist had been completed. The ID
number included *inter alia*, a general subject category and citation of
the agency conducting the evaluation (see Appendix III). The complete
ID number is also explained in Appendix III.

Each reader completed nearly two checklists per day, averaging
roughly *four hours* per evaluation report. While this amount of time is
small to evaluate any evaluation satisfactorily, it is probably more time
than most decision makers would spend on the report. As can be seen from
the checklist (Appendix II), the goal of the readings was not to evaluate
each evaluation, but rather to obtain information regarding evaluation
input, process, and outcome, and to assess in a general way the relevance
of the methodology employed, and the quality of the documentation.
[Separate project reports evaluate specific evaluations in detail.][3]

The completed checklists were read into an interactive, natural
language computer system. An illustrative checklist set of answers is
provided in Appendix IV. The interactive computer program allowed adap-
tive searching through answers, and permitted additions and modifications
in initial answers.

At this point, the research team was expanded by the presence of
Nancy Reichman, a Ph.D. student in Urban Studies and Planning, and Timothy
Eckels, a MCP student in Urban Studies and Planning. Nancy and Tim carried
part of the load of analyzing the completed checklists, and brought a
valuable sociological perspective to the research effort.

Recognizing that three readers each read roughly one-third of the

sample, and that many responses on the checklist are in large part sub-
jective in nature, an attempt was made through project meetings to pro-
vide feedback among readers to assure consistency of response. Still,
it is likely that the different readers responded somewhat differently
to various types of questions, depending often on their own areas of
primary interest. (It is not possible to test for consistency statisti-
cally since each reader was assigned a distinctly different subsample.)
However, the summary findings and conclusions of the checklist analysis
– contained in this report – do not depend on fine gradations among the
readers. An overall pattern emerges which is largely independent of
minor responder-specific variations in response style.

Each of the five researchers examined a group of questions in detail
and then drafted analytical reports on their results. Leni Berliner,
technical assistant for the project, reviewed these drafts and made re-
commendations for revision and clarification. She then took these re-
vised versions and assembled this report.

Outline of Report

This report summarizes our analysis of the 200 checklist responses,
augmented by re-examination of the original evaluation documents where-
ever necessary. Reflecting the importance of inputs, process, and out-
come, we have organized the report according to these elements of an
evaluation.

In addition to obtaining descriptive information regarding inputs,
process, and outcome, we are concerned with particular issues which
could affect the potential impact of the evaluations on decision-making.
These range from simple matters such as <u>accessibility</u> and <u>completeness</u>

of the final report, to more difficult ones, including appropriateness and application of methodology, degree of comprehensiveness, and flexibility of design and its implementation. In the conduct of the evaluation, we are concerned with amount and type of communication between program staff and evaluators and demonstrated awareness of institutional and program environmental issues affecting the program and the evaluation. The manner in which each of these issues is addressed by the evaluators affects the quality of the evaluations, hence the quality of the information contained therein, hence the quality of any subsequent decisions based (at least in part) on the evaluation.

INPUTS

As in a program, before one can analyze the process of an evaluation one must first obtain knowledge of the inputs to that evaluation. Inputs that have been found to be important include the level and type of funding, staffing patterns, timing, etc. These inputs affect the design, implementation, and outcome or quality not only of a program, but also of the evaluation of that program. For the purpose of this report, we need to know about inputs to the evaluations so that we may assess them more fairly.

Only eight (8) evaluations in our sample indicated the percentage of the program budget alloted to evaluation: this ranged from one to twelve percent. Only four evaluations reported the total funding for the evaluation. This response prompts us to ask if budgetary information should be included in the report itself. Although there is little comment on this issue in the evaluation literature, we note that evaluations are commissioned for a specific purpose. Evaluation (unlike general social science research) is conducted as a service: evaluators should be accountable

for their product. One should be able to assess the value of an evaluation given known budgetary constraints. It would be unreasonable to expect an evaluation conducted under a $3000 budget to be comparable in detail to an evaluation conducted with a $40,000 budget. Budgetary information would also provide a useful indicator or model for someone who is planning or undertaking a similar project.

The time spent on evaluation is also an important input to consider in the evaluation of evaluations. Relatively few evaluations (31 percent of the total sample) documented the time spent on the evaluative effort. The average length of time spent on evaluation was 11 months with a minimum of 1 month and a maximum of 42 months. About half those reported were conducted in less than 6 months. There were no indications that the length of time for evaluation varied significantly among agencies undertaking the evaluation. This is somewhat surprising as one would expect to find longer evaluations from established research groups such as the Police Foundation or the Urban Institute. We suspect, although we did not specifically test for it, that there may be some variation by type of evaluation study.

It is impossible to determine whether the sample over-represents evaluations of short duration. However, one could conjecture that those evaluators constrained by a short evaluation period would be more likely to highlight the time allocated to evaluation as a means of rationalizing inconclusive results. Knowing the time frame puts the evaluation in context, which enhances its overall utility.

The researchers tried to determine from the evaluation reports themselves when the evaluations were planned. The results were as follows:

TABLE 4

PLANNING OF THE EVALUATION

When was the evaluation planned?	N	%
Before Program Implementation	65	37
During Program Operations	90	52
After Program Termination	19	11
TOTAL	174	100
(missing or not applicable = 26)		

Examination of the planning of evaluation by agency responsible for it revealed that evaluations undertaken by State Planning Agencies were somewhat more frequently planned prior to program implementation than evaluations conducted by LEAA at large. This may be a direct consequence of the function of State Planning Agencies, which have clearly defined evaluation responsibilities. The evaluations in our sample conducted by the Police Foundation were planned prior to program implementation, reflecting the demonstration nature of projects undertaken by it.

It is not always clear which strategy is the best to employ. The evaluation literature suggests that evaluations should be, at the very least, in the minds of the program developers. There is, however, some danger in completely designing an evaluation before the program has had a chance to settle down. This factor proved to be a problem in several of the evaluations in our sample where program instability was cited as a reason for inconclusive results. Flexibility of evaluation design is of great importance in such instances.

The timing of the evaluation is both critical to the measures of outcome and to the utilization of evaluation findings. From reading the evaluation reports, there appears to be some confusion as to whether an evaluation was conducted during program operations or after the program had been completed. An evaluation that is conducted after program termination obviously will not provide decision-makers with timely results on program modification, though at best they may be useful for the design of future programs. On the other hand, evaluations that are conducted during program implementation run the risk of inconclusive results due to program instability, particularly if the evaluation design is inflexible. Several "evaluations" were conducted prior to program implementation. For the most part, these were small pilot studies which preceded a larger experimental program, or they were "evaluability assessments" or simulations.

TABLE 5

PLANNING OF THE EVALUATION BY AGENCY CATEGORY

Planning of the Evaluation in relation to Program Implementation by Agency Category.									
Time / Agency	Before N	(%)	During N	(%)	After N	(%)	sub-total	not applicable	Total
State Planning Agencies	14	(41)	16	(47)	4	(12)	34	5	39
LEAA	6	(20)	22	(73)	2	(7)	30	1	31
Abt Associates	0	(0)	17	(94)	1	(6)	18	1	19
Police Foundation	8	(67)	2	(17)	2	(17)	12	1	13
Miscellaneous	37	(46)	33	(41)	10	(13)	80	4	84
SUB-TOTAL	65	(37)	90	(52)	19	(11)	174	12	186

Missing = 14

TABLE 6

TIMING OF THE EVALUATION EFFORT
IN RELATION TO PROGRAM OPERATIONS

Timing:	N	%
Before implementation	9	5
During operations	118	66
After termination	53	29
TOTAL	180	100
(missing or not applicable = 20)		

The vast majority of the 200 evaluations in our sample were conducted by people who were not associated with the program itself: only 22 percent of the evaluations were conducted by in-house personnel. This was slightly lower than expected, and may be an artifact resulting from our sample. Although for the most part the evaluators appeared to have some sort of social science training, generally it was difficult to determine the background characteristics of the personnel involved. It was also difficult to determine the evaluators' professional affiliation.

Most of the evaluation reports were authored by three persons or less. This figure is most likely deceiving. The number of personnel involved in the evaluative effort often cannot be determined from reading the report. Normally one is informed only of the principal investigators and the most senior program officials. We may pursue this question further in the author and consumer questionnaires.

Knowing who conducted the evaluation (by affiliation and training) is useful in that it may help indicate the possible biases of and constraints on the evaluators and their relations with program staff. This

information, along with budgetary and other input information, helps the reader make a more informed judgement about the evaluation, and hence about the program.

The initial evaluation design is the input factor which most directly influences the conduct or process of the evaluation. Aspects of design include a statement of the end purpose of the project, and an arrangement of elements to be manipulated and procedures to follow to reach that end. We are interested in:

- whether the evaluation research was designed to yield information that would be useful in a broader context than that of the particular program,

- whether it was comprehensive or not (in those situations where it should have been) and

- whether goals and aims of the program were clearly understood and specified by the evaluators, enabling them to develop appropriate performance measures.

Examination of the criminal justice evaluations in our sample to which this question was applicable revealed that slightly over half of them were designed to yield information that would be useful in a context broader than that of the specific program under study.

TABLE 7

RESEARCH CONTEXT

Was the research designed to yield information that would be useful in a broader context than just evaluating this particular program?		
RESPONSE	N	%
YES	100	53
NO	87	47
TOTAL	187	100
(missing or not applicable = 13)		

In general the evaluations that were broad in scope seemed to be set up as demonstration projects or so called "state-of-the-art" documents for a particular field. Given the different goals, objectives, resources and perceptions of the evaluative purpose, we expected to see some differences in scope among the agencies conducting the evaluations. As one might expect, evaluations undertaken and designed by the larger research organizations (eg., the Police Foundation, the Urban Institute, etc.,) were more likely to have a broad focus than evaluations conducted by LEAA and the State Planning Agencies. With the exception of evaluations conducted on team policing programs and on technical innovations, there was remarkable consistency among subject areas in terms of the intent of the research designs. This suggests that goals, resources and perceptions of purpose vary more with the agency undertaking the evaluation than with the subject area, confirming the importance of examining the institutional seeting of evaluation (i.e., the organizational variables associated with the agency undertaking the study).

Comprehensiveness

To be considered comprehensive, a given evaluation should focus on
as many aspects or stages of the subject program as possible, from de-
sign through implementation to effect on the client population. An eval-
uation which focuses on any one of these aspects in isolation is often
less useful for either internal (program) or external (policy) purposes,
than one which is more comprehensive (although special purpose evaluations
can clearly be highly useful if carried out during the relevant stage of
the program). Obviously, financial and political constraints may restrict
the potential coverage to be provided by the evaluations. For most forms
of outcome evaluation, the evaluator will need a clear description of the
objectives that the program is pursuing: the statement of goals is the
basis for the development of several other crucial components of the eval-
uation process. These include the designation of intermediate or proxy
objectives (if necessary), the construction of a hypothesis about how
the goals (are to be) reached, the determination of performance measures
to be used, and the final conclusion about whether the program "worked".
A comprehensive evaluation may be handicapped from the very beginning
without some clear understanding of program goals. The registration of
objectives and the measurement of their attainment are rarely simple pro-
cedures, however. Established goals may be absent at the commencement
of a program, or there may be substantial disagreement about what con-
stitutes the "official" or "proper" goals. What is crucial in any case
is that adequate communication and negotiation take place among the
various actors in the evaluation process. Only then will the evaluation
be genuinely useful to decision makers or program managers. Thus, eval-
uators should at least examine the degree to which program goals were or

were not clearly focused, and should report this in their final report

for the benefit of the readers. Two questions were used in an attempt

to investigate this issue:

> 14a. Did the evaluators consider whether program goals
> were clearly specified?

> 14b. Do you feel that they were clearly specified?

The second item was a preliminary attempt to determine just how of-

ten programs did designate their goals. We will pursue this item further

as we communicate with the authors of the evaluations. Despite the fact

that we were attempting to look "through" the reports to the character

of program aims (hence the high incidence of "not applicable" or "can't

tell" as responses), disagreement over goals or lack of a clearly con-

structed set of goals was often apparent.

As shown in Tables 8 and 9 the majority of programs did appear to

have clearly stated goals, and most reports did include discussion or

presentation of them (based on rather lenient responses to the questions).

Reader comments on both questions reinforced the generally favorable trends, eg:

> "goals were clearly stated" (police-community interaction)
> "goals were specified operationally" (court scheduling)
> "goals were repeated in the evaluation and there doesn't
> seem to be any problem understanding them" (pre-trial release)
> "Spelled them out in great detail" (corrections)

Breaking down the results by topic area, however, we see that these

favorable tendencies occurred most often in the "police logistics" studies.

Given that these programs are usually more operational and focused then

programs in areas such as training or corrections, this result is not

surprising. We are primarily concerned, however, with the non-logistical

program areas where goals are often not displayed or considered by eval-

uators. This seems to be particularly the case in corrections program,

TABLE 8

GOAL SPECIFICATION I

Did the evaluators consider whether the goals were clearly specified?						
Response Program	YES N	(%)	NO N	(%)	sub- total	not applicable	Total
assistance/plan- ning/resource allocation	10	(77)	3	(23)	13	1	14
training	7	(37)	12	(63)	19	0	19
corrections	22	(46)	26	(54)	48	5	53
police- community relations	9	(53)	8	(47)	17	2	19
pre-trial release	8	(100)	0	(0)	8	0	8
courts & jury management	3	(27)	8	(73)	11	0	11
police logistics	42	(98)	1	(2)	43	2	45
AVM/CAD & information systems	12	(86)	2	(14)	14	0	14
miscellaneous	4	(80)	1	(20)	5	2	7
TOTALS	118	(65)	64	(35)	182	12	194

missing = 6

TABLE 9

GOAL SPECIFICATION II

Do you feel that the program goals were clearly specified?							
Response Program	YES		NO		N/A*		Total
	N	(%)	N	(%)	N	(%)	
assistance/plan- ning/resource allocation	4	(28)	6	(43)	4	(28)	14
training	7	(37)	8	(42)	4	(21)	19
corrections	22	(42)	15	(28)	16	(30)	53
police- community relations	6	(32)	9	(47)	4	(21)	19
pre-trial release	8	(100)	0	(0)	0	(0)	8
courts & jury management	10	(91)	1	(9)	0	(0)	11
police logistics	36	(80)	6	(14)	3	(6)	45
AVM/CAD & information systems	11	(79)	1	(7)	2	(14)	14
miscellaneous	3	(43)	2	(28)	2	(28)	7
TOTALS	106	(55)	52	(27)	36	(18)	194

*N/A = not applicable

Missing = 6

in technical assistance programs, training units, and in the Police-Community

interaction area.

A rather striking result appears in the pre-trial programs, where

all eight programs specified goals succintly. It should be added that

these programs were located in a variety of areas and that the evaluators

were sponsored by a number of different agencies. Considering that these

programs are often similar in structure to other correctional endeavors,

it may be that this was a more or less chance occurrence. On the other

hand, it may be that because pre-trial programs are more vulnerable to

criticism due to the "presumed guilt" controversy, they tend to be

more definitive about their purpose. Actors in the criminal justice sy-

stem may be less tolerant of a nebulously described operation that

is pre-trial than one which deals with the convicted offender.

When goals were not "clearly specified", this seemed to be due largely

to failure on the part of evaluators and/or program administrators to trans-

late broad goals into intermediate or sub-objectives. This may not always be

possible, but it is certainly desirable in cases where an outcome assessment is

to be applied. Some sample comments from readers:

> "refers to things like increasing professionalism without
> spelling out what that would mean" (police-community
> interaction)

> "does not clarify what 'effectiveness' means" (training)

> "goals very large, no discussion of intermediate goals,
> performance measures, or the like" (training)

> "vague terms such as 'substantial reduction', 'increase
> the clearance rate', etc." (police-logistics)

The vagueness of goals may be a result of intentional obfuscation on the

part of the evaluators. They may deliberately underspecify program

objectives in order to ensure positive results, channel the reader's attention to successful program components, and/or appease program administrators. Unfortunately, motives such as these will be difficult to ferret out, even with further investigation. For many of the poorly executed evaluations, however, they are real possibilities. Readers' comments in other portions of the report indicate strongly that a number of evaluations were no more than public relations efforts or formal legitimization of programs.

Researchers in evaluation have often referred to political or environmental influences on program goals. They point out that vague aims may be a function of the need to pass a program through the legislative process. Perhaps more germane to our study is the possibility that program administrators underspecify program goals in order to insulate themselves from unequivocal judgement of program failure. This tension between program staff, who resist operational goals, and evaluators who seek specific measures, is a recurrent problem in the evaluation field. Campbell[5] suggests that an entirely different attitude toward social programs and evaluation is needed, one where negative results from an outcome evaluation are used as steps in an improvement process rather than as the deathblow for a whole program, particularly since some social interventions may not actually have expressible goals at their outset.

Another possibility is that program purposes may be overt, but not amenable to measurement or reduction into sub-objectives. In a study of a police-community interaction program, for example, the evaluators noted that the purposes were discussed,

> "but could not be easily put into measureable objectives...
> Basically goals were that the officer would become 'more
> professional' in dealing with emotionally charged situa-
> tions, 'better able to function in an urban environment,' etc."

Readers made a limited number of similar remarks about training programs:

> "difficulty in making goals concise for a large program"
> "evaluator felt that goals were not clearly specified
> because they could not be put into measurable objectives"

Goal ambiguity may not be a sign of an inadequate program so much as an indication that a different evaluation focus is called for. In the situations described above, a greater emphasis on qualitative process evaluation would have been advisable. A process evaluation need not be goal-oriented in the same sense as an outcome assessment; the appraisors might concentrate on describing and analyzing the program activities and staff interactions, rather than attempt to develop a performance measure for an end-state of being "more professional".

Despite the desirability of clarifying objectives for an outcome evaluation, there are potentially a number of risks involved. One of these is that researchers may break down large goals into measurable objectives that inadequately represent the progress of the program. This "reductionism" issue is a recurring source of tension between evaluators and may be prominent in more cases than we can ascertain by reading the reports. A few reader comments hint at the problem:

> "....(goals were) very specific, even quantifiable...
> if anything maybe too narrow and outcome oriented."
> (drug program)

Reductionism may lead to futher problems in the evaluation design. In one program aimed at reducing police brutality, a sub-goal of reducing certain types of arrests (interfering with an officer) was used for measurement purposes. This practice raised questions about the validity of the evaluation and program outcome.

Related to reductionism is the danger that evaluators may impose certain goals in a program which have little to do with the actual staff effort. Deutscher[4] illustrates how it often happens that evaluators assume that the "official" goals as stated in the funding proposal, or in political rhetoric, are the appropriate guides for assessment. Meanwhile, the program staff may have been operating under a different set of assumption and goals. While it is not clear what an evaluator should do in such a dilemma, the problem was undoubtedly a real one in some of our sample cases. Sometimes there was discussion of what a program could not be expected to do. One evaluator of a corrections program carefully stated that he "did not propose that this or other small-scale demonstration in a field with massively ingrained problems could significantly affect recidivism rates". Thus it is important that the suitability of program goals be considered, particularly when selecting performance measures. We may address this issue (of whether or not evaluators actually assessed the goals of the program in terms other than clarity) further via the author and consumer questionnaires.

Several writers have stressed the necessity of pre-evaluation negotiation among staff, sponsor, and evaluator. Many of the tensions discussed here could have been avoided or mitigated if there had been more communication before, during, and after assessment activity.

Target Population

If an evaluation is to assess the value of a program, it is reasonable to expect attention to be given to the issue of target population. Slightly more than one-third of the evaluation in our sample discussed the issue of target population. Another third did not seem to consider the issue; in the remaining cases the response of "not applicable" generally meant that there was not enough information given in the report

for an analysis to be made, or that the program was of a different nature.

Interestingly, in our sample, the evaluators focused on discovering the population that the program was best suited to serve, as much as evaluating the program in terms of how well the intended population was being served. This may be because the target population was often not initially very clearly defined.

With the exception of pre-trial release and AVM/CAD* systems, there was little variation among substantive areas in the percentage of evaluations which considered whether the program was directed at the appropriate target population. The range was between 40 percent and 79 percent, with the average score of 53 percent. Not surprisingly, 88 percent of the studies of the pre-trial release evaluations considered this issue. The issue of who should be diverted or receive alternate treatment to incarceration was generally of central importance and some of the evaluations in this area were explicitly designed to focus on this issue.

The question of whether the program was directed at the appropriate target population was answered "not applicable" in nearly one third of the evaluations in our sample. The primary reason for this was that not enough information was provided in the evaluation to be able to answer that question. Even in cases where the evaluators made assessments, the studies generally did not provide enough information for a reader to make such an assessment. It is important to ask why so few studies explicitly considered the issue of whether the program was serving the client population it was designed to. Certainly in our sample this appeared to be

---

*AVM stands for Automated Vehicle Monitoring; CAD stands for Computer Aided Dispatch.

TABLE 10

CONSIDERATION OF TARGET POPULATION I

Response / Program	YES N	(%)	NO N	(%)	sub-total	not applicable	Total
assistance/planning/resource allocation	4	(40)	6	(60)	10	4	14
training	8	(47)	9	(53)	17	2	19
corrections	23	(47)	26	(53)	49	4	53
police-community relations	8	(50)	8	(50)	16	3	19
pre-trial release	7	(88)	1	(12)	8	0	8
courts & jury management	4	(67)	2	(33)	6	5	11
police logistics	15	(79)	4	(21)	19	26	45
AVM/CAD & information systems	1	(50)	1	(50)	2	12	14
miscellaneous	2	(40)	3	(60)	5	2	7
TOTALS	71	(52)	65	(48)	136	58	194

Missing = 6

TABLE 11

CONSIDERATION OF TARGET POPULATION II

	Do you feel that it <u>was</u> directed at the appropriate target population?						
Program \ Response	YES N	(%)	NO N	(%)	N/A* N (%)	Total	
assistance/plan- ning/resource allocation	2	(14)	1	(7)	11 (79)	14	
training	4	(21)	1	(5)	14 (74)	19	
corrections	14	(26)	9	(17)	30 (57)	53	
police- community relations	8	(42)	8	(42)	3 (16)	19	
pre-trial release	1	(12)	5	(62)	2 (25)	8	
courts & jury management	4	(36)	2	(18)	5 (45)	11	
police logistics	14	(31)	0	(0)	31 (69)	45	
AVM/CAD & information systems	3	(21)	0	(0)	11 (79)	14	
miscellaneous	0	(0)	0	(0)	7 (100)	7	
TOTALS	50	(26)	28	(14)	116 (60)	194	

*N/A = not applicable

Missing = 6

problematic, at least in the non-logistical programs.

PROCESS

Reviewing the process of an evaluation is valuable in that knowledge of the evaluation process allows us to assess the legitimacy of the results and conclusions of the evaluation. Naturally this is important for making decisions about the future of the program. Our use of the term evaluation "process" refers to the actual conduct or execution of the evaluation. Here methodology and methods are very important. Another consideration is how well the evaluators related their actual experience in conducting the evaluation to the original evaluation design or plan. In this regard we examine the evaluators' knowledge of program activities (as reflected in the evaluation reports), the degree and kind of communication between the evaluators and program staff and their flexibility (as indicated by their willingness and/or ability to adapt their design and procedures). Finally, the evaluators' knowledge and flexibility are directly related to their possessing a theory linking program activities to desired outcomes: this is of utmost importance in determining the validity and usefulness of the evaluation.

Implementation

The issue of how a program was implemented, and whether this implementation corresponded to the original plan of the program, is obviously an important one for evaluation. There is always a danger in assuming that the intended program is the same as the program in operation and then attributing the outcomes to the intended program, when in fact the two may differ quite markedly. A little less than half of the evaluations in our sample included some discussion of program implementation, or at least indirectly alluded to the fact that the

evaluators had taken a critical look before assuming that the intended

program was in operation. However, slightly more evaluations gave no

discussion of this at all, seeming to implicitly assume that the program

had been implemented as designed.

In some evaluations, where evaluators played an active role in the

program, they tried to ensure that the program did conform to the model

by carefully monitoring the implementation process. Others used partici-

pant-observation and interviewing techniques to see how staff spent their

time or how they perceived their jobs, and how this matched the original

design. A few process evaluations or comprehensive evaluations which

included a process component did historical studies, following the pro-

gram from its inception through the changes which came about during its

operation. Those evaluations which were careful to monitor and document

changes in the program were very much in the minority however. Often,

discussion of implementation was indirect or quite superficial.

There was some difference in response according to substantive area.

Evaluations which seemed particularly negligent about including discussion

of program implementation occurred in training and juvenile diversion pro-

grams. The substantive areas which fared better were courts and jury manage-

ment, AVM/CAD and information systems, and police logistics. (see Table 12).

Because discussion of program implementation tends to be inherently

negative, pointing up discrepancies between what actually occurs and

what is ideally supposed to occur, it would be interesting to compare

in-house evaluations with evaluations done by outside researchers, to

discover whether being involved in the program affected how detailed and

candid the discussion was. While on the one hand, in-house evaluators

may have a more difficult time presenting implementation problems, their

TABLE 12

DESCRIPTION OF PROGRAM ACTIVITIES

Are program activities clearly described?							
Response ⧸ Program	YES		NO		Sub-total	not applicable	Total
	N	(%)	N	(%)			
assistance/planning/resource allocation	7	(54)	6	(46)	13	1	14
training	3	(16)	15	(84)	18	1	19
corrections	23	(43)	30	(57)	53	1	54
police-community relations	13	(68)	6	(32)	19	0	19
pre-trial release	5	(63)	3	(37)	8	0	8
courts & jury management	10	(91)	1	(9)	11	0	11
police logistics	32	(73)	12	(27)	44	1	45
AVM/CAD & information systems	12	(86)	2	(14)	14	0	14
methodology	3	(100)	0	(0)	3	0	3
miscellaneous	4	(57)	3	(43)	7	0	7
TOTALS	112		78		190	3	193

missing = 7

extra contact with the program may help considerably.

A crosstabulation of responses to questions 16a and 16b (Did the evaluators consider whether the program was implemented as designed? Do you think it was?) reveals that the lack of consideration of program implementation is a more immediately pressing, though not more serious, problem than poorly implemented programs. Until evaluators do begin to consider the issue systematically, the conclusions of many evaluations will be of questionable value, since there will be little evidence that the program was properly implemented in the first place.

A clear and candid description of program activities is important in an evaluation for two reasons. First, it is important for the audience for whom the evaluation is intended to have this information, particularly if the program is under consideration for replication or if similar programs are being conducted. The evaluation report will be especially useful to its audience if it includes a detailed description of the program process and particularly a description of those aspects of the program which were faulty as well as those which were successful. Second, it is important as an indicator of the evaluators' knowledge of the actual workings of the program, and as an indicator of the extent to which process and input issues were considered by the evaluators.

In our sample slightly more than half of the evaluation reports included descriptions of program activities (See Table 12). However, this question was answered loosely, so that often a positive response was given, even though the description may not have been detailed or coherent enough to give the reader a full picture of the program. This issue is addressed further in the survey of consumers of the evaluations.

Some evaluations did an excellent job in presenting program activities. A few gave historical accounts of the projects' unfolding, including weaknesses and pitfalls as well as successes. The best descriptions seemed either to include a detailed outline of procedures for the implementation of the program or to go beyond a generalized or idealized version and supplement the picture of the total program with specific examples of program activities.

In contrast, many program descriptions were very brief. This was the most frequently cited problem. Another frequent problem was that evaluators gave an overall impression, a theoretical discussion or a very general picture without going any further. Descriptions often tended to be static, and did not consider program process.

At times evaluations went into considerable depth in discussing some aspects of the program while barely mentioning the more critical features. For example:

--"The operations of the actual subsidy program are clearly specified, but not the local correctional innovations it encouraged. These innovations are much more directly related to reductions in commitments than the mere existence of the subsidy program." (corrections)

--"While the course lectures were presented in the evaluation, no discussion of how workshop trainees and lecturers were recruited and organized, of what trying to accomplish in any specific way or of interaction among participants." (training)

Another problem even with the better descriptions is that often they had an idealized, public relations flavor, or "flowery and grandiose description." There was a sense that the actual program could not possibly fit the description given of it -- it was made to sound too smooth and neatly put together. This was particularly true for those evaluations which were intended to serve as models, and to appeal to a wide audience of decision makers who might be interested in implementing such a program.

In such cases there seemed to be fairly detailed description but minimal discussion of problems and failures.

As you may recall from the previous discussion in the section on inputs, over half of the evaluations in our sample were designed to yield information that would be useful in a context beyond that of the particular program being evaluated. Tables 13 and 14 give the breakdown by agency and by subject.

Methodology and Methods

There are two problems to be dealt with when assessing the methods used in an evaluation. The first of these concerns the appropriateness of the methods used in view of the purposes of the study and the type of program (how well the methods were chosen). A complementary concern is how well these methods were applied.

It seems best to begin a discussion of methodology and methods by briefly classifying the studies in our sample. It turned out that the most popular evaluation type was that of experimental/quasi-experimental design. There were also a large number of narrative case studies. Somewhat fewer studies opted for general outcome evaluations with no definite structure, or for a more comprehensive approach analyzing input, process and outcome components. (See Figure 5)

Evaluations of social service programs tended to be narrative case studies or input evaluations. Logistical studies were more dependent on experimental/quasi-experimental design. (See Table 15) For certain areas of inquiry, or program types, the methodology was better chosen than for others. Those studies which scored highest on "suitability of focus" ratings were innovative methods (followed by "miscellaneous"), court-

TABLE 13

AIMS OF RESEARCH, BY AGENCY

Was the research designed to yield information that would be useful in a broader context than just evaluating this particular program? (by agency)							
Response Agency	YES		NO		Sub- total	not applicable	Total
	N	(%)	N	(%)			
State Planning Agencies	14	(39)	22	(61)	36	4	40
LEAA	15	(47)	17	(53)	32	1	33
Abt Associates	8	(42)	11	(58	19	0	19
Nat'l Council on Crime & Delinquency	2	(67)	1	(33)	3	0	3
Urban Institute	2	(100)	0	(0)	2	0	2
Rand Corp./ NYC Rand	0	(0)	3	(100)	3	0	3
City Police Departments	2	(67)	1	(33)	3	0	3
MITRE	5	(72)	2	(28)	7	0	7
Public Systems Evaluation, Inc.	3	(100)	0	(0)	3	0	3
Miscellaneous	35	(54)	30	(46)	65	3	68
TOTAL	100	(53)	87	(47)	187	8	195

Missing = 5

TABLE 14

AIMS OF RESEARCH, BY SUBJECT CATEGORY

Was the research designed to yield information that would be useful in a broader context? (by subject category)							
Response Program	YES		NO		sub-total	not applicable	Total
	N	(%)	N	(%)			
assistance/plan-ning/resource allocation	6	(46)	7	(54)	13	1	14
training	6	(32)	13	(68)	19	0	19
corrections	28	(57)	21	(43)	49	4	53
police-community relations	11	(57)	8	(43)	19	0	19
pre-trial release	4	(57)	3	(43)	7	1	8
courts & jury management	6	(54)	5	(46)	11	0	11
police logistics	26	(62)	16	(38)	42	2	44
AVM/CAD & information systems	7	(50)	7	(50)	14	0	14
innovative methodology	2	(33)	4	(67)	6	0	6
miscellaneous	4	(57)	3	(43)	7	0	7
TOTAL	100	(53)	87	(47)	187	8	195

missing = 5

FIGURE 5

FREQUENCY OF STUDY FOCI


```
        5    10   15   20   25   30   35   40   45   50
```

Quasi-experimental design, N = 42

Narrative case study N = 40

TYPE
OR
"FOCUS"
OF
STUDY

General outcome evaluation, N=29

Comprehensive evaluation, N=25

Experimental design, N=17

Input evaluation or audit, N=15

Process evaluation with performance measures, N=14

Process evaluation without performance measures, N=8

Statistical models, N=3

Formal models, N=1

TOTAL N= 194
Missing = 6

TABLE 15
FOCUS OF STUDY

Program \ Focus	Quasi experimental	Narrative case study	General outcome evaluation	Comprehensive evaluation	Experimental design	Input or audit	Process *with* performance measures	Process *without* performance measures	Statistical models	Formal models
assistance/planning/resource allocation	0	7 (17.5)	1 (3)	1 (4)	0	3 (20)	0	2 (25)	0	0
training	6 (14)	5 (12.5)	6 (21)	3 (12)	0	2 (13)	0	1 (12.5)	0	0
corrections	13 (31)	11 (27.5)	6 (21)	5 (20)	4 (23)	8 (53)	3 (21)	2 (25)	0	0
police-community relations	3 (7)	8 (20)	2 (7)	1 (4)	2 (12)	2 (13)	0	1 (12.5)	0	0
pre-trial release	2 (5)	0	2 (7)	4 (16)	0	0	0	0	0	0
courts & jury management	4 (10)	0	2 (7)	2 (8)	0	0	0	1 (12.5)	0	1 (100)
police logistics	13 (31)	4 (10)	6 (21)	3 (12)	9 (53)	0	5 (36)	1 (12.5)	3 (100)	0
AVM/CAD & information systems	0	5 (12.5)	2 (7)	2 (8)	0	0	5 (36)	0	0	0
innovative methods	0	0	0	1 (4)	1 (6)	0	1 (7)	0	0	0
miscellaneous	1 (2)	0	2 (7)	3 (12)	1 (6)	0	0	0	0	0
TOTALS	42	40	29	25	17	15	14	8	3	1

Missing = 6

45

programs, AVM/CAD, resource allocation and pre-trail release. Figure 6
gives a complete breakdown by subject, giving the average score obtained by
each student, and the overall average (of all scores). Those methodologies
(or "foci") which were most often suitably selected (scored highest on
"suitability of focus" ratings) were performance evaluations with performance
measures, formal models and "comprehensive" evaluations. The lowest scores
were obtained by narrative case studies, input evaluations/audits and general
outcome evaluations. The remaining foci all clustered around the mean score.

In general, comprehensive approaches were judged to have been quite
suitable for those studies that attempted them; narrative case studies
were judged as unsuitable when such an approach was attempted. In be-
tween, the experimental/quasi-experimental design and general evaluation
types were judged by the readers as being only moderately suited to the
purpose (which means that some were much less suited than others).
Please note that suitability was rated on a scale of 1 to 7 by three
different readers: only general trends may be indicated at this stage
of the research.

The most common method appeared to be the use of descriptive
statistics. Qualitative analysis was also frequently employed, as was
statistical inference and related methods (eg. regression and analysis
of variance). Unfortunately, information on combinations of methods
used is not available due to the nature of the computer program used to
retain questionnaire responses. The average rating of the methodological
astuteness (or suitability) of the sample is 4.1, corresponding to the
mean rating of 4.3 obtained for the question on suitability of the eval-
uation focus (see Figure 7 for breakdown).

FIGURE 6

SUITABILITY RATINGS FOR STUDY FOCI

Subject of Study*	Suitability of Focus**
+ innovative methodology	6.33
+ assistance/planning/ resource allocation	4.86
- training	3.48
- corrections	4.04
+ pre-trial release	4.88
- police-community relations	3.05
+ courts	5.50
+ police logistics	4.80
- AVM/CAD and information systems	4.07
+ miscellaneous	5.57
Average score:	4.30

* += above average suitability of focus
  - = below average suitability of focus

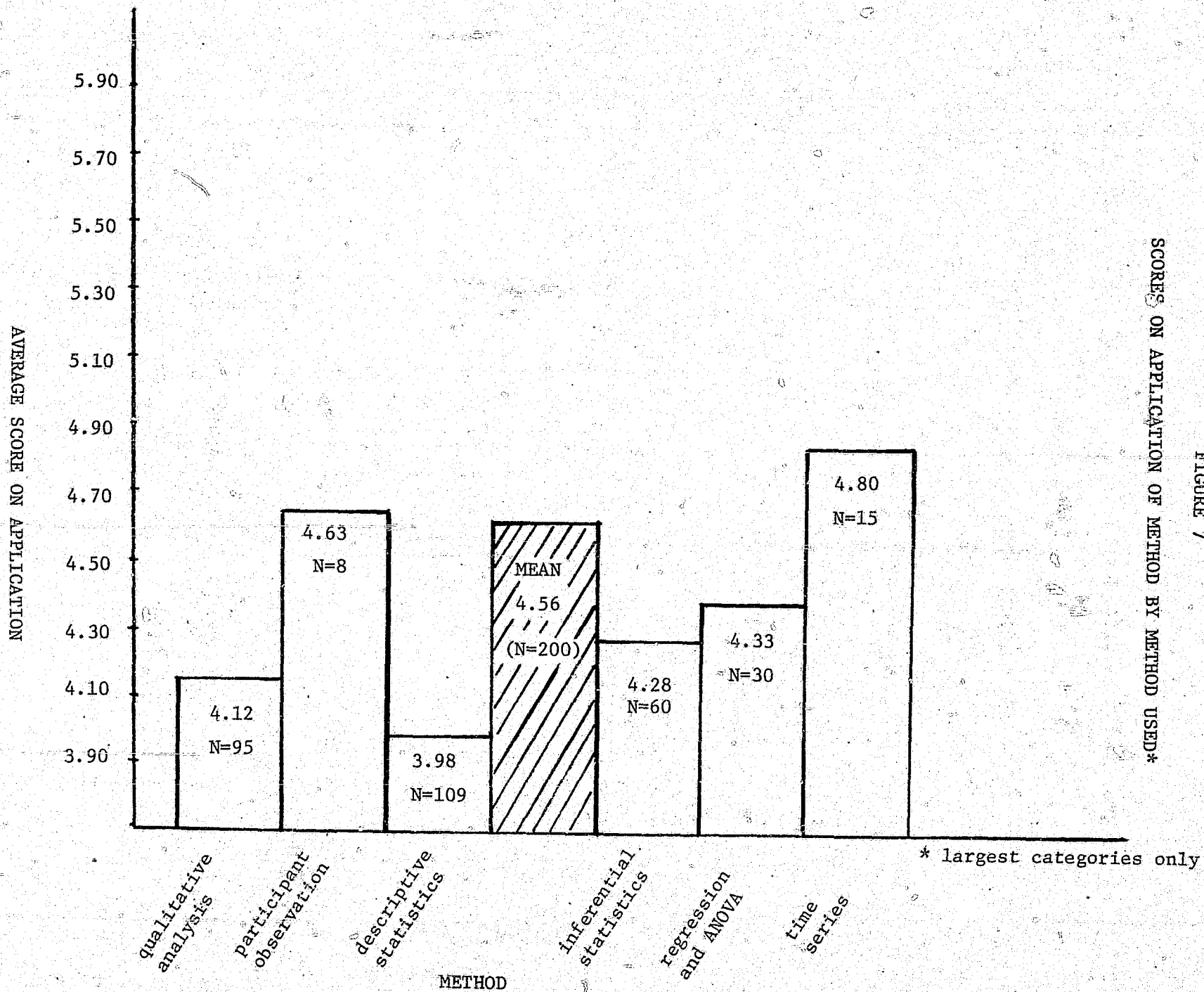** (1 = low suitability, 7 = high suitability)

FIGURE 7

SCORES ON APPLICATION OF METHOD BY METHOD USED*

* largest categories only

48

Responses to the question on adequacy of measures split about 50/50 overall. Breaking it down into categories, over 50 percent of the logistical studies received positive ratings on this question, while fewer than 50% of the social service studies received positive ratings. Two small categories were outstanding: all pre-trial release evaluations received a positive rating, as did all AVM/CAD entries.

Typical comments made on studies that didn't use measures adequately include:

> "The study placed 100% reliance on perceptual data when more objective measures could have been used." (police logistics)

While the majority of studies clearly documented their methods of analysis, the most widespread problem with methods was that the evaluators misapplied or misused common statistical techniques. For example, a comment on an evaluation where regression was used reads: "In this design, the independent variables are all linear functions of each other." (police logistics) Techniques were misused in social service studies also, resulting in excessive generalization from small samples, eg. "...the difference they are referring to here is of 2 or 3 respondents out of a total of 89. Not only is this not statistically significant, but it doesn't even seem to indicate a trend." (corrections -- juvenile diversions) Finally, in an "evaluation" of a training program where chi-square was used, the evaluators "...repeatedly manipulated categories in an attempt to get statistically significant results....which they finally did get".

As a final test of methodology, we asked "Can outcomes be directly attributed to program activities? Is there a theory linking program activities to the performance measures chosen?"

The overall response to the first part of this question was negative (28% received positive ratings, 72% received negative ratings). This response did not seem to depend on the subject area of the evaluation (see Table 16). The most common explanation given in reader comments was that the study was essentially uncontrolled (in the sense of research design). Another common comment effectively stated that causality could not be established due to a lack of information regarding the actual functioning of the program. ie., process information. Responses of "not applicable" may generally be attributed to the readers' inability to fairly answer this question, due to the fact that the evaluation reports were not sufficiently complete.

All study subject areas received at least as many positive ratings on the second part of the question as on the first part (see Table 16). Although a problem of theoretically linking program activities and performance measures does exist, this problem does not seem to be as widespread as the problem of attribution of outcomes described above. The most common comment accompanying a negative rating on this question was to the effect that the measures used were essentially measures of convenience (meaning that it is easy to collect data for them), rather than measures appropriate to the problem at hand.

Data Sources

When examining methodology and methods, it is important to go beyond the approach and techniques chosen and examine the sources of data used to evaluate the program. An evaluation may appear to be methodologically sound, while the choice of data sources results in information which is of limited value for decision-making or program management.

TABLE 16

THEORY BEHIND THE PROGRAM

| Response\Subject	Can outcomes be directly attributed to program activities?					Is there a theory linking program activities to performance measures chosen?				
	YES N (%)	NO N (%)	Sub-total	Not applicable	Total	YES N (%)	NO N (%)	Sub-total	Not applicable	Total
assistance/planning/resource allocation	0 (0)	2 (100)	2	12	14	4 (57)	3 (43)	7	7	14
training	1 (6)	20 (95)	21	2	23	7 (35)	13 (65)	20	3	23
corrections	14 (33)	29 (67)	43	11	54	23 (48)	25 (52)	48	6	54
police-community relations	3 (20)	12 (80)	15	4	19	8 (50)	8 (50)	16	3	19
pre-trial release	4 (50)	4 (50)	8	0	8	4 (50)	4 (50)	8	0	8
courts & jury management	4 (40)	6 (60)	10	1	11	8 (73)	3 (27)	11	0	11
police logistics	10 (25)	30 (75)	5	0	45	28 (68)	13 (32)	41	4	45
AVM/CAD & information systems	4 (57)	3 (43)	7	7	14	9 (90)	1 (10)	10	4	14
methodology	2 (57)	1 (33)	3	0	3	3 (100)	0	3	0	3
miscellaneous	2 (33)	4 (67)	6	1	7	5 (83)	1 (17)	6	1	7
TOTALS	44 (28)	111 (72)	155	43	198	99 (58)	71 (42)	170	28	198

missing = 2

Where did the evaluators go to collect their data? Were the sources chosen suitable and adequate to the task at hand? To explore these questions, we recorded the data sources used by the evaluators and rated the suitability of those sources on a scale of 1 to 7. We assigned low scores if data sources did not provide adequate information or were concerned with irrelevant information. High scores indicated that adequate data were gathered, and that sources were used carefully and revealingly. The results are displayed in Figure 8.

Few surprising or important results emerged from the ratings of data sources. Administrative records are by far the most commonly used data source, which is understandable given their availability and low cost. There were only minor differences in the "suitability of data sources" ratings among the various program areas.

Use of multiple data sources is one way to enhance the validity of evaluation results and provide a rich description of both program activities and outcomes from several standpoints. Many evaluations in our sample tried to use multiple sources: this was particularly the case in the social service categories, where scarcity of sources led the evaluators to try many different approaches.

One subgrouping which seemed to have a particular problem with data sources was Training. Suitability scores in this category were generally lower than average, and frequent remarks were made about the inadequate number and type of sources in these studies. Judging from other questions as well, this particular subject area seems to be posing problems for evaluators. Perhaps evaluators are using an inappropriate methodology for these programs, and something different might be attempted. Given that these are educational programs, some of the innovative approaches

FIGURE 8

SUITABILITY OF DATA SOURCES
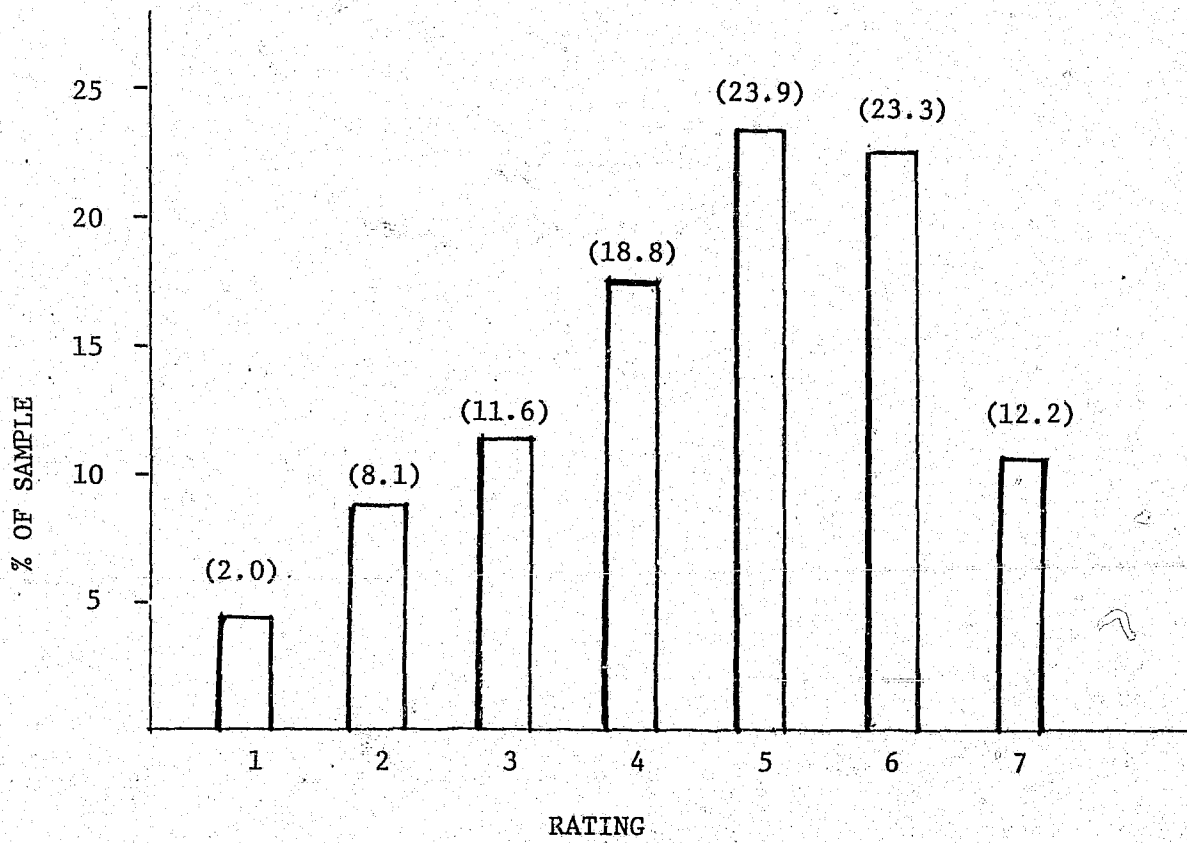
Average rating = 4.7

used in recent years in education, such as ethnography, "illuminative
evaluation" etc., might be considered.

Once data sources have been selected, the key issue becomes how
effectively they are utilized. Some evaluators simply didn't tap sources
with sufficient care and precision, a problem complementary to that of
misusing statistical techniques. Illustrative comments are given below:

> "Questions don't look very good on the questionnaire.
> No negative responses were available to subjects on
> many questions, for example." (police-community
> interaction)

> "This data was aggregated in a ridiculous manner, eg.,
> 'speed-fast, not so fast. medium, slow, very slow,'
> etc.," (police logistics)

Occasionally they did a poor job of interpreting what they collected:

> "Generally comments were quoted verbatim with no analysis
> given." (training)

The ineffective use of data also stemmed, sometimes, from information
overload, ie., too much was collected for no apparent reason.

A further problem seemed to be a lack of documentation. Evaluators
often failed to explain where data came from and how they were used to
draw various conclusions. In particular, where "softer" techniques were
used there was often no explanation of how an observation or interview
was conducted, nor were representative samples of responses given. Fair-
ly common was the assertion that "[conclusion A] is based on interviews",
yet no dialogue samples or summaries were attached. As one comment on a
police-community interaction study noted:

> "These data sources would be excellent if there were clear
> documentation of how they were used and their relationship
> to the researchers' hypotheses."

While it appears that the majority of evaluations have stated con-
clusions that are supported by the data analysis (see Table 17), several

notes of caution must be made concerning these results.

TABLE 17

CONCLUSIONS SUPPORTED BY DATA ANALYSIS

Were the conclusions supported by the data analysis?		
RESPONSE	N	%
YES	124	70
NO	53	30
TOTAL	177	100
Missing = 23		

The proportion of positive responses is too high, as it is based on a generous interpretation of the question. This is particularly so in those situations where no strong conclusions were presented: in these instances, "yes" responses were often given when conclusions were <u>not in conflict</u> with the data analysis, to give the evaluator the benefit of the doubt. Perhaps of greater importance are some of the reasons why the readers decided (when they did so) that the conclusions drawn were not supported by the data analysis. These include the following:

1. no data or data analysis

2. data were of poor quality

3. inconsistency in the data analysis

4. improper measurement

5. no comparison measure

6. recommendations were not based on findings

7. no process component which would explain outcome

8. method problem--obvious threats to internal and external validity

9. evaluators missed obvious alternate explanations

Flexibility

Since one of the concerns of this project is with the appropriateness of existing evaluation methodologies, we were concerned about whether inflexibility of available methods might be an obstacle to successful implementation of the evaluation. For instance, in some situations the demands of a rigorous experimental design utilizing randomized assignment might be highly impractical and difficult to satisfy: in such cases efforts to use randomized assignment are likely to result in distorted designs, and the supposed strength of the methodology will become a weakness. For this reason, it is desirable that a given evaluation plan be designed in such a way that it can be changed as the evaluators learn about the program from preliminary research and from interaction with program staff.

Readers' responses to the question dealing with flexibility of evaluation design were distributed as follows:

TABLE 18

DESIGN FLEXIBILITY

If policy changes or unexpected results caused a need for changes in the evaluation design, was the design flexible enough to account for this?

Response	N	%
Yes	49	25
No	59	30
Not Applicable	89	45
Total	197	100
Missing = 3		

The large number of "n.a." responses to this question reflects the fact that it was often not possible to assess how the evaluators would have or had reacted to changed circumstances; also, certain types of evaluations which were limited to descriptions or ex post facto data analysis would not really be vulnerable to such changes.

The readers' comments on this question revealed a number of ways in which evaluators may flexibly adapt themselves to changing circumstances or unexpected findings. First of all, it is clear that qualitative analyses are in general more flexible than true experimental designs. However, the following suggestions, drawn from reader comments on this question, are relevant to experimental and quasi-experimental designs as well.

1) Open-ended questions in interviews or survey questionnaires are more likely to capture results not anticipated by the evaluators than simple yes/no or multiple choice questions. For example, in discussing an evaluation of the Massachusetts Police Institute, one reader commented on the use of interview questions which were "flexible and designed to elicit program-specified responses."

2) The evaluation instruments and measures used may be modified to reflect experience gained in the early stages of the evaluation. This approach was taken in an evaluation of a Parole Aide Program. The poor rate of response to a written questionnaire among the control group prompted a shift to oral interviews in later phases of the research.

3) Additional sources of data may be employed to compensate for weaknesses in the original data collection plan. In the third year of an evaluation of a Parole Officer Aide Program, new data were gathered to address questions which had remained unresolved by earlier research.

4) Simulations or formal modeling may be used to estimate information not available from the data. This approach was used by Abt Associates in an evaluation of pre-trial intervention; similarly,

a study on jury management used simulation to study
a situation which would have been intractable by
other means.

5) Exploratory data analysis and theory-building
approaches may help to make sense of unexpected
or apparently contradictory results. One example
of this approach is shown in an evaluation of a Pro-
bation Subsidy Program. (In this case, the author's
attempt to formulate a theory which would explain
the unexpected findings was relatively weak and
unconvincing: this need not always be the case,
however.)

6) The use of multiple control groups (or non-equiva-
lent groups as a back-up for true control groups)
can help guard against failures in implementing the
experimental design. This factor turned out to be
crucial to an evaluation of pre-trial release. In
this case, the design for randomized selection of ex-
perimental and control groups was not properly im-
plemented, and it was only the planned use of multiple
non-equivalent control groups in addition which sal-
vaged the experiment.

7) The experimental or data collection period may be
extended if it is clear that conclusive results
have not yet been reached, or if problems have de-
layed program implementation. This would have been
highly desirable, for instance, in a court scheduling
study, where changeover difficulties during the
intended trial period prevented any conclusive eval-
uation of the proposed new scheduling techniques.

Communication

Obviously, the more aware evaluators are of program changes, the

easier it will be to adapt the evaluation plan accordingly. The only

way this awareness can be developed is through regular interaction and

communication with program staff. In addition, if, as is desirable, the

evaluation is being conducted during program implementation, interim re-

sults or insights may be put to best use by being fed-back to the program

staff as they come up.

The item aimed at the communication issue was "Describe any feed-back between evaluators and program staff." Responses were as follows:

TABLE 19

FEEDBACK BETWEEN EVALUATORS AND PROGRAM STAFF

Was there feedback between evaluators and program staff?		
RESPONSE	N	%
YES	49	25
NO	87	44
NOT APPLICABLE	61	31
Total = 197		
Missing = 3		

As the responses to this question do not lend themselves to definitive conclusions, it will be necessary to investigate the issue further with the author and consumer questionnaires. Some emerging patterns may be discussed, however.

Positive responses to this question frequently contained comments about the formative functions served by this feedback. The key issue here is user-orientation. Despite the obstacles to formative evaluation, such as inflexible experimental design, it is important to consider alternative ways of getting helpful information into the hands of program staff _during_ a long term evaluation.

The non-use of reports may often be traced to a time-lag problem, meaning that by the time the evaluation is completed, the program may have been substantially transformed or even discontinued.

Other comments about feedback fell into two general cases:

1. There was a high degree of communication (although summative rather than formulative) and cooperation between staffs which smoothed the way for the evaluation to proceed;

2. There was an obvious degree of tension or con-
flict between staffs.

The relatively high proportion of "n.a." responses to this question
was due to a number of factors. Based on the readers' comments, it seems
that about half of the "n.a." responses occurred in cases in which the
evaluation report simply didn't provide enough information for the read-
er to draw any conclusions. (For this reason, the question of feedback
will be pursued further through the author and consumer questionnaires.)
In about a quarter of the cases the feedback question was not directly
applicable because the evaluation was performed by the program staff.
The remaining 25 percent of n.a. responses occurred in cases which were
overviews of a field or had similar emphases.

OUTPUT

For our present purposes, the major output considerations are
accessibility and use of the final report of the evaluation. There are
two aspects to accessibility--presentation and availability.

One of the most basic requirements for a genuinely useful evaluation
is that data and the consequent analysis be presented in a clear fashion.
The information that inspired conclusions should be readily understandable
and complete enough for re-analysis by the reader: this is a way of
assuring the integrity of the research.

In order to locate deficiencies in data presentation, we asked the
readers "Is the data presentation adequate?" As shown in Table 20, the
majority (63 percent) of evaluations received a positive rating on this
item, indicating that presentation was not a major problem. The social
service categories tended to have somewhat more problems with data

TABLE 20

DATA PRESENTATION

Program \ Response	YES		NO		sub-total	not applicable	Total
	N	(%)	N	(%)			
assistance/plan-ning/resource allocation	7	(44)	9	(56)	16	1	17
training	14	(74)	5	(26)	19	0	19
corrections	27	(54)	23	(46)	50	4	54
police-community relations	8	(44)	10	(56)	18	1	19
pre-trial release	7	(88)	1	(12)	8	0	8
courts & jury management	11	(100)	0	(0)	11	0	11
police logistics	34	(79)	9	(21)	43	2	45
AVM/CAD & information systems	6	(46)	7	(54)	13	1	14
miscellaneous	5	(71)	2	(29)	7	0	7
TOTALS	119	(64)	67	(36)	186	9	195

Is the data presentation adequate?

Missing = 5

presentation than did police logistics. This may indicate that evalua-
tors have difficulties presenting qualitative data: comments made
under other questions would seem to confirm this hypothesis. Evaluators
often fail to give representative samples of interviews, or of observa-
tion results, thereby lessening the credibility of the conclusions
reached. Some reports, however, did do a good job of presenting qualita-
tive data, judging by readers' comments.

Four problems seemed to be generic among those programs which were
judged as having presented data inadequately. These were noted in most
or all program areas, although the non-logistical are the most common
offenders. These problems were:

1. Consistency. It was noted several times that the
   data presented were not consistent with the measures
   and/or sources chosen by the evaluator. Evaluators
   would conduct interviews, for example, but provide
   no data on the interviewing process or outcomes.
   Data were sometimes tossed to the reader with little
   or no explanation of how or why it fit into the
   theory or analysis;

2. Completeness. For qualitative methods, summaries
   were often used when more detail was needed. Similar
   problems occurred with quantitative data, especially
   when it came to presentation of aggregated results
   when it would have been better to present them in more
   raw form. Many evaluations did not contain enough
   information for the reader to re-analyze the report
   and confirm the conclusions. In both qualitative and
   quantitative studies the presentations were often
   far too sparse;

3. Legibility. If evaluations are going to be user-orient-
   ed they should be readable. Although "readability"
   was not frequently mentioned as a problem (by graduate
   students immersed in the field, n.b.), the presentation
   of too much data was. An overabundance of detail can
   render an evaluation almost as useless as one with no
   data presentation, especially from the practitioners'
   point of view,

4. <u>Correctness</u>. There were some comments about improper grouping of data, incorrect calculations, use of statistical tests when N's were too small, etc. (See the section on methodology and methods).

For evaluations to be useful, it is imperative that results and interpretation of results be presented legibly, completely and correctly. The possibility should be open for the reader to reinterpret the results and confirm evaluation integrity. A good many in our sample did not meet some or all of these criteria.

Presentation is only one aspect of the accessibility of final reports: the other is availability. Our efforts to obtain the evaluation reports needed to do the research may be summarized as follows:

Our original search target was 283 documents: through a process of elimination we eventually obtained 232 documents, at an average rate of 1.5 requests per document (358 requests had to be processed in order to obtain these 232 documents), using 68 different sources.

Of the original documents ordered, 210 were ordered using information from the NCJRS printout. Out of these, 34 were only obtained after trying a second source, and 2 of them required three letters, for a total of 36 NCJRS problem cases.

An additional 20 documents were ordered from the NCJRS printout and came back listed as "missing" from the NCJRS library or as being out of print at the source, i.e., these 20 were simply unobtainable. Fifty-three reports were ordered from source documents other than the NCJRS printout. The sources checked through to get these 53 documents included: individual book publishers, universities, National Technical Information Service (NTIS), Sage Publications, Abt Associates, Inc., MITRE Corporation, U.S. Government Printing Office, The Urban Institute, International City

Management Association (ICMA), Project SHARE, Rand Corporation, Police
Foundation and the International Association of Chiefs of Police. In
most cases the NCJRS Document Loan Program does have a document if it
says it does (unless it has been borrowed by another user and not re-
turned, or if it is out on loan to another user and isn't available for a
few months). The main problem occurred in cases where a company, agency
or university was listed as the principal source for a document, and in
following the lead we would discover that the source was out of business,
and would be referred to another source. In some cases a helpful individ-
ual in one of the referred agencies would send us a last copy of a docu-
ment as a favor or private loan (our requests were worded most urgently).
It is likely that if many people were to write for this same document this
practice would not be continued.

The sources are not kept up to date on the NCJRS printout. This may
be because NCJRS does not get enough feedback from users (for example,
we have not yet told them of the 36 problem reports we had). It is
likely that the average person might not want to let all that time elapse
or write a series of letters in order to obtain one document.

Finally, retrieval problems were not limited to the NCJRS printout.
Multiple requests also had to be made for about 20 of the 53 documents
ordered through other sources.

It should be clear from the above report that although evaluations
of criminal justice programs are for the most part available from the Docu-
ment Loan Service of the National Criminal Justice Research Service (NCJRS),
the procedure for obtaining them is long, complex and frustrating.
Managers of criminal justice programs may have the training, patience and
staff needed to follow this procedure through to the end: in many

instances local government officials may not have the necessary staff
and/or expertise. Community leaders and concerned citizens groups very
likely would not be able to use this method as successfully as someone as
persistent and articulate as Dorothy Green (although naturally
they would not need to get hold of as many documents).

Use

Relatively few evaluation reports indicated that evaluation findings
were used in actual decision-making. Whether evaluators felt that this
was not a significant component of the evaluation report or whether this
is a true indication of non-use is not entirely clear at this time. The
timing of the publication of the evaluation report could also be a factor
here.

The indication of use did not vary among the agencies responsible
for the evaluations. Studies conducted by the Police Foundation, for ex-
ample, were generally as likely to indicate the use of findings as were
studies conducted by the State Planning Agencies. There was little
variation among subject areas as well. Evaluation reports in the field
of pre-trial release and corrections seemed somewhat more likely to indi-
cate evaluation use than studies in other substantive areas of our sample.
It is reasonable to suspect that this may have had more to do with the
timeliness of a particular issue than with any inherent quality of the
evaluations themselves. Nonetheless one might want to review these eval-
uations to determine if one _type_ of evaluation is consistently used more
often in decision-making. These issues will be explored further via the
author and consumer questionnaires.

Timeliness is a serious problem affecting the utilization of

evaluations. There were indications in several of the evaluation reports that the results were received by program staff too late to affect the decision-making process. This problem was discussed more fully in the section on communication, and will be pursued further in the second stage of the research.

Several articles on evaluation use make note of the oral tradition in governmental decision-making. The claim is that knowledge that is transmitted by word of mouth tends to have greater impact than the written word. If this is indeed the case, evaluators might consider alternate or complementary ways of disseminating their findings. Oral presentation of findings prior to publication could be a useful tool for capturing project staff reaction and suggestions which could perhaps be incorporated into the final report. Other advantages of continuous and/or timely communication of findings were discussed in the section on communication and flexibility.

Since so few evaluators reported that their findings were or were not used in decision-making, we must ask if evaluators should give any indication of evaluation use. There are at least two reasons why such information would be useful to the reader of an evaluation:

1. Evaluations are not produced in isolation. A perfectly reasonable evaluation may fail to influence decision-making because of political reasons. It still could be useful to managers of similar programs in different jurisdictions, by indicating which political pitfalls to avoid;

2. It is important for policy-makers to be able to assess the responsiveness of the administrative and operational structure of social action programs to evaluation of their work.

The study tried to determine whether the structures of inquiry used in the evaluations could be used for periodic review by the program staff and by

others interested in a similar program. In general, the evaluations in
our sample tended to score poorly in this regard.

TABLE 21

POTENTIAL FOR PERIODIC REVIEW

Could this type of evaluation be adapted for use in periodic review?		
RESPONSE	N	%
YES	85	45
NO	105	55
TOTAL	190	100
(missing or no answer = 10)		

From the researchers' comments, it appears that an evaluation is suitable
for use in periodic review when it has a fairly strong process component
and/or a strong and appropriate methodology with accurate methods of data
collection and analysis.

CONCLUSIONS

The results of this phase of research may be summarized as follows:

-Basic inputs to the evaluation effort are rarely documented in the evaluation reports: this is the case for the budget, timing, duration and planning of the evaluation.

-Evaluations sponsored by State Planning Agencies and the Police Foundation tended to be planned prior to program implementation.

-Most (66%) of the evaluations in our sample took place during program implementation/operation.

-Those studies that were conducted prior to program implementation (5%) were either small pilot studies preceding a large experimental program, or they were "evaluability assessments" or simulations.

-Only 22% of the evaluations in our sample were conducted by in-house personnel; it was difficult to determine the background of people conducting the evaluation for almost the entirety of our sample.

-The goals, resources and perceptions of purpose of evaluation vary more by agency undertaking the evaluation than by subject area.

-Police Logistics and pre-trial release reports and programs scored well on goal specification, while social service reports and programs gave a relatively poor showing.

-The issue of target population was not discussed by the evaluators in 1/3 of our sample. A slight majority of reports did not consider whether the program had been implemented as designed, and description of program activities is frequently inadequate as well.

-Experimental and quasi-experimental design are the most popular evaluation types, followed by narrative case studies; there was almost no use of statistical or formal models.

-Narrative case studies, input evaluations (or audits) and general outcome evaluations were generally poorly matched to the substantive program they were being used to evaluate.

-The most common methods used were descriptive statistics, several types of qualitative analysis, regression and ANOVA. Suitability scores on methods were not particularly outstanding.

-All pre-trial release and AVM/CAD studies received a positive rating on the adequacy or proper use of measures. The most widespread problems were misapplication of common statistical techniques and difficulties in attributing outcomes to program activities, i.e., poor choice of performance measures.

-Administrative records are the most commonly used source of data.
Attempts were made to use multiple sources, particularly in the
social service categories. Evaluations of training programs tended to
use less than suitable sources of data.

-There is a generalized lack of documentation of data collection
procedures, and data were sometimes poorly used once obtained.
A complementary problem is poor presentation, more so in qualita-
tive than in quantitative studies.

-From the evaluation reports it is often difficult or impossible to
assess how the evaluators would have or had reacted to changed cir-
cumstance. A number of avenues are open to evaluators to enhance
the flexibility of their studies.

-Relatively few evaluation reports indicated whether evaluation findings
were used in actual decision-making, with no variation by agency.

Many of the above points are manifestations of the basic problem

with the criminal justice evaluations in our sample, namely that quite

frequently the evaluation methodology used is not well matched to the

type of program being evaluated. There has been little or no attempt,

not to say success, at combining methodologies to achieve the purpose of

evaluating a particular program.

A contributing factor seems to be that it is not common practice to

have well-structured hypotheses or mental models concerning how the

program should work. It is very important that the evaluator have some

notion of how program activities are linked to desired outputs and to

other social, economic and political activities in the subject community:

in many instances, the use of statistical or other formal models would

help immensely. The point of stressing the need for articulated hypoth-

eses is to wean evaluators away from the textbook formulas to which they

were taught to adhere with little regard for circumstances. As selection of

hypotheses by administrative means is equally undesirable, it is necessary

to develop some sort of "performance measures" by which to assess the effi-

cacy of theoretical models that might be used in criminal justice

evaluation in the future.

On a different level, difficulties in applying various types of social science methods and measures were frequently manifested. They occur throughout our sample, however, a major problem seems to be that evaluators attempt, in good faith, to apply quantitative methods wherever they thought they might be appropriate. This could be due to input constraints on the evaluation, or to misunderstanding of the technical complexity of some of the measures. Common sense occasionally gets lost in the pursuit of elegant methods. In terms of evaluation usability, one of the more serious problems confronting a decision-maker is the final report (of the evaluation) itself. Final reports frequently arrive too late to influence decisions to which they may have had relevance. Whatever the timing, they often present very incomplete pictures of the program and of the evaluation: information on the inputs to and process of the evaluation, or of the program, is often impossible to obtain from the reports. This is unfortunate because it prevents the decision-makers from making a fair assessment of the program, and inhibits improvement of evaluation and hence program practice.

In fact, in order for us to make recommendations regarding evaluation practice, we must first uncover just this kind of input and process information. To do this, we are administering questionnaires to the authors (evaluators) and consumers (program managers and funders) of the evaluation reports which we read. These questionnaires are designed to elicit information on the budgeting, timing, planning and design of the evaluation (inputs), interaction between program staff and evaluators, e.g., communication (process), and on the ultimate use of the evaluation.

Responses to these questionnaires will be cross-checked with each other and with the original reader questionnaires that were used to evaluate the reports.

FOOTNOTES

1. Some of the more well-known evaluations, say those supported by the Police Foundation or by LEAA's National Institute of Law Enforcement Criminal Justice, were selected independently of the NCJRS listings.

2. R.C. Larson, "A Framework for the Comprehensive Evaluation of Evaluations," to appear.

3. Kaplan, Edward H. (1978) Evaluating The Effectiveness of One- Versus Two-Officer Patrol Units, Technical Report No. 153, Operations Research Center, M.I.T.

   Kaplan, Edward H. (1979), Models For The Evaluation of Treatment-Release Corrections Programs, Technical Report, Operations Research Center, M.I.T.

4. Irwin Deutscher, "Toward Avoiding the Goal Trap in Evaluation Research" in Francis G. Caro (ed.) Readings in Evaluation Research, 2nd edition (New York: Russell Sage Foundation, 1977) pp. 221-239.

   Donald T. Campbell, "Reforms as Experiments" in Elmer L. Struening and Marcia Guttentag (eds) Handbook of Evaluation Research, Vol. 1 (Beverly Hills: Sage Publications, 1975) pp. 71-137.

Title	Author(s)	Published	Source/Status
1. "Evaluation Report on the Model Cities Team Policing Unit of the Holyoke (MA.) Police Dept."	J. Angell, et al.	1972	NCJRS Document Loan
2. "Cleveland Impact Cities Program- Deterrence, Detection, and Apprehension, Operating Program Final Evaluation Report"	Clev. Impact Cities Program	1975	NCJRS Document Loan
3. "Holyoke (MA.) Team Police Experiment-Evaluation Report" and Supplemental Reports & Appendices	Holyoke Police Department	1973	NCJRS Document Loan
4. "Oakland Police Department's Digital Communications System"	Oakland Police	1972	Project Owned
5. "PATRIC (Pattern Recognition and Information Correction) Evaluation Report, March 5-April 21/73"	Calif. Council on Criminal Justice	1973	NCJRS Document Loan
6. Police Personnel Management Information Systems: The Dallas and Dade County Experiences	Police Foundation	1977	Project Owned
7. The Crime Control Team - An Element of an Offensively Deployed Municipal Police Department	General Electric Elec. Laboratory	1970	Project Owned
8. PAR (Pooling All Resources) Policing - Final Report	-	1973	NCJRS, Microfiche Project Owned
9. Albuquerque Police Department, Property Crime Reduction Program	W.V. Niederberger W.F. Wagner	1974	NCJRS, Microfiche Project Owned
10. Crime-Specific 2 Burglary Program - Report on Evaluation	System Development Corporation	1975	NCJRS, Microfiche Project Owned
11. Baltimore-Impact Courts Program - Evaluation Report	Baltimore Mayor's Coordinating Council on Criminal Justice	1974	NCJRS, Microfiche Project Owned
12. Evaluation of San Francisco Police Computer-Assisted Dispatch Project	W.R. Partridge	1975	NCJRS, Microfiche Project Owned
13. Community Sector Team Policing	Cincinnati Police	1974	NCJRS, Microfiche Project Owned

Title	Author(s)	Published	Source/Status
14. Team Policing Planning Guide	Los Angeles P.D.		NCJRS, Microfiche Project Owned
15. Delaware County Total Information System Project - Final Evaluation Report	L. Polansky	1975	NCJRS, Microfiche Project Owned
16. "Evaluation of Operation Identification as Implemented in Illinois"	H.W. Mattick, et al.	1974	Project Owned
17. Hartford (Ct.) Housing Authority Security Patrol - Evaluation Report	J. Carmen, et al.	1972	NCJRS, Microfiche Project Owned
18. New Haven (Ct.) Case Incident Regional Reporting System (CIRRS) Exemplary Project Validation Report	Abt Associates	1976	NCJRS, Microfiche Project Owned
19. Newark (Ca.) Police Department - Project Stop (Master Plan) 1st Year, October/71-September/72	Newark Police Department	1972	NCJRS, Microfiche Project Owned
20. Atlanta - Impact Program - Master Plan	Atlanta Regional	-	NCJRS, Microfiche Project Owned
21. Riverside (Ca.) Police Department Project ACE (Aerial Crime Enforcement)	R.E. Hoffman	1972	NCJRS, Microfiche Project Owned
22. "Special Police Units in Michigan: An Evaluation"	R.G. Lewis, et al.	-	Project Owned
23. "Penn. Community Treatment Services: An Evaluation and Proposed Evaluation Information System"	Informatics, Inc.	1972	Project Owned
24. "Systems Analysis Training Program for Beverly Hills Police Department" Volumes I, II, and III	Beverly Hills Police Department	1972	Project Owned(I), NCJRS Document Loan (II, III)
25. "Social Evaluation Research - The Evaluation of Two Police Patrolling Strategies" Journal of Applied Behavior Analysis, Vol. 8, No. 4	Schnelle, et al.	1975	Project Owned
26. "Computerized Scheduling of Police Manpower, Volumes 1 and 2"	N.B. Heller	1973	Project Owned

Title	Author(s)	Published	Source/Status
27. "Evaluation of Operations Neighborhood"	P.B. Bloch and D.I. Specht	1973	Project Owned
28. "Improving Patrol Productivity, Volume I, Routine Patrol; Volume II, Specialized Patrol"	W.G. Gay, et al.	1977	Project Owned
29. "Final Evaluation for Grant #1161 (Burglary Reduction Program)"	Seattle Police	1975	Project Owned
30. "Evaluation of a Police-Implemented AVM System: Phase I," A Summary Report	Public Systems Evaluation, Inc.	1976	Project Owned
31. "An Evaluation Report of the Worcester Crime Impact Program" and Appendices	Public Systems Evaluation, Inc.	1975	Project Owned
32. "Examination of Three Data System Projects – High Impact Anti-Crime Program"	L.A. Greenfeld	1975	NCJRS, Microfiche Project Owned
33. "Salt Lake City – Strategic Patrol and Coordination Effort (SPACE) – Federal Grant Evaluation, 73/74"	Salt Lake City	1974	NCJRS, Microfiche Project Owned
34. "Albuquerque/Bernalillo County – Criminal Justice Action Plan – 1975 – Concepts in Criminal Justice"	Albuquerque Met. Crim. Just. C.C.	1975	NCJRS, Microfiche Project Owned
35. "Report on Investigative Effectiveness – A Comparison of Three Investigative Models"	The Urban Inst.	1974	NCJRS, Microfiche Project Owned
36. "St. Petersburg – Police Department – Aviation Unit – Evaluation Report"	St. Petersburg Police Department	1974	NCJRS, Microfiche Project Owned
37. "Crime-Correlated Area Model – An Application in Evaluating Intensive Police Patrol Activities"	F.S. Budnick	1972	NCJRS Document Loan
38. "Evaluation of the Community Centered Community Team Policing Program"	Wright State Univ.	1971	NCJRS Document Loan
39. "Philadelphia – Police Department – West Philadelphia Strike Force Act 1; Act 2, Final Report	J. Cooper, et al.	1974	NCJRS Document Loan

	Title	Author(s)	Published	Source/Status
40.	"Hawaii – Law Enforcement Program – Evaluation of Selected Projects"	C.T. Araki, et al.	1975	NCJRS Document Loan
41.	"Los Angeles Police Department – Crime Specific Team Policing-Final Evaluation Team 28"	Los Angeles Police Department	1974	NCJRS Document Loan
42.	"Personal Patrol Car Programs – Evaluation Report – Prince George's County Police Department"	Anon	1973	NCJRS Document Loan
43.	"A Review of Six Research Studies on the Relationship Between Police Patrol Activity and Crime"	J.S. Dahmann	1974	NCJRS Document Loan
44.	"An Examination of the Impact of Intensive Police Patrol Activities"	F.S. Budnick	1971	NCJRS Document Loan
45.	"Crime Reduction in Albuquerque – Evaluation of Three Police Projects"	P.M. Sears	1973	NCJRS Document Loan
46.	"Examination of Police Patrol Effectiveness – High Impact Anti-Crime Program"	J.S. Dahmann	1975	NCJRS Document Loan
47.	Policewomen on Patrol (Final Report)	P.B. Bloch and D. Anderson	1974	Project Owned
48.	Patrol Staffing in San Diego (One- or Two-Officer Units)	J.E. Boydstun, et al.	1977	Project Owned
49.	Field Evaluation of the Hypercube System for the Analysis of Police Patrol Operations: Final Report	The Institute for Public Program Analysis	1977	Project Owned
50.	An Experiment to Determine the Feasibility of Holographic Assistance to Fingerprint Identification	Project Search	1972	Project Owned
51.	Random Digit Dialing, Lowering the Cost of Victimization Surveys	Police Foundation	1976	Project Owned
52.	The Cincinnati Team Policing Experiment, A Technical Report (Volume I & II)	The Urban Institute	1977	Project Owned

Title	Author(s)	Published	Source/Status
53. Managing Investigations: The Rochester System	P.B. Bloch and J. Bell	1976	Project Owned
54. The Kansas City Preventive Patrol Experiment A Technical Report	The Police Foundation	1974	Project Owned
55. Police Response Time, its Determinants and Effects	Midwest Research Institute	1976	Project Owned
56. San Diego Field Interrogation, Final Report	J.E. Boydstun	1975	Project Owned
57. Countywide CAPER Project, 1974-1975 Report	CAPER	1975	NCJRS Document Loan
58. "Patrol Evaluation Research: A Multiple Baseline Analysis of Saturation Police Patrolling During Day and Night Hours" Journal of Applied Behavior Analysis	J.F. Schnelle, et al.	1977	Project Owned
59. The Indianapolis POlice Fleet Plan	Urban Institute	1970	Project Owned
60. An Evaluation Report of an Alternative Approach in Police Patrol: The Wilmington Split-Force Experiment	Public Systems Evaluation, Inc.	1977	Project Owned
61. Response Time Analysis, Executive Summary, Vol. I/II	Kansas City P.D.	1977	Project Owned
62. Three Approaches to Criminal Apprehension in Kansas City: An Evaluation Report	T. Pate, et al.	1976	Project Owned
63. Operational Analysis of Police Field Force Command & Control in San Jose	Sylvania Electronic Systems	1968	Project Owned
64. "Connecticut Justice Commission, Summary of Evaluations"	Conn. Justice Comm.	1978	Project Owned
65. "Urban Field Test of Four Vehicle Location Techniques"	MITRE Corp.	1973	Project Owned
66. "Vehicel Location & Status Reporting System (LOCATES), Phase II, Final Report"	City of Montclair Police Department	1974	Project Owned

CONTINUED

1 OF 2

Title	Author(s)	Published	Source/Status
67. "Automated Status Reporting Police Communications Study: Final Report	Applied Technology for Chino P.D.	1971	Project Owned
68. Atlanta Impact Program - Master Plan	Atlanta Regional Commission	1972	NCJRS, Mictofiche Project Owned

Reader: Cheryl Mattingly

69. San Diego Community Profile Final Report	J.E. Boydstun M.E. Sherry	1975	Project Owned
70. Project New Pride, Denver Colorado	C.H. Blew, et al.	1977	Project Owned
71. New York City Police Department Street Crime Unit - An Exemplary Project	Abt Associates	1975	NCJRS, Microfiche Project Owned
72. The Police and Interpersonal Conflict: Third-Party Intervention Approaches	Police Foundation	1976	Project Owned
73. Community-Based Corrections Program, Polk County	-	1973	Project Owned
74. Community-Based Adolescent Diversion Program	-	1977	Project Owned
75. "Program Area Local Correctional Institution Rehabilitative System Management and Service Delivery" Final Report	Evaluation Unit of the New Jersey State Law Enforcement Planning Agency	1977	Project Owned
76. "Program Area Improvement of Police Service to Juveniles"	As Above	1977	Project Owned
77. Preventing Delinquency Through Diversion - The Sacramento County Probation Department 601 Project-A Third Year Report	R. Baron and F. Feeney	1974	NCJRS, Microfiche Project Owned
78. Pre-Service and In-Service Training of Georgia Correctional Personnel, An Interim Evaluation	J.A. Nosin	1975	NCJRS, Microfiche Project Owned
79. First Year Development of a Master's Degree Program in Judicial Administration	J.L. Fazio	1972	NCJRS, Microfiche Project Owned

Title	Author(s)	Published	Source/Status
80. "Executive Summary: Model Evaluation Project: Youth Service Bureaus in Michigan"	Criminal Justice Systems Center Staff	–	Project Owned
81. Kansas City (Mo.) Peer Review Panel - An Evaluation Report	Pate, et al.	1976	Project Owned
82. "The Elementary School Child's Perception of Police and the Police Function" An Evaluation Study of the Officer Friendly Program	University of Chicago	1972	Project Owned
83. "Experiment in Delinquency Prevention and Control"	R. Pooley	1971	Project Owned
84. Pivotal Ingredients of Police Juvenile Diversion Programs	M. Klein	1975	NCJRS, Microfiche Project Owned
85. "Santa Clara Criminal Justice Pilot Program - Final Report, Phase I"	American Justice Institute	1971	Project Owned
86. "Evaluation of Five Drug Treatment and Rehabilitation Projects - Cluster Evaluation"	G. Fink, et al.	1974	NCJRS, Microfiche Project Owned
87. "Cluster Evaluation of Five Diversion Projects - Final Report"	A.K. Bean and F.R. Campbell	1974	NCJRS, Microfiche Project Owned
88. "Evaluating Drug Treatment Programs - A Review and Critique of Some Studies on Programs" Drug Forum	B. Sugarman	1974	Project Owned
89. "Northeast Denver Youth Services Bureau Final Evaluation Report"	Behavioral Research Institute	1974	Project Owned
90. "Philadelphia Neighborhood Youth Resources Center - An Exemplary Project"	Anon	1975	Project Owned
91. "Portland State University - Police-Community Relations Evaluation Report"	Portland State University	?	Private Loan
92. "Effect of Using the Police Radio in Teaching the New Criminal Code to Kentucky Police Officers"	D.J. Wiechman	1973	NCJRS Document Loan

	Title	Author(s)	Published	Source/Status
93.	"Tennessee - Law Enforcement Training Academy - Evaluation Report"	Bettiol, et al.	1974	NCJRS Document Loan
94.	"Massachusetts Police Institute - Evaluation"	A.D. Little	1976	NCJRS Document Loan
95.	"Evaluation of the Washington Criminal Justice Education and Training Center"	Fogarty, et al.	1974	NCJRS Document Loan
96.	"National Crime Prevention Institute - Final Progress Report, Fiscal Year 1975/1976"	National Crime Prevention Institute	1976	NCJRS Document Loan
97.	"Police Management Career Development Seminars Part 2 - Evaluation Report"	Anon	1976	NCJRS Document Loan
98.	"Pilot Police Project - A Description and Assessment of a Police-Community Relations Experiment in Washington"	R.M. Kelly	1972	NCJRS Document Loan
99.	"Arrest Decisions as Prelude to ?: An Evaluation of Policy Related Research, Volume I: Administrative and Training Script"	M.G. Neithercutt	1974	NCJRS Document Loan
100.	"Arrest Decisions as Preludes to ?: As Above, Volume II: Study Design, Findings, and Policy Implications"	M.G. Neithercutt W.H. Moseley	1974	NCJRS Document Loan
101.	"Police-Community Action - A Program for Change in Police-Community Behavior Pattterns"	T. Eisenberg	1973	NCJRS Document Loan
102.	"Zero-Zero-Nine: A Report"	D.H. Burns	?	NCJRS Document Loan
103.	"Evaluation Report of the Narcotics Education League's Residental Treatment Program for Chicano Heroin Addicts"	J.H. Langer	1975	NCJRS Document Loan
104.	"Minnesota - 180 Degrees, Inc. - Community Re-Entry Center - A Preliminary Evaluation Report"	Minn. Governor's Comm. on Crime Prevention & Control	1974	NCJRS Document Loan
105.	"Minnesota - Police-School Liaison Projects - An Evaluation"	As Above	1973	NCJRS Document Loan

	Title	Author(s)	Published	Source/Status
106.	"Evaluation of the Temple University Human Relations Training Program for the Philadelphia Police Dept."	Bartell Assoc.	1974	NCJRS Document Loan
107.	"New Jersey - Police Training Commission - Mobile Training Units, V.1, Final Report Project Evaluation"	L.A. Cullo	1968	NCJRS Document Loan
108.	"Police Foot Patrol - An Evaluation"	R.T. Lynch	1971	NCJRS Document Loan
109.	"Community Service Officer - Cluster Evaluation"	Booz-Allen and Hamilton, Inc.	1974	NCJRS Document Loan
110.	"West Virginia - Comprehensive Training Program for Correctional Personnel -Final Narrative Report, November 1, 1969 - May 31, 1970"	West Virginia University	1970	NCJRS Document Loan
111.	"Answer to a Challenge - Final Report on In-Service Training for Correctional Personnel"	Eastern Kentucky University	1969	NCJRS Document Loan
112.	"Oregon Correctional Training Systems Project Developmental Phase Final Report"	D.R. Rinehart	1969	NCJRS Document Loan
113.	"Statewide Penal Code Training - An Evaluation of Grants A-70-176-53, A70-176-53CC, and A71-2111-152"	N. Robb	1973	NCJRS, Microfiche Project Owned
114.	"Community Treatment and Research Program for Juvenile Offenders - Final Report"	M.G. Itkin	1975	NCJRS, Mocrofiche Project Owned
115.	"Des Moines Project - Coordinating Community Corrections - Training Handbook"	D. Boorkman, et al.	1975	NCJRS Document Loan
116.	"In-Service Training for Law Enforcement Personnel - Final Report"	M.L. Brazeal	1970	NCJRS Document Loan
117.	"Developmental Laboratory for Correctional Training - Interim Report"	Southern Ill. Teaching & Research Ctr.	?	NCJRS Document Loan
118.	Santa Clara County Pre-Delinquent Diversion Project, Second Year			
119.	Santa Clara County Pre-Delinquent Diversion Project, First Year			

Title	Author(s)	Published	Source/Status
120. Criminal Justice Education: The End of the Beginning	John Jay College of Crim. Justice	1978	
121. Crisis Intervention Training - An Experimental Evaluation Program in Immediate Action Decision Making	Trenhelm & Whiteneck	1974	NCJRS Document Loan
122. National Youth Project Using Mini-Bikes - Annual Report, Nov. 16, 1973 - Dec. 31, 1974	F.Y. Hoshiyama	1975	NCJRS Document Loan
123. National Youth Project Using Mini-Bikes Annual Report, September 2, 1972 - Nov. 15, 1973	F.Y. Hoshiyama	1974	NCJRS Document Loan
124. "Impact of Training on Job Related Decisions - An Evaluation of the Colorado Youth Workers Training Center"	Larson & Whiteneck	1973	NCJRS Document Loan
125. Federal Juvenile Delinquency Programs, Second Analysis & Evaluation, Vol. 1	Office of Juvenile	?	
126. Time Out: A National Study of Juvenile Correctional Programs	R. Vinter	1976	
127. Models for Police-Public Interaction and Police Policy Development Procedures in Portsmouth, Virginia	Fitch, Morrow, and Gray	1973	Project Owned
128. Theory and Practice of Delinquency Prevention in the United States: Review, Synthesis and Assessment	J.P. Walker Ohio State Univ.	1976	NCJRS Document Loan
129. Police Community Relations Training Program	DeYoung	1968	
130. "Evaluation of the Development and Upgrading of Basic Judicial Skills: National College of the State Judiciary, Final Report"	Liacouras, et al.	1974	NCJRS Document Loan
131. "Behavior & Attitude Modification in a Jail Setting"	Met. Dade County	1971	NCJRS Document Loan

SELECTED SAMPLE OF 200 EVALUATION DOCUMENTS – APPENDIX B

Title	Author(s)	Published	Source/Status
132. Community Treatment and Research Program for Juvenile Offenders – Final Report	M.G. Itkin	1975	
133. "Alternate Routes: A Diversion Project in the Juvenile Justice System," Evaluation Quarterly	G.R. Gilbert	1977	Project Owned
134. "Evaluation and Case Study of a School-Based Delinquency Prevention Program"	P. Higgins	1978	

Reader: Vicki Bier

Title	Author(s)	Published	Source/Status
135. Court Planning and Research: The Los Angeles Experience	D. McGillis & L. Wise	1976	Project Owned
136. The Major Offense Bureau (MOB) Bronx County District Attorney's Office, New York		1977	Project Owned
137. "Court Information Systems" Phase 1 Report	NEP	1977	Project Owned
138. One Day/One Trial: Jury System, Wayne County, MI	K. Carlson	1977	Project Owned
139. Central Police Dispatch (CPD), Muskegon County, MI		1975	Project Owned
140. Administrative Adjudication Bureau (AAB) New York State Department of Motor Vehicles	A. Halper	1975	Project Owned
141. Law Officer Project in the Family Court of New York City – An Evaluation	NYU	1973	NCJRS, Microfiche Project Owned
142. Philadelphia – Family Court – Research and Planning Unit, Refund Evaluation Report	D. Duffee & K. Wright	1975	NCJRS, Microfiche Project Owned
143. Philadelphia – Family Court – Research and Planning Unit – Update Evaluation	D. Duffee & K. Wright	1975	NCJRS, Microfiche Project Owned
144. "Pre-Trial Diversion/Intervention"	Minn. Gov. Comm.	1976	Project Owned

Title	Author(s)	Published	Source/Status
145. "Measurement Practice in Intensive & Special Adult Probation" Evaluation Quarterly, Vol. 2, No. 1.	Banks, Rardin	1978	Project Owned
146. New Haven (Ct.) Case Incident Regional Reporting System (CIRRS) Exemplary Project Validation Report	Abt Associates	1976	NCJRS, Microfiche Project Owned
147. St. Louis Court Improvement Project - Project Review and Evaluation Report	St. Louis Comm. on Crime & Law Enforcement	1975	NCJRS, Microfiche Project Owned
148. Philadelphia - Court of Common Pleas - Release on Recognizance Program	R.A. Wilson	1975	NCJRS, Microfiche Project Owned
149. Evaluation of the National Center for Prosecution Management, 1971 - 1973	Greenwood, et al.	1974	NCJRS Microfiche Project Owned
150. Seattle Pre-Sentence Counseling Program Exemplary Validation Report	Anon	1975	NCJRS, Microfiche Project Owned
151. Analysis of Minnesota's Criminal Justice System	S. Coleman	1976	NCJRS, Microfiche Project Owned
152. Los Angeles County - Sheriff's Department - Reserve Deputy Sherrif Program, Exemplary	Anon	1975	NCJRS, Microfiche Project Owned
153. Dayton - Police Department - Neighborhood Assistance Officer Program, Exemplary	Anon	1975	NCJRS, Microfiche Project Owned
154. "Philadelphia - Probation Department - Research and Development Unit - Final Evaluation Report"	J.C. Sternbach	1975	NCJRS, Microfiche Project Owned
155. "Philadelphia - Court of Common Pleas - Research and Program Development Unit"	J.C. Sternbach	1975	NCJRS, Microfiche Project Owned
156. "Criminal Justice Research Assistance Project - Final Report"	Zimmerman, et al.	1976	NCJRS, Microfiche Project Owned
157. "Denver - Intensive Probation and Parole Supervision Project - Final Report"	P.W. Hemingway	?	NCJRS, Microfiche Project Owned

Title	Author(s)	Published	Source/Status
158. "Lancaster County (Pa.) - Volunteers in Probation and Parole - Final Report and Evaluation"	D.W. Trexler	1975	NCJRS, Microfiche Project Owned
159. "Pretrial Release with Supporting Services for 'High Risk' Defendants - Three-Year Evaluation"	Brown, et al	1973	NCJRS, Microfiche Project Owned
160. "Evaluation of the Specialized Units Project of the Pennsylvania Board of Probation and Parole"	M.V. Lewis, et al	1974	NCJRS, Microfiche Project Owned
161. "Massachusetts Police Institution - Evaluation"	A.D. Little	1976	NCJRS Document Loan
162. "Attorneys As Friends Out of Court: An Evaluation of the 1974 California Parole Aide Program"	G. Feis, et al	1975	Project Owned
163. "Pre-Trial Release in Maryland - A Study of Maryland District Rule 777"	National Council on Crime & Delinquency	1974	Project Owned
164. "Cluster Evaluation of Narcotics Coordination Projects Including County-Wide Comp. Narc. Project"	R.E. Dizinno, et al	1974	NCJRS Document Loan
165. "Community Sponsors and Support Teams in Corrections - An Experiment and It's Evaluation	M.V. Lewis, et al	1974	NCJRS Document Loan
166. "Probation Officer Case Aide Project - Final Evaluation - Phase I"	D.W. Beless, E.R. Rest	?	NCJRS Document Loan
167. "Probation Officer Case Aide Project - Final Evaluation - Phase II"	E.R. Rest, G.J. Busiel	?	NCJRS Document Loan
168. "Orange County (CA) Evaluation Progress of the Alternative Routes Project - Following 19 Months of Development and Demonstration"	G.W. Carter, et al	1973	NCJRS Document Loan
169. "Parole Officer Aide Program in Ohio - An Exemplary Project"	R.R. Priestino, H.E. Allen	1975	NCJRS Document Loan
170. "Only Ex-Offenders Need Apply - Exemplary Project"	C.H. Blew, et al	1976	NCJRS Document Loan

Title	Author(s)	Published	Source/Status
171. "Report on Administration of the Program to Reduce Crime in Minnesota - Law Enforcement Assistance Administration - Minnesota Governor's Commission on Crime Prevention and Control"	U.S. General Accounting Office	1974	NCJRS Document Loan
172. "Voluntary Selection of Drug Treatment Program Project - Final Report"	CA Dept. of the Youth Authority	1973	NCJRS Document Loan
173. "Evaluation of the California Probation Subsidy Program"	L.L. Kuehn	1973	NCJRS Document Loan
174. "Evaluation of the Manhattan Criminal Court's Master Calandar Project - Phase I - February 1 - June 30, 1971"	J.B. Jennings	1972	NCJRS Document Loan
175. "Volunteer Parole Aide Program, 2 Year Evaluation, 1972-1974"	Lipstein and Stebbins	1975	Project Owned
176. "California - Office of Criminal Justice Planning - Strategic Evaluation Plan, Volume I"	J. Fisk	1975	NCJRS, Microfiche
177. "Volunteer Programs in Corrections: A Survey Report"	Amer. Bar Assoc., ABA Commission on Correctional Fac. and Services	1975	Project Owned
178. Uniform Parole Reports: A National Correctional Data Systems	M.G. Neithercutt, et al	?	Project Owned
179. "Criminal and Juvenile Courts in Mecklenburg County, North Carolina: Measurement and Analysis of Performance"	Clarke		Project Owned
180. "Analysis of the Night and Weekend Arraignment Parts in the Bronx and Queens Criminal Courts"	Jennings		Project Owned
181. "Scheduling Techniques for Municipal Court Traffic Sessions"	Fath		Project Owned

	Title	Author(s)	Published	Source/Status
182.	"Juror Waiting Time Reduction"	Pabst		Project Owned
183.	"National Volunteer Parole Aide Program Final Report and Evaluation – First Funding Period, 1971-1973"	Cochran		Project Owned
184.	"Norfolk Fellowship Foundation, Inc.: Massachusetts Correctional Institution, Norfolk"	Abt Assoc.		NCJRS Document Loan
185.	"Legal Services for Prisoners, Inc. – Topeka, Kansas"	Abt Assoc.		NCJRS Document Loan
186.	"Performance Measures Used in the Impact Program"	Siegel and Garse		NCJRS Document Loan
187.	High Impact Anti-Crime Program – National Level Evaluation – Final Report"	E. Chelimsky		NCJRS Document Loan
188.	"Marin County TASC Evaluation, 1975"	Pick, et al		NCJRS Document Loan
189.	"The New Haven Pretrial Diversion Program – A Preliminary Evaluation (May 16, 1972 – May 1, 1973"	Freed, et al		Private Loan
190.	"Pre-Trial Intervention: A Program Evaluation of Nine Manpower-Based Pre-Trial Intervention Projects Developed under the Manpower Administration, U.S. Dept. of Labor, Final Report"	Abt Assoc.		Project Owned
191.	"Santa Clara County Pretrial Release Program – Pretrial Release Program in an Urban Area"	Santa Clara County Pretrial Release Program		Project Owned
192.	"St. Louis High Impact Anti-Crime Program – Impact Evaluation Plan and Evaluation Progress Report"	MO Law Enforcement Assistance Council – Region 5		NCJRS Document Loan
193.	"Evaluation and Case Study of a School-Based Delinquency Prevention Program: The Minnesota Youth Advocate Program"	P.S. Higgins	1978	Project Owned

	Title	Author(s)	Published	Source/Status
194.	"Inmate Self-Government and Attitude Change"	J.L. Regens, W.B. Hobson	1978	Project Owned
195.	"Cost-Effectiveness of Residential Community Corrections: An Analytical Prototype"	C.M. Gray, et al	1978	Project Owned
196.	"Exemplary Project Screening and Validation Report: Southeastern Correctional Management Training Council, Athens, Georgia"	Abt Assoc.		NCJRS Document Loan
197.	"Exemplary Project Screening and Validation Report: Project Teletraining, County of Suffolk Police Dept., New York"	Abt Assoc.		NCJRS Document Loan
198.	"Exemplary Project Screening and Validation Report: Community Education on Law and Justice, Chicago, Illinois"	Abt Assoc.		NCJRS Document Loan
199.	"Phase I Evaluation of Pretrial Release Programs - Work Product Four - Assessment of the Present State of Knowledge Concerning Pretrial Release Programs, February 1976"	National Center for State Courts		NCJRS Document Loan
200.	"Knowledge Assessment - Phase I Evaluation of Intensive Special Probation Projects"	Banks, et al		Project Owned

*****   *****   *****   *****   *****   *****   *****

Note: Both Tim Eckels and Nancy Reichman have read the following twenty-five reports:
       (numbers taken from this appendix listing)

       7, 27, 30, 47, 53, 54, 60, 62, 69, 75, 76, 81

       82, 86, 89, 92, 126, 144, 149, 155, 160, 162, 172, 179, 180

An Empirical Study of Methods Used In Criminal Justice Evaluations
Final Phase II Checklist
June 12, 1978

1. READER & DATE: _____

2. TITLE: _____

   _____

   _____

   _____

3. SUBJECT & ID#: _____

   _____

4. AUTHOR & ORGANIZATION: _____

   _____

   _____

5. PUBLICATION DATE: _____

6. FUNDING ORGANIZATION, STATE PLANNING AGENCY, OR SPONSOR: _____

   _____

   _____

7. NCJRS # OR SALES AGENCY: _____

8. CHECK ONE (if applicable):   _____ Exemplary Project

                                   _____ Exemplary Validation Report

                                   _____ National Evaluation Program (NEP)

                                   _____ State Planning Agency

                                   _____ High Impact Anti-Crime Program

9. PERCENT (%) OF BUDGET ALLOCATED TO EVALUATION: _____

10. TOTAL FUNDING OF EVALUATION: _____

_____

_____

11. TIME ALLOCATED TO EVALUATION: _____

_____

_____

12. TIMING OF EVALUATION (e.g., before, during or after program implementation):
(a)

_____

_____

(b) PLANNING OF EVALUATION (before, during, or after program implementation):

_____

_____

13. NUMBER OF EVALUATION PERSONNEL AND BACKGROUNDS:
     Same as program personnel? What relationship with program?

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

14. DID THE EVALUATORS CONSIDER WHETHER PROGRAM GOALS WERE CLEARLY SPECIFIED?
(a)

_____

_____

_____

_____

(b) DO YOU FEEL THAT THEY WERE CLEARLY SPECIFIED? _____

_____

_____

_____

_____

15. DID THE EVALUATORS CONSIDER WHETHER THE PROGRAM WAS DIRECTED AT THE APPROPRIATE
(a) TARGET POPULATION?

_____

_____

_____

(b) DO YOU FEEL THAT IT WAS DIRECTED AT THE APPROPRIATE POPULATION?

_____

_____

_____

16. DID THE EVALUATORS CONSIDER WHETHER THE PROGRAM WAS IMPLEMENTED AS DESIGNED?
(a)

_____

_____

_____

(b) DO YOU FEEL THAT IT WAS IMPLEMENTED AS DESIGNED? _____

_____

_____

_____

17. FOCUS OF EVALUATION:

_____Narrative Case Study   _____Outcome Evaluation

_____Input Evaluation or Audit   ___experimental design

_____Process Evaluation (i.e.,   ___quasi-experimental design
     program monitoring)
                                                ___statistical models
     ___w/ performance measures
                                                ___formal models
     ___w/o performance measures

                                              _____Comprehensive Evaluation

     SUITABILITY OF EVALUATION FOCUS: 1   2   3   4   5   6   7
                                            least   most

     COMMENTS: _____

     _____

     _____

     _____

18. WERE THE MEASURES ADEQUATE?: _____

     _____

     _____

     _____

19. Evaluability--CAN OUTCOMES BE ATTRIBUTED DIRECTLY TO PROGRAM ACTIVITIES?
(a)

     _____

     _____

     _____

     _____

(b) Evaluability--IS THERE A THEORY LINKING PROGRAM ACTIVITIES TO THE PERFORMANCE
        MEASURES CHOSEN?

     _____

     _____

     _____

20. ARE PROGRAM ACTIVITIES CLEARLY DESCRIBED IN THE EVALUATION? _____

_____

_____

_____

_____

_____

_____

21. Describe feedback between program staff and evaluation staff. If policy
     changes or unexpected results cause a need for changes in the evaluation
     design, was the design flexible enough to account for this?

_____

_____

_____

_____

_____

_____

_____

_____

22. Was the research designed to yield information that would be useful in a
     broader context than just evaluating this particular program?

_____

_____

_____

_____

_____

23. DATA SOURCES:

___Informal Material   ___Observational Data

___Administrative Records   ___Questionnaire Data

___already available   ___Interview Data

___gathered esp. for the eval.   ___Written Documents

SUITABILITY OF DATA SOURCES: 1   2   3   4   5   6   7
                                   least   most

COMMENTS: _____

_____

_____

_____

_____

_____

24. METHODOLOGY:

___Qualitative Analysis   ___Regression and ANOVA

___Participant Observation   ___Time Series Analysis

___Content Analysis (or other   ___Factor Analysis
   analysis of written materials)

___Descriptive Statistics   ___Formal Models

___Statistical Inference   ___Other (specify) _____

                                        _____

SUITABILITY OF METHODOLOGY: 1   2   3   4   5   6   7
                                   least   most

COMMENTS: _____

_____

_____

_____

_____

25. IS THE DATA PRESENTATION ADEQUATE? _____

_____

_____

26. ARE THE METHODS OF ANALYSIS CLEARLY DOCUMENTED? _____

_____

_____

27. ARE THE CONCLUSIONS SUPPORTED BY THE DATA ANALYSIS? _____

_____

_____

28. IS THERE SOME DISCUSSION OF PROGRAM IMPLEMENTATION AND POLICY IMPLICATIONS?

_____

_____

_____

29. IS THERE AN INDICATION THAT THE EVALUATION FINDINGS INCLUENCED
(a) ACTUAL DECISION-MAKING ?

_____

_____

(b) COULD THIS TYPE OF EVALUATION BE ADAPTED FOR USE IN PERIODIC REVIEW? _____

_____

_____

_____

30. WOULD THIS REPORT BE ACCESSIBLE TO THE AUDIENCES WHO WOULD FIND IT USEFUL?

_____

_____

_____

_____

_____

31. IS THERE A DISCUSSION OF PROBLEMS MET IN IMPLEMENTING THE EVALUATION PLAN?

_____

_____

_____

_____

_____

FINAL COMMENTS: _____

_____

_____

_____

_____

_____

_____

_____

_____

_____

COMPUTER CODE LISTS

The total system for identifying evaluation reports used in our study consists of a 12-digit number, representing 8 different components. A 6-digit (ID) is assigned to each report and its questionnaire. The first two digits denote the subject code of the report; the second two are the entry number in a subject code, representing the agency or organization that publishes the report. This completes the main ID number. The next single digit indicates the report belongs to either the National Evaluation, or a High Impact Anti-crime program, or a State Planning Agency program. Following this, a single digit number is used to denote Exemplary Projects or Exemplary Validation Program reports. The sixth component in the system indicates the focus of the evaluation, the seventh indicates the source of data used, and the eighth represents the methodology used. A sample printout is as follows:

```
Enter a request.
--> list 50-00-05 to 81-00-05
```

Report ID	SUBJ	ENT NO	AGENCY	NEP	EXEMP	FOCUS	DATA	METH1	METH2
50-01-70	50	01	70			3	7	1	4
50-02-99	50	02	99			7	2	4	5
50-03-99	50	03	99			6	1	3	5
50-04-61	50	04	61	0	3	7	6	4	

COMPUTER CODE LIST

SUBJECTS

Code	Subject
00	GENERAL REFERENCE
01	EVALUATION MANUALS: GENERAL
02	EVALUATION MANUALS: SPECIFIC
03	METHODOLOGY: QUALITATIVE
04	METHODOLOGY: EXP. DESIGN
05	METHODOLOGY: TIME SERIES
06	METHODOLOGY: STATISTICS
07	METHODOLOGY: MODELS
08	METHODOLOGY: OTHER
10	TECHNICAL ASSISTANCE PROGRAMS
11	IN-HOUSE R&D OR PLANNING UNITS
12	TRAINING UNITS
13	OTHER SPECIALIZED ADMINISTRATIVE UNITS
20	CORRECTIONS: INCARCERATION
21	CORRECTIONS: PROBATION & PAROLE
22	CORRECTIONS: JUVENILE DIVERSION
23	CORRECTIONS: PROGRAMS FOR DRUG ABUSERS
24	CORRECTIONS: OTHER ALTERNATIVES TO INCARCERATION
30	POLICE COMMUNITY RELATIONS: GENERAL
31	POLICE COMMUNITY RELATIONS: STREET SAFETY/CRIME REDUCTION
32	POLICE/SCHOOL RELATIONS PROGRAMS
33	POLICE/COMMUNITY TRAINING

40	PRE-TRIAL RELEASE
41	JURY MANAGEMENT SELECTION
42	COURT SCHEDULING
50	PATROL STRATEGIES: PATROL INTENSITY
51	PATROL STRATEGIES: ONE/TWO OFFICER
52	PATROL STRATEGIES: DISPATCHING
53	PATROL STRATEGIES: SECTOR DESIGN
54	PATROL STRATEGIES: RESPONSE TIME
55	PATROL STRATEGIES: OTHER
60	TEAM POLICING: COMMUNITY RELATIONS
61	TEAM POLICING: DECENTRALIZED MANAGEMENT
70	AVM SYSTEMS
71	CAD (COMPUTER AIDED DISPATCH)
80	INFORMATION SYSTEMS: STATISTICAL
81	INFORMATION SYSTEMS: GEOGRAPHICAL
90	RESOURCE ALLOCATION: COST
91	RESOURCE ALLOCATION: PRODUCTIVITY
99	MISCELLANEOUS

AGENCIES

01	NEW JERSEY SPA
02	MINNESOTA SPA
03	CALIFORNIA SPA
04	CONNECTICUT SPA
05	PENNSYLVANIA SPA
06	MICHIGAN SPA
07	HAWAII SPA
08	OREGON SPA
09	VIRGINIA SPA
10	NORTH CAROLINA SPA
11	GEORGIA SPA
12	NEW YORK SPA
13	MISSOURI SPA
20	WASHINGTON CRIMINAL JUSTICE EDUCATION & TRAINING CENTER
50	URBAN INSTITUTE
51	ABT
52	LAW ENFORCEMENT ASSISTANCE ADMINISTRATION
53	NATIONAL COUNCIL ON CRIME & DELINQUENCY
54	THE RAND CORPORATION
55	NATIONAL SCIENCE FOUNDATION
56	SYSTEM DEVELOPMENT CORPORATION
57	LOS ANGELES POLICE DEPARTMENT

AGENCIES (continued)

58   KANSAS CITY POLICE DEPARTMENT

59   SEATTLE POLICE DEPARTMENT

60   NEW YORK CITY RAND

61   MITRE

62   AMERICAN BAR ASSOCIATION

69   PUBLIC SYSTEMS EVALUATION

70   POLICE FOUNDATION

99   MISCELLANEOUS

Reader & Date: Cheryl, June 21
NCJRS Number:
Report Title: Evaluation of the Massachusetts Police Institu
Author & Organization: Arthur D. Little
Publication Date:
Funding Organization: Massachusetts Committee on Criminal Jutsice
Abstract:

End of Report.
SUBJ 3
ENT NO 12
AGENCY 99
NEP 0
EXEMP 0
FOCUS 9
DATA 3
METH1 0
METH2
METH3
An Empirical Study of Methods used in Criminal Justice Evaluations.
8.
9. not answered.
10. not answered.
11. not answered.
12(a). during
12(b). during
13.

14(a). no
COMMENT:
Not explicitly, though certainly they spelled them out in great
detail.

14(b). yes

15(a). yes
COMMENT:
Evaluators questioned whethered the program should serve smaller
police departments primarily and whether they in fact did so.

15(b). yes

16(a). yes
COMMENT:
They considered in their discussion of program implementation, in
specific cases they reviewed. Evaluators did not directly
address this issue.

16(b), yes
COMMENT:
It seems to be. There is not enough discussion of overall trends
in the program to be able to tell easily.

17. Comprehensive Evaluation
SUITABILITY OF EVALUATION FOCUS: 6
COMMENT:
Evaluators had both a process evaluation and an outcome
evaluation in the form of a quasi-experimental design. There was
an interesting use of a quasi-experimental design to do a
qualitative, process-oriented study of a large program containing
many projects. They used a randomized, stratified sampling for
the experimental group. It was a quasi-experimental design
because the control and comparison groups were chosen form a
different population. They could not do a strict experimental
design because the evaluation was done after the inception of the
program. However, the groups were very carefully chosen and
randomization was employed whenever possible. They chose as the
population for their control group those police departments who
had requested assistance from MPI but had not yet recieved it.
While they had per- formance measures, these were specific to
particular parts of the overall program-- certain tangible
criteria the evaluators felt to be measures of the program's
overall success.

18. yes
COMMENT:
On the whole quite good. Evaluators were careful to consider
that satisfaction of clients only a partial measure of the
program's success and that impact in behavior, concrete outcomes,
were the most important measures.

19(a). yes
COMMENT:
Evaluators were careful to look at what activities resulted in
what outcomes.

19(b). yes

20. yes
COMMENT:
Not simply a general overview of a typical program activity but
actual historical accounting of how activities unfolded in
different projects.

21(a). yes
COMMENT:
Yes in sense that program staff objected to or questioned certain
feedback which the staff then elaborated as a result. This
implies, if not continuous feedback, some interval feedback
before the final report.

21(b). yes
COMMENT:
Design flexible because interview questions, the major data
source, were flexible and designed to elicit program-specific
responses. Also apparant in the ingenuity with which evaluators
constructed experimental and control groups.

22. no
COMMENT:
Not really testing a hypothesis. However, the program they are
evaluating is a very large one which has many smaller pro- jects
under it.

23. Observational Data
SUITABILITY OF DATA SOURCES: 6
COMMENT:
Rather informal telephone interviews but structed in that there
is a definite set of questions as a base. The basic questions,
give
 in the appendix, seem quite thoughtful-- investigative reporting
style which is important to unearth the police cheif's
perceptions of what really happened. While there is no direct
observation, the interviews are not simply standardized surveys
but are designed to get a rich, situation-specific response.
While main data source was interview data, evaluators used
administrative records for a determination of the organization's
cost-effectiveness.

24. Qualitative Analysis
SUITABILITY OF METHODOLOGY: 6
COMMENT:
Notes about this in final comments.

25. yes
COMMENT:
They give examples of responses as well as evaluator's
interpretation and aggregation.

26. no
COMMENT:
Telephone interviews are stated as the only form of data
gathering but probably information was gathered from MPI files
and perhaps some observation of on-going programs as well. (This
should got under question 25.)

27. yes

28. yes
COMMENT:
They particularly listed those which potential clients said they
felt would be useful.

29(a). no

29(b). yes
COMMENT:
Rather efficient because use of telephone interviewing and
process oriented rather than oriented solely to measuring impact
on long-term goals.

30. yes
COMMENT:
Executive summary. First sections of report confusing but
overall fairly readable and results informative.

31. no

FINAL COMMENT:
This is one of the most interesting evaluations I have seen.   It
combines features of a process evaluation --- while done on
aspects of the program completed before the evaluation took place
--- and features of an experimental design which allows you to
look carefully at a small sample in a systematic way and feel
relatively confident your findings can be generalized to the
larger population you are actually interested in. The evaluators
were faced with looking at a multi- faceted program which had
been in operation long before they arrived on the scene. They
could not examine, with any depth, all of the things the program
had done yet they wanted to evaluate the program overall.   Their
solution was quite creative. They carefully shose a stratified
random sample of police departments who had made use of the MPI.
These were matched, randomizing where possible, with a control
group of clients awaiting services from MPI and a comparison
group of police departments who had never made requests. (Why
they chose the former as a control group is unclear and a bit
problematic to me.) They then did structured telephone
interviewing of subjects in various groups. Where "treatment"
had been given, they probed not only for the attitudes of
subjects but for documentation as to what had actually changed in
their depart- ments as a result of the treatment. While this
"investigative reporting" method of data collection does not
substitute for direct observation, the interviewing was situation
specific enough (as revealed by reported subject responses) to go
beyond the initial subjective responses of the clients. One
problem which casts some doubt on the evaluators interpretations
of data is that they did sound a bit too anxious to praise the
projet staff where they could.

End of Questionaire.

END