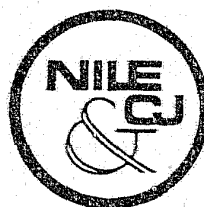




# HIGH IMPACT ANTI-CRIME PROGRAM

An Assessment of Project-Level Evaluation Reporting



36126

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

# NATIONAL IMPACT PROGRAM EVALUATION

An Assessment of Project-Level Evaluation Reporting

By

M. B. FISCHER

G. W. KUPERSMITH

A. P. TREHAN

THE MITRE CORPORATION

NCJRS

AUG 26 1975

OCTOBER 1975

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

This document has been prepared by The MITRE Corporation,  
Washington Operations, under Contract J-LEAA-028-75 for  
the Law Enforcement Assistance Administration.

# ABSTRACT

This document presents an assessment of project-level evaluation reporting in the High Impact Anti-Crime Program; the reports are examined in terms of their number, type and quality. The assessment criteria used to define evaluation reporting quality and evaluation approach applicability are presented. Results of the assessment process are presented for the program as a whole, for different types of projects, and for the individual Impact cities. This document was prepared by The MITRE Corporation in conjunction with the National Institute of Law Enforcement and Criminal Justice as part of the national-level evaluation of the High Impact Anti-Crime Program.

MITRE Department  
and Project Approval:

*Chen Chelini*

# TABLE OF CONTENTS

	<u>Page</u>
LIST OF ILLUSTRATIONS	vii
EXECUTIVE SUMMARY	ix
1.0 INTRODUCTION	1
2.0 EVALUATION IN THE IMPACT PROGRAM	3
3.0 ASSESSING IMPACT PROJECT EVALUATION EFFORTS	5
3.1 Focus of the Document	5
3.2 Assessment Approach	5
3.2.1 Evaluation Reporting Quality	5
3.2.2 Evaluation Approach Applicability	11
3.3 Data Collection	17
3.4 Analysis Approach	18
4.0 OVERVIEW	20
4.1 Review Coverage	20
4.2 Type of Report	21
4.3 Reliability of Reported Findings	25
4.4 Evaluation Reporting	28
4.5 Evaluation Approach Applicability	33
5.0 ANALYSIS RESULTS	38
5.1 Project Focus	38
5.1.1 Review Coverage	39
5.1.2 Type of Reports Reviewed	41
5.1.3 Evaluation Reporting Quality	43
5.1.4 Evaluation Approach Applicability	43
5.2 City Analysis	47
5.2.1 Review Coverage	47
5.2.2 Type of Reviewed Reports	51
5.2.3 Evaluation Reporting Quality	54
5.2.4 Approach Applicability	57
6.0 ANALYSIS OF STAGES OF THE EVALUATION PROCESS	62
7.0 ADDITIONAL ASPECTS OF EVALUATION REPORTING	69
7.1 Presence of Limitations or Qualification of Findings	69
7.2 Description of Reported Limitations	70
7.3 Operational Problems and Recommendations	78
7.3.1 Reporting of Operational Problems	78
7.3.2 Description of Reported Problems	79
7.3.3 Reporting of Operational Recommendations	81
8.0 SUMMARY AND CONCLUSIONS: PROJECT-LEVEL EVALUATION IN THE IMPACT PROGRAM	86

## TABLE OF CONTENTS (CONCLUDED)

	<u>Page</u>
APPENDIX I PROJECT-LEVEL EVALUATION TECHNICAL REVIEW DATA COLLECTION FORM	91
APPENDIX II DESCRIPTION OF ALLOCATION OF IMPACT PROJECT- LEVEL EVALUATION RESPONSIBILITY	105

## LIST OF ILLUSTRATIONS

<u>Table Number</u>		<u>Page</u>
I	Assessing Evaluation Reporting	6
II	Evaluation Reporting Typology	9
III	Criteria Required for Evaluation Reporting Quality Classifications	10
IV	Assessing the Applicability of an Evaluation Strategy	13
V	Evaluation Approach Typology	15
VI	Criteria Required for Evaluation Approach Applicability Classifications	16
VII	Operational Period of Projects for which Reports were not Received and Reviewed	22
VIII	Distribution of Projects by Evaluation Reporting Quality	29
IX	Distribution of Project Documentation by Evaluation Reporting Criterion	31
X	Distribution of Approach Applicability of Projects	34
XI	Distribution of Evaluation Approach Types	37
XII	Distribution of Technically Reviewed Projects by Project Focus	40
XIII	Type of Report Reviewed by Project Focus	41
XIV	Evaluation Reporting Quality by Project Focus	44
XV	Evaluation Approach Applicability by Project Focus	45
XVI	Proportion of Projects Documented and Reviewed (by Impact City)	48
XVII	Type of Evaluation Documentation Reviewed for each Impact City	52
XVIII	Evaluation Reporting Quality by Impact City	55
XIX	Evaluation Approach Applicability by Impact City	58
XX	Comparison of Expected and Actual Mean Approach Applicability Scores by Impact City	60
XXI	Evaluation Reporting Quality by Evaluation Approach Applicability	63
XXII	Evaluation Planning Quality by Evaluation Approach Applicability	66
XXIII	Evaluation Planning Quality and Evaluation Reporting Quality	68
XXIV	Presence of Limitations in Impact Evaluation by City Documentation	72
XXV	Breakdown of Reported Limitations	74
XXVI	Operational Problems Reported	80
XXVII	Presence of Recommendations in Reports of Assessed Projects by City	83
XXVIII	Presence of Recommendations in Project Reports by Project Focus	84

# LIST OF ILLUSTRATIONS

<u>Figure Number</u>		<u>Page</u>
1	Distribution of Reviewed Project Documentation by Report Type	24
2	Justification of Project Evaluation Findings	26

## EXECUTIVE SUMMARY

As part of the national-level evaluation of the Law Enforcement Assistance Administration's High Impact Anti-Crime Program, project evaluation documentation was technically reviewed for 119 Impact-funded projects. These documents were examined to assess the quantity, type and quality of project-level evaluation conducted in the Impact program.

Specific areas addressed in the document include:

- the criteria used to define and assess evaluation reporting quality and approach applicability;
- the overall character of Impact program project evaluation efforts;
- the proportion of Impact projects documented and reviewed;
- the type of documents reviewed;
- the evaluation reporting quality of reviewed Impact project evaluation documents;
- the applicability of evaluation approaches used to identify observed changes in targeted crime problems;
- the relationships observed among evaluation planning quality, evaluation reporting quality and evaluation approach applicability;
- the number and type of evaluation limitations cited in reviewed documentation; and
- the frequency and nature of operational problems and recommendations included in the reviewed project evaluation documentation.

For each aspect of the evaluation process listed above, analysis results are presented for the program as a whole, for each Impact city, and for each of the three types of projects implemented in the program (e.g., crime-reduction, recidivism-reduction and systems/other).



Important findings based on this analysis include:

(a) Review Coverage

To obtain a general sense of the extent to which project evaluation responsibilities were fulfilled, an examination was made of the proportion of Impact-funded projects for which evaluation documentation was received. These proportions were viewed as indicative of the review coverage achieved for each Impact city and for each focus category. It was found that:

- of the 233 Impact-funded projects, 140 or 60 percent had forwarded at least one evaluation report as of 1 July 1975. Of these 140 projects, documentation for 119 projects (51 percent of the 233) was considered suitable to be subjected to technical review for use in this assessment;
- assessment of project-level evaluation could not be conducted for almost half of the funded Impact projects, thereby limiting the amount of information available for use in this document (as well as in other efforts designed to assess the Impact program);
- of the 93 projects for which no evaluation documentation was available for use in this assessment, only 24 percent were recently funded (less than one year);
- the proportion of projects documented did not vary by the focus of the project (e.g., crime-reduction, recidivism-reduction, systems/other); approximately 50 percent of the projects funded for each focus had been documented and have been reviewed;
- Impact cities varied widely in terms of review coverage; Cleveland and Baltimore had the widest coverage (approximately 75 percent of all city projects) while Newark had the lowest (33 percent);
- recent project fundings in Atlanta and Denver (approximately 40 percent of the projects for which evaluation documentation was not available have been funded less than one year) may partially account for documentation delays.

(b) Type of Report Reviewed

To capture differences and similarities in the amount and type of activity and outcome information presented in reviewed evaluation documentation, documents were classified as either status, progress, preliminary evaluation or full-fledged evaluation reports. Use of this classification scheme provided the following findings:

- 57 percent of the 119 projects reviewed posed and answered, in relative detail, questions regarding project outcomes (full-fledged evaluation reports); an additional 15 percent of the projects were the subject of preliminary evaluation reports providing only cursory information on project outcomes. Thus, 72 percent of the 119 (85) reviewed projects provided documentation containing at least some information regarding project outcomes;
- documentation reviewed for a majority (70 percent) of crime-reduction focused projects were considered to be full-fledged evaluation reports; a greater proportion of status and progress reports were received and reviewed for recidivism-reduction and systems/other projects (26 percent and 46 percent respectively compared to 9 percent for crime-reduction projects);
- for four cities (Atlanta, Denver, Portland and St. Louis) more than 75 percent of the projects subject to review provided documentation considered to be full-fledged evaluation reports. No full-fledged reports were received from Newark;
- Portland and Atlanta's high proportion of full-fledged evaluations are noteworthy as both cities placed the primary responsibility for outcome evaluation with organizations not directly involved with the day-to-day concerns of Impact projects.

(c) Evaluation Reporting Quality

The manner in which evaluative information was conveyed in project documentation was examined in terms of the following set of criteria: readability, presentation of activity information, specification of limitations, inclusion of data for face validation. Based on the extent to which these criteria were met, evaluation documentation was classified and analyzed according to the quality of the project evaluation reporting effort. The findings were that among the 119 projects reviewed, variations in evaluating reporting quality existed.

Specifically:

- only 42 of the 119 (35%) projects reviewed provided evaluation documentation with enough information to permit a meaningful interpretation of results;
- the evaluation reporting quality of documentation reviewed for projects of differing foci (i.e., crime-reduction, recidivism-reduction, and systems/other) did not vary. Mean reporting quality scores calculated for projects in each of the three foci are almost identical - clustering around 2.2 (out of a possible 4.0), indicating the general descriptive rather than explanatory orientation of documents reviewed;
- evaluation reporting quality varied among the eight Impact cities:
  - Portland and Denver are noteworthy for their good evaluation reporting with mean reporting quality scores of 3.5 and 2.9 respectively,
  - project evaluation documentation reviewed from Atlanta, Dallas, and St. Louis (with reporting quality scores 2.5, 2.4, and 2.2 respectively) were of reasonably good quality. In each of these cities, over 40 percent of the projects for which documents were reviewed included limitations important to a fair interpretation of reported findings, and
  - the evaluation reporting efforts of Baltimore, Cleveland and Newark were viewed as being the least impressive with mean reporting quality scores of 2.1, 1.8, and 1.6 respectively;
- documentation from 80 percent of the reviewed projects provided the reader with a description of the projects' activities. The extent of activity information provided, however, would probably not suffice for use by those interested in project replication or transfer;
- data needed for face-validation of the findings presented were provided in documentation for only 40 percent (47) of the 119 reviewed projects.

(d) Evaluation Approach Applicability

Variations existed in the applicability of evaluation approaches used to gauge project outcomes. These evaluation approaches were assessed by examining the extent to which selected approaches permitted the identification of changes in the targeted crime problem and the attribution of such changes to project activities. Analysis results led to the following findings:

- over half (64) of the 119 projects reviewed were documented in the absence of an evaluation approach;
- only 14 percent (17) of the projects reviewed used what was considered to be a rigorous evaluation approach to assess project outcomes;
- of the 85 projects explicitly reporting project findings, such findings were viewed by MITRE staff to be justified and substantiated 50 percent of the time, regardless of the nature of these findings;
- of the 55 projects using at least some type of evaluation approach, 42 percent relied solely upon a before/after design; 29 percent combined the key aspects of the before/after approach with some type of comparison base;
- variations in evaluation approach were observed among projects of differing foci. Specifically:
  - crime-reduction projects were most highly assessed with a mean approach applicability score of 1.87 (out of a possible 3.0). This compares to a mean of 1.63 for recidivism-reduction projects, and a mean of 1.38 for systems/other projects;
  - crime-reduction projects were more likely to be evaluated using at least some type of evaluation approach than recidivism-reduction projects; both fared better overall than systems/other projects; and
  - rigorous evaluation approaches were equally prevalent among projects of all three foci;
- substantial variations in evaluation approach applicability were observed among the eight Impact cities; these variations override project focus considerations;



- three cities are noteworthy for the use of rigorous evaluation approaches: Portland, Atlanta, and Denver,
- projects reviewed from Baltimore and Cleveland consistently relied upon evaluation approaches that were not rigorous, and
- all projects reviewed from Newark were documented without the use of an evaluation approach.

(e) Stages in Evaluation Process

Interrelationships among evaluation planning quality, evaluation reporting quality and evaluation approach applicability were examined for the 119 documented projects. The findings include the following:

- evaluation reporting quality and evaluation approach applicability are related to one another in a positive, symmetrical fashion;
- both evaluation reporting quality and evaluation approach applicability appear to be positively associated with more comprehensive evaluation planning;
- however, good evaluation reporting and the use of rigorous evaluation approaches are possible, nonetheless, for those projects with minimal or missing evaluation plans.

(f) Reporting of Evaluation Limitations

To gauge the extent to which Impact evaluators sought to encourage the proper interpretation and application of reported findings, the existence and type of limitations cited in reviewed documentation were recorded and analyzed. The results include the following:

- of the 85 Impact projects providing findings, 87 percent tempered these findings by citing limitations in the interpretation of findings;
- all of the crime-reduction projects with findings specified limitations regarding their interpretation; 89 percent of the systems/other projects provided such an interpretive context, in contrast to 80 percent of the recidivism-reduction projects;
- Impact cities did not differ significantly in the reporting of limitations. More than 80 percent of the reviewed

projects for each city which provided findings included within their evaluation documentation explicit limitations with respect to the interpretation of these findings;

- data constraints (i.e., unavailability, limited quantity and reliability) were the most frequently reported evaluative limitations, accounting for 54 percent of the 208 limitations recorded;
- design approach problems (e.g., lack of comparability among control group/area; seasonality, etc.) were the second most frequently reported limitation, accounting for 33 percent of the 208 recorded;
- variations in the type of limitations reported were observed among projects of differing foci:
  - design problems (limitations) were most prevalent among crime-reduction projects (53 percent) and can be largely explained by the fact that projects reviewed of this focus frequently (33 percent) tempered findings with the admission that attribution of changes to project activities was not possible or within the scope of project evaluators, and
  - data problems predominated among recidivism-reduction (51 percent) and systems/other projects (54 percent) reflecting the dependency of projects of these foci upon other than established data sources.

(g) Reporting of Operational Problems and Recommendations

Project evaluation documents were reviewed to assess the extent to which they included statements concerning problems encountered in project operations and recommendations for improvement. Findings indicated that:

- operational problems were cited in the evaluation documentation of 79 percent (94) of the 119 projects reviewed;
- two cities - Portland and Baltimore - reported operational recommendations significantly less often than other cities;
- the organizational responsibility for project-level evaluation seemed to have no effect upon the extent to which evaluation documentation communicated operational problems;

- personnel problems (i.e., staffing and training) were the most frequently mentioned difficulty cited in the reviewed documentation accounting for 35 percent of all problems cited;
- problems concerning the development of lines of communication essential to project operations (i.e., establishing client referral sources, inter-agency cooperation, community support) was the second most frequent difficulty mentioned in the reviewed documentation accounting for 23 percent of all problems cited;
- problems regarding project funding were rarely (3 percent) cited in the reviewed project documentation;
- recommendations for the improvement of project operations were cited in the evaluation documentation of 55 percent of the 119 reviewed projects;
- project-level evaluations written by project staff did present recommendations more frequently than those prepared by other agencies;
- recommendations reported in the reviewed project evaluation documentation typically corresponded to reported operational problems or consisted solely of general statements urging project refunding; and
- recommendations reported by the reviewed projects generally were not logical extensions of evaluation results nor specific enough to adequately inform those responsible for resolving project difficulties.

Based on the above findings it is clear that the quantity, type and quality of Impact program evaluation efforts varies considerably across cities. The heterogeneity observed suggests the need to provide guidelines for the staffing, conduct and organization of project evaluation efforts if uniformity, and thus comparability of results, is to be realistically achieved. Additionally, the results suggest that the differences in the ease with which different types of projects may be evaluated, rests, to some extent, on the availability and accessibility of established data sources for use in the evaluation effort.

## 1.0 INTRODUCTION

Evaluation has become a slogan for policymakers in the 1970's. Its popularity stems, at least in part, from the recognition that the expenditure of funds does not automatically insure viable solutions to pressing social problems. That is, the large influx of federal monies into the public arena has not been accompanied by decreases in crime or poverty, for example. Faced with a multiplicity of possible problem solutions and decreasing resources, policymakers have begun to look toward evaluation as a means for obtaining objective information about the potential success of various social programs and projects.

Calls for evaluation have frequently been made, however, without adequate consideration of the true costs of sound evaluation. Obtaining reliable, objective information requires a level of coordination, planning and technical expertise which frequently transcends the resources, capabilities, and interests of program funding and implementing agencies. Implementing agencies, faced with limited resources for service provision, are understandably reluctant to allocate sufficient funds for evaluation. Program designers/funders and operating personnel are frequently in conflict over the priority to be accorded to evaluation efforts and to the type of evaluative information which is to be produced. The former are typically interested in long-term evaluation results, particularly as they compare to the results of other approaches to the same problem, while the latter seek to gather and analyze other data viewed as more pertinent to their decision-making needs. Additionally, rigorous evaluation entails a reliance upon skills and expertise often unavailable among program planning staffs and project operators. As a result, extra personnel or technical assistance may be required, further escalating evaluation costs and coordination problems.

In addition to these constraints to evaluation, projects vary considerably in their amenability to rigorous evaluation. Problems in operationally defining and measuring desired outcomes such as "the quality of justice", "health", or "the equality of educational opportunity" make it difficult if not impossible to assess the success of some projects in precise, quantifiable terms. Similarly, when services are provided to a diverse, unbounded population or area, assembling information on the effects a project may have had on such groups or areas becomes exceedingly difficult. Other obstacles to evaluation, such as privacy considerations and constraints on the selection of treatment subjects (which are inherent in particular types of projects) add to the difficulties. Suffice it to say, then, that the difficulty with which a program or project may be evaluated is an important factor frequently ignored amidst cries for evaluation.

Given these constraints and obstacles to evaluation, what can realistically be expected from large-scale evaluation efforts? What are the ingredients essential for translating the idea of evaluation into a sound, thoughtful effort? And finally, what are some of the limitations and problems which may be expected to impact efforts to evaluate short-term action projects?

Answering these questions is a driving force behind this document. Through a description and assessment of the project-level evaluation activities and experience of eight cities engaged in a large-scale anti-crime program, this document seeks to provide a basis for a better understanding of what can realistically be expected from project evaluation in the area of criminal justice.

## 2.0 EVALUATION IN THE IMPACT PROGRAM

The High Impact Anti-Crime Program is an LEAA effort designed to assist eight large cities<sup>1</sup> in reducing the incidence of stranger-to-stranger crime and burglary in their local jurisdictions. Developed in response to the rising crime rates reported in the late 1960's, Impact allocated about \$20 million in discretionary funds to each of the eight cities, encouraging them to produce comprehensive, concentrated approaches to crime within broad Impact policy guidelines.

Guidelines surrounding the use of Impact monies reflect the basic concepts and innovations of the program. Awards were to be contingent upon the preparation of master plans detailing the specific crime problems faced in a given city, its problem-solving priorities, and the anti-crime strategies and tactics which would be brought to bear on them. Additionally, each city was expected to prepare plans for evaluating not only the overall program implemented in the city, but the specific anti-crime projects which comprised the overall city program. Taken together, these plans contained the essential elements of a key feature of the Impact program--the development and application of a model for comprehensive Crime-Oriented Planning, Implementation and Evaluation (COPIE).

Translating the components of this model into action was the responsibility of newly established city-based Crime Analysis Teams (CATs). These teams, comprised of researchers, criminal justice planners and functional area specialists, were charged with conducting or overseeing the planning, implementation and evaluation activities dictated by the COPIE-cycle model. While the actual substance and conduct of these activities varied from city to city, the essential thrust of the

---

<sup>1</sup>Atlanta, Baltimore, Cleveland, Dallas, Denver, Newark, Portland and St. Louis

program - reduction of serious crime in 8 urban areas - has been maintained within and through these program features since Impact's inception in January 1972.

Impact's effects, both as an anti-crime agent and a vehicle for the introduction of the COPIE-cycle, are currently being examined from several different perspectives. These perspectives, corresponding to different levels of evaluation, run the gamut from localized city-initiated evaluations of individual anti-crime projects (project-level evaluations) to a large, national-level examination of overall program strengths and weaknesses. As part of the national-level evaluation of the Impact program, this document provides the results of an assessment of the evaluations performed of the individual anti-crime projects-- project-level evaluation. Succeeding sections will present a description and assessment of the number, type and quality of project-level evaluations, conducted by (or overseen by) the Crime Analysis Teams in each of the 8 Impact cities.

### 3.0 ASSESSING IMPACT PROJECT EVALUATION EFFORTS

#### 3.1 Focus of the Document

The present assessment of Impact project-level evaluation efforts is based upon an examination of the performance of the eight Impact cities along two dimensions of the evaluative process. The first dimension to be addressed is evaluation reporting. That is, the written product used by evaluators to transmit information about the progress a project is making in solving the crime problem it targets and in providing planned activities. The second dimension of the evaluative process to be examined relates to the overall strategy for gathering this information. That is, the approach used to measure changes in the targeted crime problem and to estimate the extent to which such observed changes may be attributable to project activities. Taken together, these two dimensions provide a focal point for determining the extent to which project-level evaluations performed in the Impact program were viable mechanisms for gathering objective information for use by decision-makers.

#### 3.2 Assessment Approach

Assessing Impact project-level evaluation efforts entailed reviewing over 200 evaluation documents produced and disseminated by the Impact cities as of 1 July 1975. These documents, addressing 119 Impact projects, were reviewed by three members of the MITRE technical staff according to the set of criteria discussed below.

##### 3.2.1 Evaluation Reporting Quality

As discussed earlier (see page 1), the need for quality evaluation reporting revolves around the belief that sound evaluation yields information which may be used for decision-making purposes. With this in mind, the present assessment identifies a number of criteria as central to a successful evaluation reporting effort. Table I presents these criteria, the rationale for their selection, and the corresponding operational definitions which were used in the review process.

TABLE I  
ASSESSING EVALUATION REPORTING

CRITERIA	RATIONALE	OPERATIONAL DEFINITION
• Readability	People have limited time to devote to information acquisition. If the information presented in the report is not readily accessible and understandable, there is little chance such information will be extracted and used as a basis for operational or policy decisions.	Evaluation report should: <ul style="list-style-type: none"> <li>• include a brief description of what project is all about;</li> <li>• summarize the major findings in terms of objectives;</li> <li>• tie conclusions to the reported findings;</li> <li>• provide a lucid description of the method used to assess objective attainment; and</li> <li>• present a detailed data analysis on an objective-by-objective basis which logically integrates data/tables into text.</li> </ul>
• Activity Information Presented	A sense of activities (treatment) is needed to draw linkages between observed outcomes and actual project operations.	Present information regarding: <ul style="list-style-type: none"> <li>• target population/target area;</li> <li>• kind of services delivered; and</li> <li>• amount of services delivered.</li> </ul>
• Limitations of Approach Specified	A statement of limitations is essential to provide the reader with a context in which to interpret analysis results--there is a need to understand what analysis results really mean in terms of the project's time frame, target area/population, and activities to use information appropriately. Such information allows the reader to identify possible alternative explanations for observed changes and assess their plausibility.	Results placed within context of: <ul style="list-style-type: none"> <li>• project activities;</li> <li>• time frame of project operations;</li> <li>• target area/population;</li> <li>• methodological problems inherent in design approach used; and</li> <li>• evaluative problems encountered (data commensurability, inappropriate selection of comparison area/group).</li> </ul>
• Data Presented in Report	The inclusion of raw data is essential for critical reading of the reports. It provides the opportunity for the reader to get a sense of the basis upon which results were derived; greater (or less) confidence in reported results.	Report includes: <ul style="list-style-type: none"> <li>• absolute numbers used to assess changes in targeted problem;</li> <li>• data for entire time frame upon which results are based; and</li> <li>• data (absolute figures, preferably) used in calculation of any statistics.</li> </ul>

First, an evaluation report ideally presents information on both project activities and outcomes in a readable, logical fashion. The readability of a report is viewed as essential--not to be compensated for by the presence of other elements characteristic of good evaluation reporting. This requirement is based on the assumption that decision-makers are unwilling to spend valuable time untangling and sorting out information, regardless of its validity, which is haphazardly presented in an evaluation report.

Similarly, the inclusion of information regarding project activities is also considered important. Documentation must provide a clear picture of the basic dimensions of a project so that those wishing to apply the results of an evaluation know what activities were presumably responsible for the reported results. While the emphasis in recent years has been upon outcome-oriented evaluation, the reporting of information about project activities is nonetheless crucial for a full understanding of the context in which the evaluation has been performed.

Third, good evaluation reports are written with the non-technically oriented policy-maker and practitioner in mind. Reports should present information not only on the activities of the project and its outcomes, but also provide the reader with a clear understanding of the context in which this information is to be interpreted. Such an interpretive context may be provided by clearly specifying limitations in the research design used, citing changes made in project activities during the period of project operations being evaluated, and data problems encountered during the course of the evaluation. When this is done, the adequacy of information provided in the report may be a meaningful basis upon which decisions regarding project funding and operations can be made.



Finally, the documentation of data upon which reported findings are based is considered essential for a critical reading of an evaluation report. Figures underlying conclusions of project success or failure, when integrated into the text of the document, provide the reader with the opportunity to gauge the face validity of reported findings.

Based on these general criteria--readability, presentation of activity information, specification of limitations regarding the interpretation of findings, and inclusion of data to "validate" findings--a typology was developed (see Table II) to permit the classification of reviewed Impact project documentation according to some overall measure of evaluation reporting quality. Based upon this typology, four kinds of evaluation reports have been defined, each labeled in terms of a most comprehensive level of information provided:

- (a) No information
- (b) Descriptive information
- (c) Explanatory information
- (d) Substantiated information

When the manner in which information was presented in the report precluded an understanding of its source, purpose and meaning; the report was viewed as providing "no information." Also included in this category were those project documents which, while minimally acceptable in terms of readability, totally failed to provide the reader with an idea of what the project was about in terms of its activities and desired outcomes.

Project documents which only contained a description of project activities, in addition to any outcome information being reported, and which were considered to be logical, readable reports, were classified as providing "descriptive information." Failure to provide the reader

TABLE II  
EVALUATION REPORTING TYPOLOGY

TYPE (LEVEL)	CHARACTERISTICS	OPERATIONAL DEFINITION
1) No Information	The manner in which information is presented in the report precludes an understanding of its source, purpose and meaning in terms of the project's activities and outcome objectives.	<ul style="list-style-type: none"> <li>• Evaluation report judged "unacceptable" in terms of readability criterion (See Table I).</li> </ul>
2) Descriptive Information	Information is presented which describes the project's activities. The report may also describe outcomes observed but fails to interpret these outcomes within the context of the project's activities and the evaluation approach utilized.	<ul style="list-style-type: none"> <li>• Evaluation report judged "acceptable" in terms of readability criterion.</li> <li>• Report presents information regarding project activities (see Table I).</li> </ul>
3) Explanatory Information	Information is presented which provides a context for interpreting and understanding reported findings. Such contextual information serves as a basis from which to explain the occurrence of observed outcomes.	<ul style="list-style-type: none"> <li>• Evaluation report judged acceptable in terms of readability criterion.</li> <li>• Report provides information on project activities.</li> <li>• Report presents information which addresses possible limitations (caveats) in the interpretation of the findings/analysis results reported (see Table I).</li> </ul>
4) Substantiated Information	Information and data are presented which allow the reader to critically assess the plausibility of reported findings in terms of the methodological problems and assumptions inherent in the selected evaluation approach and the manner in which services were actually delivered and monitored.	<ul style="list-style-type: none"> <li>• Report judged "good" or "excellent" on readability criterion.</li> <li>• Report provides activity information.</li> <li>• Report addressed major limitations in interpretation of findings.</li> <li>• Report presents data upon which findings are based (see Table I).</li> </ul>



TABLE III  
CRITERIA REQUIRED FOR  
EVALUATION REPORTING QUALITY CLASSIFICATIONS

EVALUATION REPORTING QUALITY TYPES (LEVELS)	CRITERIA			
	READABILITY/ LOGIC	ACTIVITY INFORMATION	LIMITATIONS	DATA TO VALIDATE
LEVEL 1 (NO INFORMATION)	X <sup>1</sup>			
LEVEL 2 (DESCRIPTIVE INFORMATION)	X	X		
LEVEL 3 (EXPLANATORY INFORMATION)	X	X	X	
LEVEL 4 (SUBSTANTIATED INFORMATION)	X <sup>2</sup>	X	X	X

<sup>1</sup>Level 1 also includes those evaluation documentation for which information was presented in such a way that the reviewers were unable to understand its source, purpose, and meaning.

<sup>2</sup>To be categorized as a Level 4 evaluation, documentation was required to be judged "good" or "excellent" on the readability criterion.

with a context in which to interpret reported findings prevented this group of documents from being considered suppliers of explanatory information.

Documents viewed as being transmitters of "explanatory information" met three of the four major criteria used to guide the classification process: (a) readability, (b) a description of project activities, and (c) the inclusion of limitations and/or related information providing an interpretive context for reported findings. Missing from these documents were data needed to permit even a cursory validation of reported findings regarding project success or failure.

Those documents classified as providing "substantiated information" achieved this status by meeting all of the criteria used to define good evaluation reporting. In addition to presenting information in a logical, readable fashion, these documents also included a description of project activities, cautioned the reader regarding the interpretation and use of reported findings and included data needed to check the face validity of these findings. In brief, such documents were viewed as being excellent transmitters of information and as being structured in such a way as to encourage the application of reported findings.

### 3.2.2 Evaluation Approach Applicability

Transmitting evaluative information in an understandable, usable fashion does not in any way reflect nor insure the validity of the information being conveyed. Questions regarding the validity of reported findings must be answered through an examination of the strategy or approach used to gather this information. Thus, an examination of the type of evaluation approach used to gather reported information is essential to obtain a balanced judgment regarding the overall quality of an evaluation effort.

Assessing the applicability of an evaluation approach or strategy used to identify changes associated with a project's activities entailed a more technical review of the content of Impact documentation than the assessment of evaluation reporting quality. Of major concern was the extent to which observed changes in the targeted crime problem could reasonably be attributed to project activities. An evaluation approach which permits an analysis of attribution, by virtue of its ability to control or adjust for outside factors, is viewed as being the most suitable or applicable approach for measuring the success of Impact projects. The rigor of such an approach helps rule out other plausible explanations for observed outcomes. Starting from this ideal, a number of general characteristics (see Table IV) have been identified which, in conjunction with one another, permit the identification and classification of a range of evaluation approaches.

The key characteristic of a bona fide project evaluation approach is the specification of some basis against which observed values in the outcome measure may be compared and interpreted. For instance, the reporting of crime rates observed in an area patrolled by foot patrolmen funded through Impact monies is not especially meaningful unless it is possible to determine whether or not these observed rates represent an improvement in the area's crime picture. Thus, an evaluation approach must at least provide for the identification of changes in the outcome measure used to signal changes in the targeted problem. Changes observed must then be related to what would have happened without the project's implementation.

The type of comparison base employed in the evaluation effort is important; the degree to which the attribution goal of evaluation can be achieved is directly dependent upon the type of comparison base used in the effort. The logic underlying the use of a comparison base centers on the assumption that such use permits an estimate of

TABLE IV  
ASSESSING THE APPLICABILITY OF AN EVALUATION STRATEGY

CRITERIA	RATIONALE	OPERATIONAL DEFINITION
Allows for identification of changes in outcome measure.	Introduction of treatment predicated on the desire to change targeted problem. Thus, an evaluation approach should minimally provide some basis for identifying such changes.	Specification of some base/group/area whose status on the outcome measure will be compared to that of the treatment area/base/group.
Allows for control of non-project influences.	Logic underlying the use of comparison base/group/area centers on the assumption that its use permits an estimate of what would have happened to targeted problem if nothing had been done. Use of a comparison base/group/area which (before treatment) differs significantly from the targeted area/group may <u>severely</u> bias an estimate and render conclusions based on changes in outcome measures invalid.	Target area/group and comparison area/group appear to be similar with respect to pre-treatment experience, targeted problem, and with respect to key socio-economic or systems variables.
Allows for unbiased estimate of project effects.	Manner in which comparison selected influences the extent to which biases may creep into results. Random allocation of treatment insures against such biases. Statistical approximations often are a reasonable substitute.	Approach used centers on random allocation of treatment and/or explicit identification and monitorings of possible outside influences for estimation of linkage between project activities and outcomes.

what would have happened to the targeted problem if project intervention had not occurred. Thus, the use of a comparison base which, before treatment, differs significantly from the targeted area or group may severely bias this estimate and render invalid conclusions based on observed changes in the outcome measures. With this in mind, MITRE staff attempted to gauge the similarity with the area or group being targeted by paying close attention to documentation describing the use of the comparison base in the project evaluation effort. Of particular concern during this review were the criteria used to select the comparison base and the process by which the actual area or group comprising this base were selected.

Using the characteristics described above, three types of evaluation approaches (see Table V) have been defined:

- (a) No approach;
- (b) Change measurement; and
- (c) Attribution analysis.

Project evaluation documentation which either did not present any outcome findings or presented absolute figures with no point of comparison was classified as using "no approach." An evaluation approach viewed as merely allowing for the identification of changes in the targeted crime problem without controlling for other factors was classified as "change measurement." Finally, the most rigorous approach identified was characterized by its ability to control, explain or adjust for the influence of outside factors on observed changes, in addition to the use of at least two comparable data points to detect such changes. This approach was labeled "attribution analysis." This typology allows for the classification of Impact evaluation approaches according to the degree to which they permitted, by virtue of their inherent nature and the manner in which they were used, the identification of changes in the target problem, and the attribution of these observed changes to project activities.

TABLE V

EVALUATION APPROACH TYPOLOGY

TYPE (LEVEL)	CHARACTERISTICS	OPERATIONAL DEFINITION
1) No Approach	No information available to measure change in outcome or intermediate variable	Approach does not have two data points corresponding to either a before/after, treatment/control or observed/expected comparison.
2) Change Measurement	Information relating to size and direction of change in the outcome measure or surrogate vis à vis a control or comparison group/area	Approach must include two data points corresponding to either a before/after, treatment/control or observed/expected comparison.
3) Attribution Analysis	Observed changes in measure of performance are related or linked to project activities	Approach must include above, plus a plan to explain, control or adjust for other factors which could have influenced observed changes.

TABLE VI

CRITERIA REQUIRED FOR  
EVALUATION APPROACH APPLICABILITY CLASSIFICATIONS

EVALUATION APPROACH APPLICABILITY TYPE (LEVEL)	CRITERIA		
	NO COMPARISON POINTS	MEASUREMENT AGAINST CONTROL OR COMPARISON BASE	CONTROL FOR OUTSIDE FACTORS
LEVEL 1 (NO APPROACH)	X		
LEVEL 2 (CHANGE MEASUREMENT)		X	
LEVEL 3 (ATTRIBUTION ANALYSIS)		X	X

### 3.3 Data Collection

Classifying project-level evaluations according to the reporting and approach typologies discussed in the previous section was accomplished by using a two-stage review process. First, all of the documentation<sup>2</sup> received by 1 July 1975 for a specific project was assembled and technically reviewed on a first-come, first-serve basis by one of the three MITRE staff members involved in the review process. The results of these individually conducted technical reviews were recorded on data collection forms (see Appendix I) which were formatted to permit the recording of information corresponding to each of the review criteria discussed in the previous section. These collected data, processed as technical review packages, were then subject to revision and discussion by the entire MITRE group involved in the project evaluation review process.

These group discussions or conferences constituted the second stage in the assessment process. To encourage consistency in the classification of project evaluation documentation, the technical review package completed for each project was jointly read and discussed before assigning it to an evaluation reporting and evaluation approach applicability category. Typically these conferences required individual reviewers to recommend and justify, on the basis of the information they had recorded on the data collection forms, classification of the project into a specific category. In this way, the subjectivities inevitable in any review process were a matter of debate and thus checked to avoid systematic reviewer biases in the assessment process.

<sup>2</sup> Documentation included project grant applications and evaluation plans as well as evaluation reports. Reliance upon documentation other than project evaluation reports provided reviewers with a more complete picture of project activities and intentions. This picture was essential for realistically determining the applicability of the evaluation approach.

### 3.4 Analysis Approach

The analysis of data derived from the technical review of Impact evaluation documentation centers upon answering several basic questions concerning the conduct of evaluations in the Impact program. Among these are: what kind of evaluation was typically conducted? what coverage was provided by Impact evaluation funding incentives? and what was the overall quality of these project evaluation efforts? To answer these questions, program-wide distributions were developed for the following key variables: proportion of projects evaluated, type of documentation submitted, evaluation reporting quality and evaluation approach applicability. To permit an overview of program-wide performance, program-wide means were calculated for key variables and used to supplement discussions based on observed frequencies.

Next, analyses were performed to identify project-type factors which may have influenced the amount and quality of Impact project evaluation efforts. To accomplish this, variables related to the type of project being evaluated, such as project focus and type of implementing agency, were cross-tabulated with variables used to reflect evaluation performance (e.g., evaluation reporting quality, evaluation approach applicability, report type). The results of this analysis are presented in the first part of Section 5.0.

An assessment of each Impact city's evaluative performance then was undertaken. This analysis focuses upon (a) each city's overall response to its project-level evaluation responsibility, and (b) the quality of those project evaluations received and reviewed by MITRE staff for each of the eight cities. Relative city performances, based on a comparison of reporting quality and approach applicability scores, are discussed and then further interpreted in light of the proportion of implemented projects for which evaluation documentation has been reviewed.

Following the city-by-city analysis, an examination of the relationships among several aspects of the Impact project planning and evaluation process is presented. Associations observed among variables corresponding to various aspects of the Impact evaluation process (e.g., evaluation planning, evaluation reporting) are discussed in Section 6.0.

Finally, an analysis was conducted of the reported operational and evaluative limitations encountered by Impact personnel and reported via evaluation documentation. Variations in the frequency and nature of these problems were examined by cross-tabulating city, project focus and type of report with variables relating to the reporting and description of these problems. The results of this analysis are presented in Section 7.0.

#### 4.0 OVERVIEW

The belief that project-level evaluation could serve as a viable mechanism for gathering information to gauge the success of specific anti-crime tactics caused LEAA to mandate the evaluation of all Impact-funded projects. Incentives for evaluation were built into the program from the outset: monies specifically earmarked for evaluation were provided on a no-match basis, as well as the promise of technical assistance to be coordinated through the research arm of the LEAA. While evaluation monies were indeed allocated (\$625,000 per city), technical assistance to the individual cities was not supplied in a coordinated, consistent fashion. Thus, project evaluation has been truly a state or local function supported primarily through financial incentives.

With the program evaluation goal and incentives in mind, this section will provide a general description of what happened in the Impact program in terms of project-level evaluation. Specifically, it details the number of project-level evaluations known to have been carried out, the type of evaluation conducted to date, the nature of the evaluated projects, and the overall quality of these efforts given the federal monies expended. A more detailed view and discussion of analysis results, centering upon variations among the project types and Impact cities, is presented in Section 5.0.

##### 4.1 Review Coverage

The most recent count of Impact projects indicates that 233 projects have been funded under the Impact program. Of these, 140 (or 60 percent) had forwarded at least one evaluation report as of 1 July 1975. Documentation for only 119 projects (51 percent) was actually subjected to technical review for use in this assessment.<sup>3</sup>

<sup>3</sup> Documentation received for 21 projects was not considered relevant for this analysis since it consisted essentially of plans for future project implementation or evaluation.

This means that technical reviews could not be conducted for almost half of the funded Impact projects, thereby limiting the amount of information available for use in this document (as well as in other efforts designed to assess the Impact program).

This gap in project information is only partially due to the slowness with which Impact projects were designed and funded.<sup>4</sup> In fact, as shown in Table VII, the majority (76 percent) of those projects for which documentation was not received and reviewed, had been funded for over one year as of 1 July 1975. Given the LEAA evaluation guidelines and their emphasis upon timely, interim evaluations, some documentation could have been expected for this subset of 71 projects.

##### 4.2 Type of Report

The type of information contained in the documentation for the 119 reviewed projects varied widely. This variation reflects, in part, the operational period of evaluated projects as well as decisions made among the Impact cities regarding the manner in which evaluation responsibilities would be fulfilled. Appendix II provides a description of the manner in which the project-level evaluation responsibility was met in each Impact city. Because of these differences, reviewed project documentation was classified into four categories according to the amount and type of information presented. This classification provides, therefore, a more accurate picture of the nature of the Impact evaluation documentation reviewed. These four groups are characterized below:

- (a) Status reports present minimal information. Questions regarding project success in solving targeted crime problems are not addressed. Information contained in

<sup>4</sup> For an analysis of Impact project implementation see, The MITRE Corporation Technical Report MTR-6961, "An Assessment of the Implementation Quality of High Impact Anti-Crime Projects," dated August 1975.



TABLE VII  
OPERATIONAL PERIOD OF PROJECTS FOR WHICH REPORTS  
WERE NOT RECEIVED AND REVIEWED

PERIOD FROM DATE  
OF INITIAL AWARD TO  
JULY 1975

1 YEAR OR LESS	22	24%
1 - 2 YEARS	40	43%
2 YEARS OR MORE	31	33%

N = 93 (PROJECTS NOT RECEIVED  
AND REVIEWED)

these reports is restricted to issues such as project start-start-up, hiring and training of staff and site selection. Activity data concerning the fulfillment of stated activity objectives are not presented.

- (b) Progress reports also do not address questions of project success regarding targeted crime problems. These reports do, however, address issues related to the provision of designated services and activities.
- (c) Preliminary evaluation reports provide initial or sketchy information regarding project success in solving a targeted crime problem. The rudimentary nature of this information is typically due to insufficient operating time or numbers of clients served.
- (d) Full-fledged evaluation reports contain in-depth information regarding project success. Information presented is sufficient in scope to provide a relatively detailed picture of project outcomes.

Projects were assigned to one of the above categories based on all evaluation documentation received for a given project. That is, where more than one report existed for a project, classification decisions rested upon the degree to which information in the combined reports satisfied the classification criteria described above. For instance, a project may have submitted two reports, one which dealt solely with implementation and another presenting tentative outcome information based on short term or incomplete data. Documentation for such a project would be classified as "preliminary evaluation" to reflect the provision of preliminary evaluation information in at least one report.

Figure 1 displays the distribution of the 119 projects for which documentation was reviewed when classified in the above fashion. Documentation submitted for over half of these projects (57 percent) could be considered to constitute full-fledged evaluation reports. When this percentage is added to the percentage of reports assessed as preliminary evaluations (15 percent), it is clear that almost

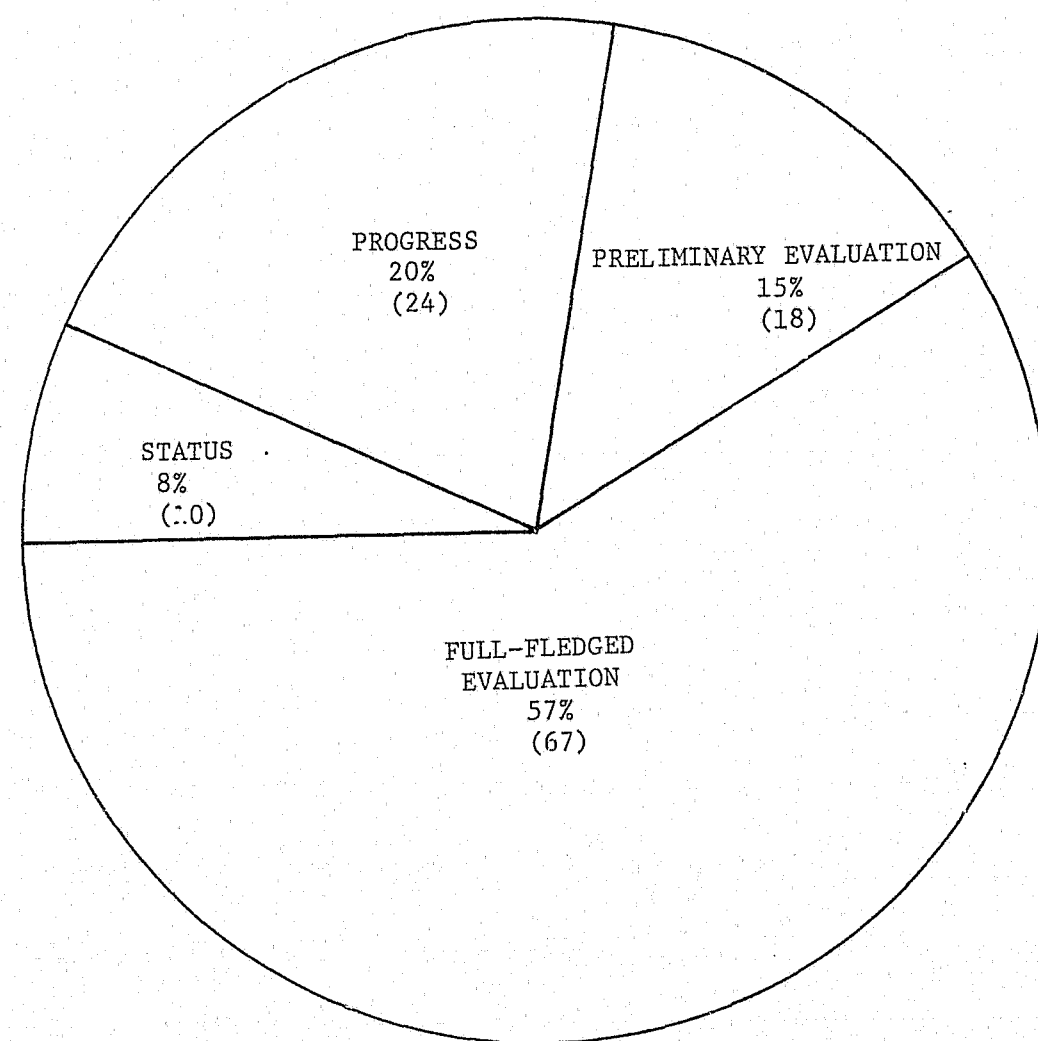


FIGURE 1  
DISTRIBUTION OF REVIEWED PROJECT  
DOCUMENTATION BY REPORT TYPE

three-fourths (72 percent) of the reviewed projects submitted documentation containing some information regarding project outcomes. This percentage is less encouraging when all 233 Impact-funded projects are considered. Information about the success of anti-crime tactics implemented in the Impact program was available for only 36 percent (67 "full-fledged" evaluations and 18 "preliminary" evaluations out of a total of 233) of these projects. Yet the evaluation of 85 projects constitutes no mean achievement at the local level, and presents evidence of real progress in the evaluative successes produced to date by federally-funded social action programs. These 85 projects account for \$48,573,000 in federal funds, or 35 percent of all the Impact monies awarded.

#### 4.3 Reliability of Reported Findings

An evaluation report typically presents findings and conclusions regarding the project's success in solving its targeted crime problem. These findings and conclusions are often based on a large and diverse set of data generated using a variety of evaluation approaches (or no approach at all). Assessing the extent to which such findings are reliable and valid (and thus justified) in light of the evaluation approach used in an essential aspect of any critical review of evaluation documentation. To realistically assess the overall Impact project evaluation effort, it is therefore necessary to consider the reliability of those findings resulting from this effort. Figure 2 depicts the results of MITRE's assessment of the extent to which project-level evaluation findings (in terms of the targeted crime problem) were considered justified. Of the Impact projects documented and reviewed, findings related to project outcomes were provided 71 percent of the time (85 of 119 projects). The presence of findings regarding project outcomes does not mean, however, that the project was considered a success. In fact, of those projects for which findings were reported in the reviewed documentation (85), 32 percent or 27 indicated that

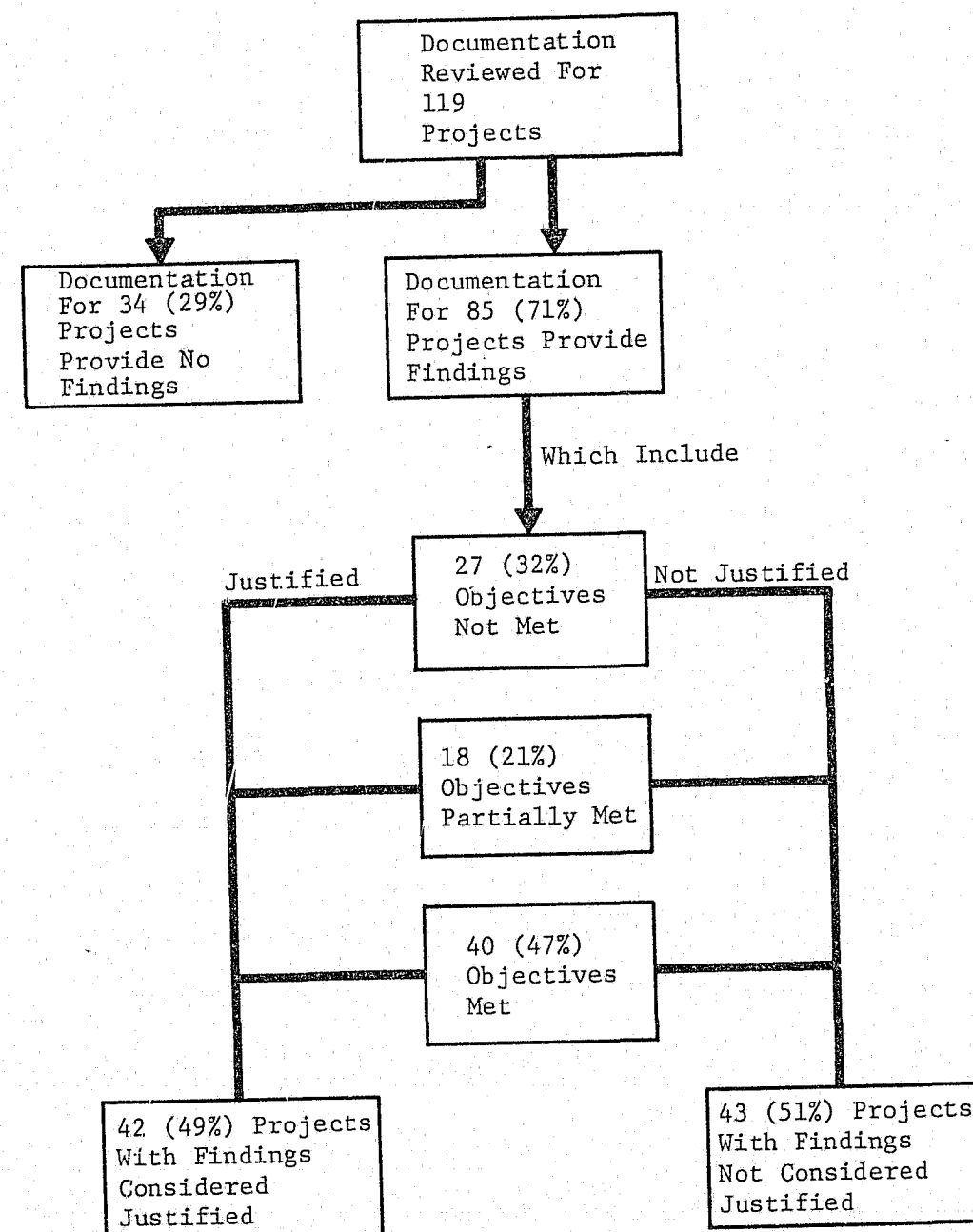


FIGURE 2  
JUSTIFICATION OF PROJECT EVALUATION FINDINGS

their major crime-specific objectives were not met. Additionally, 21 percent of the projects reviewed indicated that the project's major outcome objectives had been only partially met, leaving 47 percent explicitly claiming project outcome success.

Claims of project success, partial success, or failure are not always justified. Justification depends largely upon the approach used to generate and analyze data underlying such assertions. To provide an accurate picture of the success of Impact-funded projects, claims of project success or failure were weighed by MITRE staff in light of the evaluation approach and data used. In essence, MITRE staff determined the credibility of these assertions based on information provided in the report. Again, Figure 2 presents the results of this aspect of the assessment process.

Evaluation findings were considered justified and substantiated for almost half (49 percent) of the 85 projects for which reviewed documentation included statements concerning project success or failure. The percentage of findings considered justified did not vary significantly with the reported success or failure of the project. Claims of project success in achieving major crime-specific objectives were considered justified for 17 of the 40 projects (42 percent) while claims of objectives being partially met appeared substantiated in 56 percent of the cases (10 of 18). Similarly, statements of failure were also believed 52 percent of the time (15 of 29).

In sum, the manner in which Impact evaluators collected, analyzed and presented information to support conclusions regarding the success of projects in meeting their outcome objectives raised questions in the minds of MITRE reviewers, regarding the legitimacy of such conclusions almost half of the time, regardless of the nature of the reported findings.

#### 4.4 Evaluation Reporting

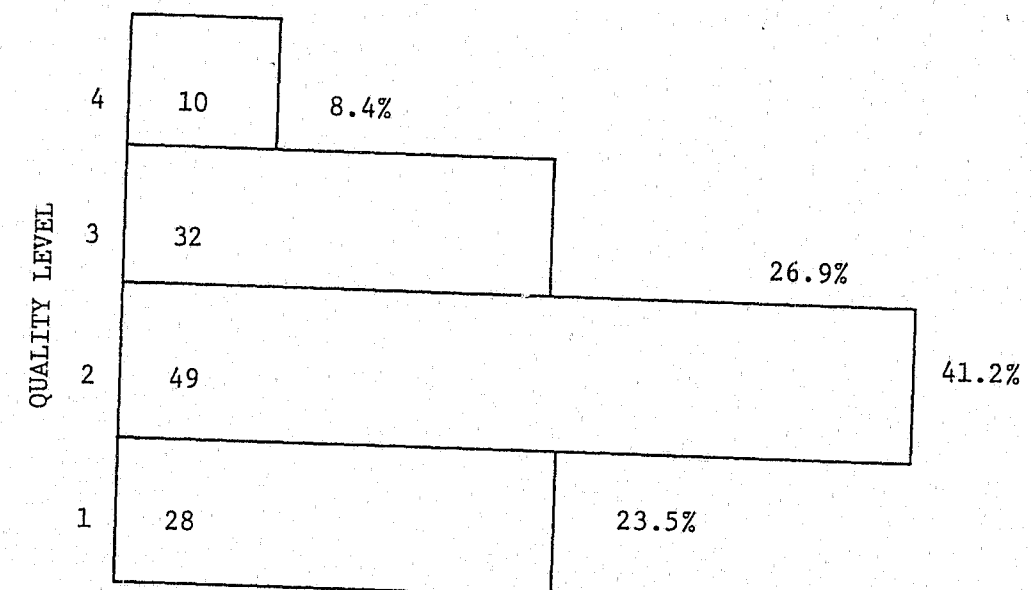
Reviewed Impact evaluation documents varied in their ability to transmit a clear picture of the progress projects had made in achieving their activity and outcome objectives (see Table VIII). Evaluation documentation for 28 projects (23.5 percent of the total number of reviewed projects) were viewed as providing "no information" (quality level 1) according to the evaluation reporting typology developed in Section 3.0. This means that the manner in which information was presented for these projects precluded an understanding of its source, purpose or meaning. Documents from these projects either essentially failed to provide a sense of what the project was trying to accomplish or presented the information in what was considered an unreadable fashion.

The greatest proportion of projects (41.2 percent) submitted documentation which provided strictly "descriptive information" (quality level 2). That is, these documents contained a readable description of project activities but did not provide the reader with a context for interpreting any outcome information which may have been included.

Over one quarter (26.9 percent) of the projects fell into the category labeled "explanatory information" (quality level 3). These are projects whose documentation was viewed as providing the reader with a good understanding of the project. These reports were readable and included limitations regarding the interpretation of findings as well as a cogent description of project activities. These documents did not include, however, data needed to permit even a cursory validation of reported findings.

"Substantiated information" (quality level 4) was presented for only 10 or 8.4 percent of the projects reviewed. This means that the information transmitted for these projects included data which allowed

TABLE VIII  
DISTRIBUTION OF PROJECTS BY EVALUATION REPORTING QUALITY



N = 119

#### LEGEND:

- Level 1 = No Information
- Level 2 = Descriptive Information
- Level 3 = Explanatory Information
- Level 4 = Substantiated Explanation

a critical assessment of analysis results, in addition to stating major limitations and providing activity data. Additionally, the documentation for these projects was considered to be extremely well written, with information presented in a coherent, logical fashion.

Specific weaknesses observed in Impact evaluation reporting efforts are more easily discussed when the criteria comprising the overall measure of evaluation reporting quality are individually examined. Table IX permits a discussion of each individual criterion.

Despite the overwhelming emphasis upon outcome-oriented evaluative information, which characterized the type of project evaluation to be conducted in the Impact program, 80 percent or 95 of the projects for which documentation was reviewed provided the reader with a description of the project's activities. While project activity information was obviously not forsaken in lieu of strictly outcome-oriented information, it should be recognized that the fulfillment of this specific activity information criterion did not require a detailed accounting of specific project activities. Rather, reports had only to provide a general description of what the project was doing to solve the targeted crime problem. This type of description, while sufficient for decisions based on project outcomes, would not suffice for users of these reports who were interested in project replication or transfer.

Overall, the majority of Impact project evaluations were considered to be acceptably written. Only 7 percent (8 of the 119 reviewed projects) were documented in what was considered an unacceptable fashion. In these documents information was presented in a totally disorganized fashion, making it virtually impossible to tie findings reported either to the project's stated objectives or to tables included in the report. Almost half of the projects reviewed (57 of 119, or 48 percent) submitted documentation which

TABLE IX

DISTRIBUTION OF PROJECT DOCUMENTATION  
BY EVALUATION REPORTING CRITERION

ACTIVITY INFORMATION PROVIDED

YES	NO	
80% (95)	20% (24)	N = 119

RELATIVE READABILITY

UNACCEPTABLE	ACCEPTABLE	GOOD	EXCELLENT	
7% (8)	48% (57)	38% (46)	7% (8)	N = 119

LIMITATIONS REPORTED

YES	NO	
78% (93)	22% (26)	N = 119

DATA INCLUDED FOR FACE VALIDATION

YES	NO	
39% (47)	61% (72)	N = 119

was viewed as being "acceptable" in terms of the readability criterion. Documentation from these projects, while presenting information in a fairly logical fashion, typically did not present a lucid description of the method used to assess objective attainment or provide detailed data analysis on an objective-by-objective basis. Projects whose documentation was considered "good" in terms of readability (46 of 119, or 38 percent) generally presented a thorough description of the approach used to gauge objective attainment. These projects also reported findings and data on an objective-by-objective basis. Shortcomings of these reports in terms of their readability were typically a failure to fully integrate data and tables into the text of the report and to provide cogent summaries of major findings. Only 7 percent (8 of 119) of the projects reviewed provided documentation considered to be of excellent quality (i.e., the same number precisely as the total of reports that were unacceptably written). These documents possessed all of the characteristics associated with a "good" rating in terms of readability, also integrated data and tables into the text, and provided succinct summaries of major findings.

The "limitations" criterion is perhaps the most important of all the criteria used to define good evaluation reporting. By reporting limitations encountered in the course of identifying and assessing the relationship between observed outcomes and project activities, evaluators facilitate the accurate interpretation and use of reported findings. Reflecting the importance assigned to this aspect of evaluation reporting, a detailed discussion of this criterion is presented in Section 7.2. For the purposes of this overview, suffice it to say that 78 percent, or 93 of the 119 projects reviewed, reported limitations regarding their project evaluation efforts.

Finally, Impact evaluation reporting efforts typically fell short when it came to providing data needed to critically assess the

reliability and validity of reported findings. Only 39 percent of the projects for which documentation was reviewed reported absolute figures used in the calculation of statistics and provided as well a complete display of relevant percentages. In contrast, 61 percent, or 72 of the 119 Impact projects reviewed did not include the data necessary for the face validation of reported results, thus precluding a truly critical reading of the evaluation report.

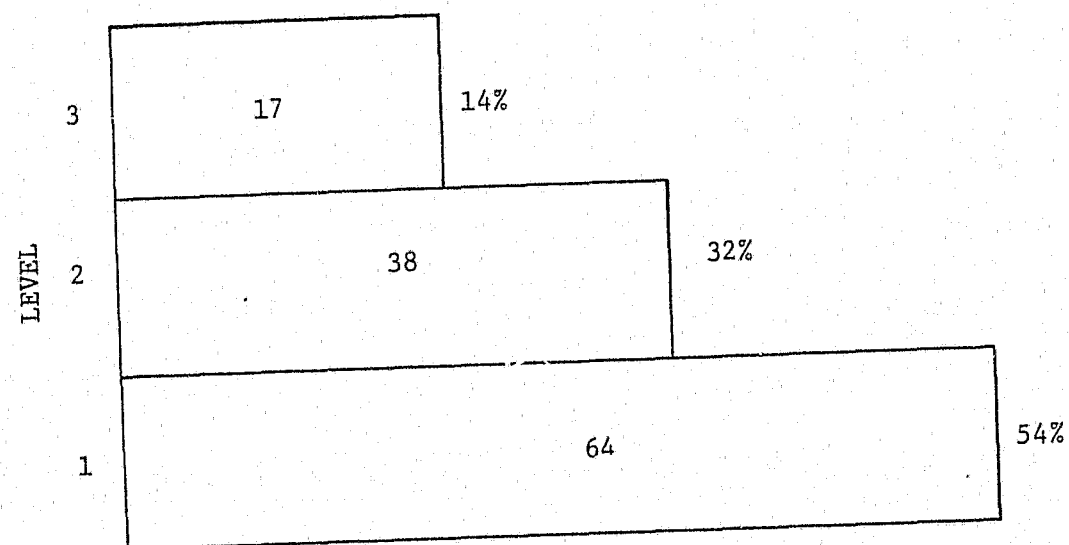
#### 4.5 Evaluation Approach Applicability

Impact city evaluators relied upon a variety of evaluation approaches to identify changes in targeted crime problems. These approaches, by their inherent nature, allow for differing degrees of confidence in the results obtained regarding a project's performance vis-à-vis its outcome objectives. Documentation received and reviewed for the 119 Impact projects was assessed according to the evaluation approach applicability criteria defined in Section 3.2.2 (see page 11). The extent to which these criteria are met reflect, in part, the degree of confidence that may be placed in the reported findings.

Project-level information was gathered and reported by city evaluators for 54 percent of the reviewed projects in the absence of any discernible evaluation approach (see Table X below). This means that information for these projects did not provide insights as to changes which may have occurred in the targeted crime problem. Rather, information presented included, at best, absolute figures regarding the major outcome measure of the project. For example, a recidivism-reduction project designed to change the criminal behavior patterns of youth may have reported that 10 of the 100 clients served by the project in a year were rearrested. In the absence of additional information concerning the rearrest rates which might be expected among project clients without project intervention, it is impossible to determine if the absolute figures presented (10 rearrests for 100



TABLE X  
DISTRIBUTION OF APPROACH APPLICABILITY  
OF PROJECTS



N = 119

LEGEND

- 1 = No Approach
- 2 = Change Measurement
- 3 = Attribution Analysis

clients) represent any improvement for project clients. Evaluation documents providing information similar to or less rigorous than that contained in the above example were classified as having no evaluation approach.

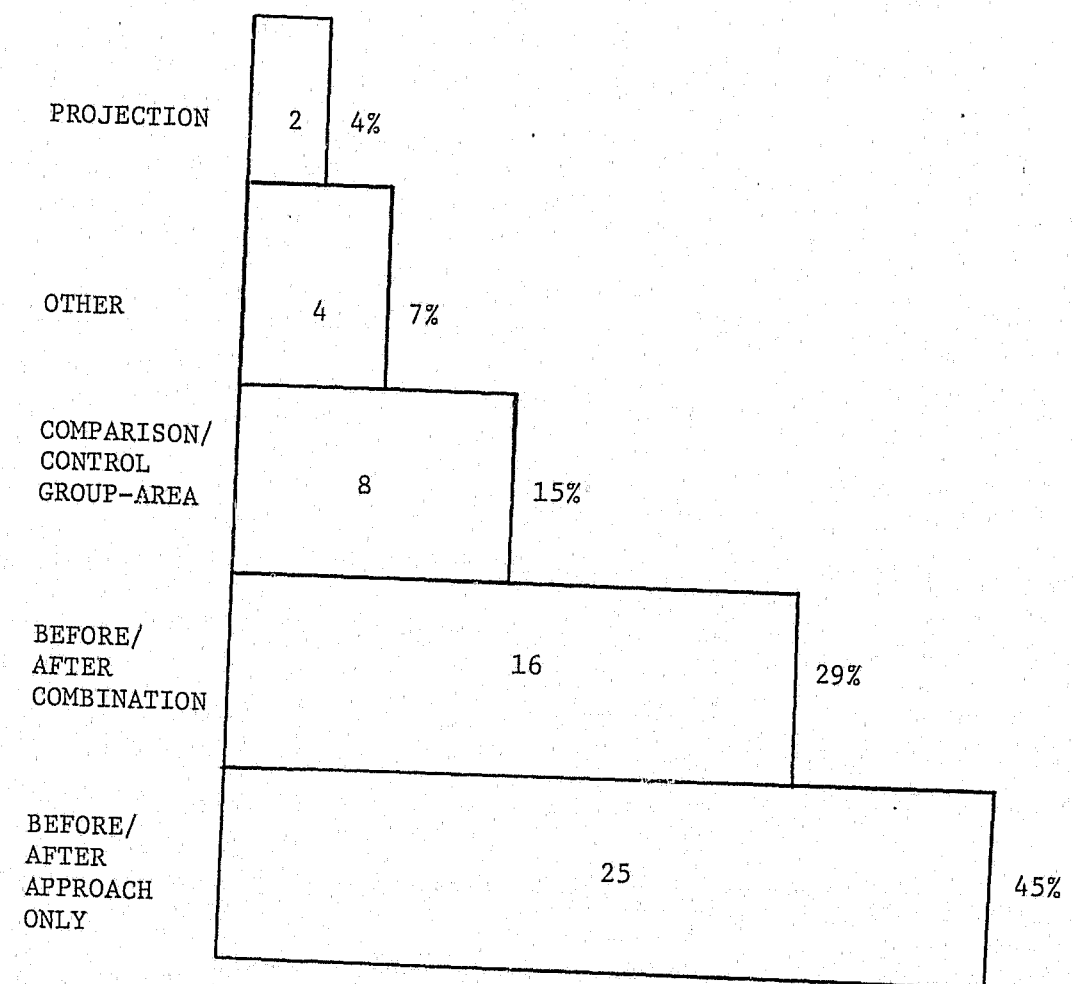
As shown in Table X, almost one third (32 percent) of the projects reviewed used an evaluation approach which was classified as providing "change measurement." Such an approach allows for the identification of changes in the targeted crime problem without controlling for outside factors which are necessary to link them to observed changes in the activities of the project. In the example presented above, comparing the 10 percent rearrest rate (10 rearrests for 100 clients) to such standards as rearrest rates for clients in a similar program, accepted national standards, or client criminal behavior before project participation, would allow for the identification of changes that occurred among project clients. These approaches while providing for the measurement of changes do not, however, permit the direct attribution of these changes to project activities.

Finally, 14 percent (17 of 119) of the projects reviewed used an evaluation approach considered to be rigorous. These approaches were characterized by the fact that they control, explain, or adjust for the influence of outside factors on observed changes and thus permit the attribution of these changes to project activities. Using the above example once again, random selection and assignment of potential project clients to treatment and non-treatment (control) groups would have permitted a more rigorous assessment of the effects of the project's treatment activities. Observing a 5 percent rearrest rate among the treatment group as compared (for example) to a 10 percent rate among a randomly selected control group could then be interpreted as an indication of positive change in the behavior of project clients which is reasonably attributable to the activities of the project.

A sense of the type of evaluation approaches which were evidently feasible within the context of the Impact program, is provided by looking more closely at the 55 projects (46 percent of the 119 reviewed projects) which used some type of evaluation approach. Table XI indicates that for nearly half (45 percent) of the projects reviewed which used some type of evaluation approach, Impact evaluators relied solely upon a before/after design. Next in frequency of use were those approaches that combined the essential aspects of the before/after design with some type of comparison area or group (29 percent of those projects having an evaluation approach). Evaluations of this nature relied upon, for example, comparisons of crime rates observed before and after project implementation in a targeted and non-targeted area to determine project success. The remaining project documentation reviewed contained information derived from a variety of evaluation approaches; these included the sole use of a comparison area or group (15 percent), projections (4 percent) and a mixture of different statistical methods (7 percent).

Based on this overall assessment of the type and applicability of evaluation approaches used in Impact evaluation efforts, it appears that 86 percent (54 percent using no approach plus 32 percent with a "change measurement" approach) of the projects reviewed were evaluated using designs which were considered less than rigorous in evaluation research. These approaches, as discussed in Section 4.3, permit considerable latitude in the interpretation and therefore in the reliability of reported findings. This latitude is evidenced by the large proportion of projects (i.e., 51 percent as indicated in Figure 2) whose documentation contained findings which were not considered to be justified by MITRE staff. As a result, information needed to identify successful anti-crime tactics has not been accumulated to the fullest extent.

TABLE XI  
DISTRIBUTION OF EVALUATION APPROACH TYPES



N = 55 (Projects using some approach)

## 5.0 ANALYSIS RESULTS

Looking at the current status of project evaluation for the Impact program as a whole, MITRE found that about half of the Impact projects were documented as a result of project evaluation efforts. Additionally, the type of evaluation conducted to date typically focuses upon questions of project outcomes, although answers to these questions are often necessarily tentative due to the relatively short time many of the projects have been operating. And finally, it appears from the data at hand that rigorous evaluation approaches were not used extensively for assessing project outcomes, possibly because of the free-form (New Federalist) character of the program and the overriding demands for service provision.

Given the diversity of projects funded through Impact, and the differing city contexts in which these projects were planned, implemented, and evaluated, it is important to know whether the type and quality of Impact project evaluations described above were uniform or different among varying types of projects and across the eight cities. In the following sections, observed similarities and differences will be discussed.

### 5.1 Project Focus

Impact monies were used to design and implement a variety of anti-crime approaches. These approaches or tactics for reducing crime may be classified according to their general target or focus. First, there are those tactics which target crime in a direct fashion. Projects implementing this approach to the crime problem concentrate on the direct prevention and control of crime in specific geographical areas. For instance, foot patrol projects and street lighting projects both seek to reduce crime by increasing the risk associated with either committing a crime in areas more frequently patrolled due to the presence of foot patrolmen or areas more visible due to

increased lighting. Projects of this type are classified, therefore, as having a direct "crime-reduction" focus. Second, there are those tactics which target or focus upon those who commit crimes. Projects implementing this approach to the crime problem emphasize the prevention or reduction of recidivism on the part of potential or established criminals. For instance, juvenile rehabilitation projects and diversion projects both seek to affect the behavior of people who have shown evidence of criminal involvement. Projects of this type are classified as having a "recidivism-reduction" focus.

Third, there are those tactics which seek to address the crime problem by improving various aspects of the criminal justice system. Projects implementing this less direct approach to crime problems include court improvement projects, information systems projects and research projects. Projects of this type are classified as having a "systems/other" focus. Given these three types of projects, each with its own distinctive focus, the type and quality of evaluation conducted to assess their effectiveness as anti-crime approaches is explored below.

#### 5.1.1 Review Coverage

Before discussing the quality of Impact evaluation efforts by project focus, it is necessary to know the proportion of projects in each focus category which were documented and reviewed. These proportions may be considered as indicators of the priorities of Impact evaluators and of the directions in which Impact evaluation energies were channeled.

Table XII below presents the number of projects in each focus category along with information regarding the proportion of these projects which have been documented and reviewed by MITRE staff. It is clear, first of all, that the number of projects funded differs widely across project foci.

TABLE XII  
DISTRIBUTION OF TECHNICALLY REVIEWED PROJECTS  
BY PROJECT FOCUS

PROJECT FOCUS	PROJECT FOCUS AS A PERCENT OF TOTAL IMPACT FUNDED PROJECTS	PROJECTS TECHNICALLY REVIEWED AS A PERCENT OF FUNDED PROJECTS
CRIME-REDUCTION	19% (45)	53% (23)
RECIDIVISM-REDUCTION	47% (109)	57% (61)
SYSTEMS/OTHER	34% (79)	46% (35)
TOTAL	100% (233)	51% (119)

Almost half (47 percent) of the funded projects were designed to reduce recidivism. One-third of the Impact projects were aimed at criminal justice system improvement, while crime was directly targeted by projects designated as incorporating crime reduction tactics in only 45 out of a total of 233 projects (19 percent).

Looking at the proportion of projects documented and reviewed, there is a remarkably high degree of consistency across project foci. For each focus, approximately half of the funded projects have

submitted documentation for review by MITRE staff. While the proportion of systems/other projects documented and reviewed (46 percent) lags slightly behind that observed for recidivism-reduction projects and crime-reduction projects (57 and 53 percent respectively), this difference does not suggest a bias in the submission of documentation for different types of projects.

#### 5.1.2 Type of Reports Reviewed

Variations in the type of reports reviewed are evident among projects of differing foci. Figures presented in Table XIII indicates that reports reviewed dealing with direct crime-reduction projects were overwhelmingly classified as full-fledged evaluation reports.

TABLE XIII  
TYPE OF REPORT REVIEWED BY PROJECT FOCUS

PROJECT FOCUS	N	TYPE OF REPORT			
		STATUS	PROGRESS	PRELIMINARY	FULL-FLEDGED EVALUATION
CRIME-REDUCTION	23	8.7% (2)	0% (0)	21.7% (5)	69.6% (16)
RECIDIVISM-REDUCTION	61	6.5% (4)	19.7% (12)	14.8% (9)	59.0% (36)
SYSTEMS/OTHER	35	11.4% (4)	34.3% (12)	11.4% (4)	42.9% (15)
	119	8% (10)	20% (24)	15% (18)	57% (67)

Over ninety percent of the documents submitted and reviewed for projects of this focus provided information gauging the extent to which project outcome objectives, in this case crime rate reductions, were being met. That is, they were classified as preliminary or full-fledged evaluation

reports. Additionally, none of the documents reviewed for crime-reduction projects were classified as progress reports. This finding is not surprising since periodic assessments of crime rate variations may be initiated concurrently with project implementation due to the continuous and standardized measurement of reported crime rates based on police statistics.

Information dealing with project outcomes was also prevalent among the reports reviewed for recidivism-reduction projects, although to a lesser extent than that observed for direct crime-reduction projects. This may be due in part to the problems associated with the measurement of recidivism as an indicator of project performance. Measuring recidivism often involved the establishment of new data sources and the follow-up of client behavior both during and after participation in an Impact project. These problems perhaps explain the lower proportion of recidivism-focused projects providing information dealing with project outcomes (73.8 percent for recidivism-reduction compared to 91.3 percent for crime-reduction).

For those projects designed to improve the functioning of the criminal justice system, (i.e., systems/other) one finds a greater proportion of the reviewed documentation restricted to an accounting of project activities and operations (11.4 percent status and 34.3 percent progress). This concentration on activity information as opposed to findings related to project outcomes is not unusual given the inherent nature of these projects. For these types of projects, the entire sequence of events comprising the system to be modified must frequently be played out before improvements may be measured. Furthermore, many of these projects were not designed to permit the type of outcome-oriented evaluation which was intended to characterize the overall Impact project evaluation effort.

#### 5.1.3 Evaluation Reporting Quality

Little variation is observed when looking at the evaluation reporting quality of documents (see Table XIV) reviewed for each of the three project foci. Mean reporting quality scores for projects in each of the three foci are almost identical all clustering around 2.2. This lack of variation among project foci is not surprising. There is little inherent in the focus of a project which would be expected to affect the manner in which information about that project is conveyed via a project evaluation report. Thus, while project activities and limitations encountered in actually carrying out the evaluation may differ among different types of projects, the careful discussion of these limitations or the provision of data where such is available (as evidenced by reported findings) is not affected.

#### 5.1.4 Evaluation Approach Applicability

Transmitting readable, usable evaluation reports is only one aspect of the project evaluation process. At least as critical to the overall evaluation effort is the approach used to gather the information on project outcomes presented in the report. Different approaches provide varying degrees of confidence regarding the extent to which observed outcomes may be attributed to project activities. The applicability of the approach actually used in the evaluation is often colored, however, by real-life constraints and considerations.

Table XV provides a look at the three types of projects implemented in the program (i.e., crime-reduction, recidivism-reduction and systems/other) in terms of evaluation approach applicability. The percentage distributions presented in this table indicate considerable variation in the applicability of evaluation approaches used to assess the success of projects of varying foci. Calculating the mean for evaluation documentation reviewed for each project, crime-reduction projects used evaluation approaches which were most

TABLE XIV  
EVALUATION REPORTING QUALITY BY PROJECT FOCUS

PROJECT FOCUS	EVALUATION REPORTING QUALITY PERCENT IN EACH LEVEL*					MEAN QUALITY**
	N	1	2	3	4	
CRIME-REDUCTION	23	22% (5)	43% (10)	22% (5)	13% (3)	2.26
RECIDIVISM-REDUCTION	61	28% (17)	39% (24)	25% (15)	8% (5)	2.13
SYSTEMS/OTHER	35	17% (6)	43% (15)	34% (12)	6% (2)	2.29
TOTAL	119	23.5% (28)	41.1% (49)	27% (32)	8.4% (10)	2.20

\* Level 1 indicates the lowest reporting quality (No Information) while Level 4 (Substantiated Information) indicates the highest.

\*\* The mean quality for each focus is obtained by multiplying the number of projects by the reporting quality level, summing, and dividing the total by the number of projects. Thus for Crime-Reduction:  

$$\frac{(1 \times 5) + (2 \times 10) + (3 \times 5) + (4 \times 3)}{23} = \frac{52}{23} = 2.26$$

TABLE XV  
EVALUATION APPROACH APPLICABILITY  
BY PROJECT FOCUS

PROJECT FOCUS	N	EVALUATION APPROACH APPLICABILITY PERCENT IN EACH LEVEL*			MEAN SCORE
		1	2	3	
CRIME-REDUCTION	23	26% (6)	61% (14)	13% (3)	1.87
RECIDIVISM-REDUCTION	61	53% (32)	31% (19)	16% (10)	1.63
SYSTEMS/OTHER	35	74% (26)	14% (5)	12% (4)	1.38
TOTAL	119	54% (64)	32% (38)	14% (17)	1.60

\* Level 1 indicates the lowest approach applicability (No Approach) while Level 3 (Attribution Analysis) indicates the highest.



highly assessed ( $X = 1.87$ ); followed by recidivism-reduction projects (1.63); and systems/other projects (1.38).

Among all projects reviewed, approximately 75 percent of those aimed toward direct reductions in crime used an evaluation approach which allowed for the identification of changes in the targeted problem (i.e., levels 2 and 3). Additionally, crime-reduction projects, when compared to other project types, had the lowest percentage (26 percent) of reports reviewed classified as providing no basis for comparing observed outcomes (i.e., level 1).

Gauging the success of projects designed to reduce recidivism was apparently a less straightforward job for Impact evaluators. As indicated in Table XV, questions of recidivism-reduction were addressed using some type of evaluation approach for about half (19 projects in level 2 and 10 in level 3 out of the 61 recidivism projects) of those reviewed. However, one-third (10 of 29) of this subset of projects provided documentation that contained findings generated via the use of an evaluation approach considered rigorous within the Impact context. (This is more than three times the number achieving this status among crime-reduction focused projects.)

Finally, only 26 percent (9 of 35) of the systems/other projects reviewed presented findings based on some point of comparison. This means that for almost three-fourths of the systems/other projects subject to the review process, quantitative estimates of changes resulting from the project's implementation and operation were not obtainable.

Table XV also shows that the proportionate use of a rigorous evaluation approach (level 3) did not differ significantly (13 percent, 16 percent, and 12 percent) among the three project foci. However, the proportion of projects which lacked a discernible

evaluation approach (level 1) was substantially higher among systems/other and recidivism-reduction projects than projects with a direct crime-reduction focus.

## 5.2 City Analysis

In this section, the project-level evaluation performance of the eight Impact cities will be explored in detail. An overview of relative city performances is provided based on a comparison across cities of the proportion of projects documented and reviewed, the type of reports reviewed, mean reporting quality and mean approach applicability scores. These comparisons provide a picture of how well the cities performed relative to one another for those projects which were documented and could be reviewed by MITRE staff.

### 5.2.1 Review Coverage

In order to obtain a realistic picture of each Impact city's evaluative performance, the proportion of funded Impact projects evaluated and reviewed by MITRE staff is discussed below by city. This discussion focuses upon the number of projects documented from each city and the number of projects which have been funded for over one year. This latter figure reflects the number of projects for which evaluation information could realistically have been expected. In conjunction with one another, these figures provide a sense of the progress each city has made to date in fulfilling Impact project evaluation requirements. Table XVI facilitates this comparison.

Clearly, variations exist among cities in the proportion of city projects represented in this review process; they range from a low of 33 percent in Newark to a high of 77 percent in Cleveland and average 51 percent across cities. This average means that for almost one-half of the Impact projects, no information was available for use in the overall evaluation of the Impact program experience as of

TABLE XVI  
PROPORTION OF PROJECTS DOCUMENTED  
AND REVIEWED (BY IMPACT CITY)

	TOTAL NUMBER IMPACT PROJECTS	NUMBER OF PROJECTS FOR WHICH DOCUMENTS WERE REVIEWED	REVIEWED PROJECTS AS A PERCENT OF ALL PROJECTS	PERCENT OF PROJECTS WITH NO DOCUMENTA- TION (FUNDED LESS THAN ONE YEAR)
ATLANTA	20	10	50%	40%
BALTIMORE	27	20	74%	17%
CLEVELAND	39	30	77%	0%
DALLAS	19	11	58%	33%
DENVER	37	14	38%	43%
NEWARK	27	9	33%	33%
PORTLAND	17	6	35%	9%
ST. LOUIS	47	19	40%	0%
TOTAL	233	119	51%	24%

1 July 1975. Furthermore, this lack of information is not restricted to newly-funded projects whose activities have not existed long enough to permit the gathering of information on either its progress or outcomes. As shown in Table VII (see Section 4.0, page 22), approximately 75 percent of all Impact-funded projects have been funded for well over a year.

Evaluation documentation (as reflected in Table XVI), seems to have been especially noteworthy in Baltimore and Cleveland with respect to the proportion of projects for which evaluation documents were received and could be technically reviewed. The most comprehensive coverage of Impact-funded projects, in terms of this review process, was provided by Cleveland evaluators, with 77 percent of all projects having been subject to evaluation. Additionally, two reports were generally submitted for each documented project, with later reports consisting of an assessment of project performance from a strictly management perspective.

Reviewing Baltimore's evaluation documentation entailed examining reports prepared by the SPA for almost three-fourths (74 percent) of all Impact projects funded in the city. At the time these documents were reviewed, information regarding each of the 20 projects was combined and provided in an evaluation package which was prepared for release in the fall of 1974.

For four of the Impact cities, (Denver, Newark, Portland, and St. Louis), less than half of city projects were reviewed due to the absence of documentation specifically earmarked for evaluation purposes. In the case of Denver, 43 percent of the projects which were not reviewed by MITRE staff had been funded for less than one year (see Table XVI), and as such could not be expected to have been subject to a full-scale evaluation. Additionally, the Denver effort was structured to produce and disseminate quarterly monitoring

reports written by project directors. These reports, designed to provide information for use in nine-month interim evaluation reports prepared by the CAT evaluation staff, were not included in an accounting of the evaluation documentation received from Denver. While not specifically labeled as providing evaluative information, the dissemination of these monitoring reports to MITRE staff nonetheless reflects an unusual effort to provide some type of information for the majority of funded projects.

Newark's low participation (33 percent) was largely the result of evaluation management problems in the CAT office, further complicated by the turnover of key evaluation personnel. Because of this, Newark has fallen drastically behind in the fulfillment of LEAA evaluation requirements. Additionally, one-third of Newark Impact projects not submitting evaluation documents have been funded for less than one year.

There are several reasons for the relatively small number of evaluation documents reviewed from Portland. First, Portland projects were subject to substantial delays which resulted in slow project start-ups. Second, the Portland evaluation effort cannot be characterized as an on-going process with incremental evaluation reporting. Evaluators for Portland's projects apparently chose to wait until comprehensive information was available before providing documentation on project activities and outcomes. This reporting style, in conjunction with slow project start-up, provides an explanation for Portland's low proportion of projects for which evaluation documentation was reviewed.

In St. Louis, organizational difficulties, especially the resignation of members of the Impact Evaluation Unit, at a critical time in the evaluation effort, may have contributed to the low proportion of projects available for review.

Finally, Atlanta and Dallas are noteworthy in that their proportions of funded projects with documentation available for review corresponds approximately to the average observed for the cities as a whole. Project evaluation documentation in both cities was not a one-shot effort; interim and quarterly reports were disseminated prior to completion of comprehensive annual reports. Additionally, 40 percent of the Atlanta projects not submitting evaluation documentation have been funded for less than one year. One-third of the non-documented projects in Dallas were similarly classified.

The previous discussion provides an indication of the progress each of the cities made in preparing and submitting evaluative information for use by policy-makers and decision-makers outside of the immediate local context. The type of information conveyed for those projects which were documented and reviewed from each of the cities is discussed below.

#### 5.2.2 Type of Reviewed Reports

Table XVII permits a comparison of the type of evaluation documents reviewed for each city. Based on the typology developed to classify the evaluation documentation of Impact projects according to the kind and amount of information provided (see Section 4.2, page 21), four cities were noteworthy for submitting high proportions of full-fledged evaluation reports (see Table XVII). More than 75 percent of the projects reviewed from Atlanta, Denver, Portland and St. Louis provided documentation that was considered to warrant classification as full-fledged evaluation. This means that project documentation in these cities typically included information regarding the extent to which project objectives were being met. Thus, questions dealing with project outcomes (e.g., changes in reported crime rates, recidivism reductions), were posed and at least partially answered in almost all of the reports reviewed from these four cities.

TABLE XVII  
TYPE OF EVALUATION DOCUMENTATION REVIEWED  
FOR EACH IMPACT CITY

IMPACT CITY	N	TYPE OF REPORT			
		STATUS	PROGRESS	PRELIMINARY	FULL-FLEDGED EVALUATION
ATLANTA	10			10% (1)	90% (9)
BALTIMORE	20	15% (3)	20% (4)	40% (8)	25% (5)
CLEVELAND	30	7% (2)	23% (7)	13% (4)	57% (17)
DALLAS	11	9% (1)	27% (3)	18% (3)	46% (5)
DENVER	14		14% (2)	7% (1)	79% (11)
NEWARK	9	33% (3)	45% (4)	22% (2)	
PORTLAND	6		17% (1)		83% (5)
ST. LOUIS	19	5% (1)	16% (3)		79% (15)
TOTAL	119	9% (10)	20% (24)	15% (18)	56% (67)

Specifically, Portland and Atlanta are noteworthy because virtually all of the documentation submitted by these two cities qualified as full-fledged evaluation reports. There is a common aspect in the evaluation effort in these two cities which may have colored the type of reports submitted. Both placed the primary responsibility for outcome evaluation with organizations not directly involved with the day-to-day concerns of Impact projects.

Of the remaining four cities, Cleveland and Dallas are strikingly similar with respect to the type of evaluation documentation submitted and reviewed by MITRE staff. As shown in Table XVII, in both of these cities over half of the projects provided documentation containing at least some preliminary outcome information, with a greater proportion of the projects (57 percent and 46 percent respectively) with documentation classified as full-fledged evaluation reports.

Full-fledged evaluation reports were not characteristic of the type of documentation reviewed for Baltimore and Newark. These two cities do, however, differ in several important respects. Documentation submitted by Baltimore runs the gamut from status to full-fledged evaluation reports. It is not surprising to find 40 percent (8 out of 20 projects) of these projects with documentation classified as preliminary evaluation reports, since they were designed primarily to meet the needs of the SPA in its project refunding process and not specifically to fulfill the LEAA project-level evaluation requirement. The decision was apparently made to postpone the collection and analysis of detailed outcome information. In the meantime, evaluation efforts have been limited to the compilation of information needed to generally portray which projects merit refunding. While this typically included some information about the outcome of the projects, in only a few instances (25 percent) did it contain enough outcome information to warrant classification as full-fledged rather than preliminary evaluation reports.

Finally, Newark is unusual for its poor performance in providing either preliminary or full-fledged evaluative information (see Table XVII). No full-fledged evaluation reports have been produced and documentation provided from only 2 (or 22 percent) of the projects reviewed could even be considered as preliminary evaluation reports.

### 5.2.3 Evaluation Reporting Quality

In many instances, the primary source of information regarding the success of Impact projects is the project evaluation report. With the production of these reports rests the responsibility to provide a comprehensive and realistic picture of project strengths and weaknesses, thereby providing the opportunity for the widespread dissemination and use of valid evaluative information. The performance of each of the Impact cities in the fulfillment of this responsibility is detailed below.

Variations in evaluation reporting quality are observed when the distribution of reports falling into the various reporting quality levels are broken down according to the Impact city responsible for the production of individual project evaluation documentation. Table XVIII facilitates this comparison of evaluation reporting quality among the eight Impact cities.

Two of the Impact cities stand out for having produced top-notch evaluation reports. Portland evaluators warrant particular attention for the scope and thoroughness of their evaluation reporting efforts. With one exception (a status report detailing problems encountered in evaluating Project Picture), all of the Portland reports were extremely well written, provided sufficient activity information,

TABLE XVIII  
EVALUATION REPORTING QUALITY BY IMPACT CITY

CITY	N	EVALUATION REPORTING QUALITY PERCENT IN EACH QUALITY LEVEL				MEAN QUALITY $\bar{X}$
		1	2	3	4	
ATLANTA	10	0% (0)	60% (6)	30% (3)	10% (1)	2.5
BALTIMORE	20	15% (3)	65% (13)	20% (4)	0% (0)	2.1
CLEVELAND	30	37% (11)	50% (15)	13% (4)	0% (0)	1.8
DALLAS	11	10% (1)	45% (5)	45% (5)	0% (0)	2.4
DENVER	14	0% (0)	29% (4)	57% (8)	14% (2)	2.9
NEWARK	9	67% (6)	11% (1)	22% (2)	0% (0)	1.6
PORTLAND	6	17% (1)	0% (0)	0% (0)	83% (5)	3.5
ST. LOUIS	19	32% (6)	26% (5)	32% (6)	10% (2)	2.2
TOTAL	119	23.5% (28)	41.1% (49)	27% (32)	8.4% (10)	2.2

cited a host of considerations crucial to a fair and unbiased interpretation of reported findings, and included detailed tables containing data upon which reported findings were based. In short, the six Portland evaluation reports reviewed were judged to be excellent transmitters of evaluative information.

Also noteworthy are the reports for the fourteen Denver Impact projects which were technically reviewed. Denver evaluators did an excellent job of providing detailed activity information, thus providing the reader with a context in which to interpret reported findings. Limitations were cited for the majority of the projects reviewed (71 percent): data needed to get a feel for the validity of the reported findings was included for only 14 percent of the Denver projects for which documentation was reviewed. This low percentage is in part a result of the fact that for many of the Denver project evaluations MITRE reviewers found the findings relied upon recidivism rates derived from a city-wide recidivism study, which were not fully presented nor explained in individual evaluation reports. Although the MITRE staff had access to the Denver document which carefully lays out the design of the recidivism study and the manner in which expected recidivism rates were derived, the quality of the Denver reporting effort was assessed from the perspective of a more general audience which may not have had ready access to this document. While each evaluation report could not be expected to provide a reiteration of the recidivism study design and results, each report would have been greatly enhanced by providing brief descriptions of the sample design and method for determining baseline recidivism rates.

Among the remaining Impact cities, project evaluation documentation reviewed from Atlanta, Dallas and St. Louis was of reasonably good quality. In each of these cities, over 40 percent of the projects for which documents were reviewed included limitations

important to a fair interpretation of reported findings. Only a few of the project reporting efforts in Atlanta and St. Louis (10 percent each), however, included data needed for reader validation.

Evaluation reporting efforts in Baltimore, Cleveland, and Newark were less impressive. The relatively poor reporting effort reflected in documentation produced by Baltimore and Newark is partially the result of the scarcity of full-fledged or preliminary evaluations conducted in these cities (see Table XVII). This was not true, however of Cleveland evaluators who produced a high percentage of full-fledged evaluation reports. Cleveland's relatively poor performance was in large part a result of the fact that the documentation reviewed failed to meet the criteria used to assess evaluation reporting quality (see Section 3.2.1).

#### 5.2.4 Approach Applicability

Table XIX permits a comparison of mean evaluation approach applicability scores for the eight Impact cities. Based on documents reviewed by MITRE staff, three cities stand out for having used rigorous evaluation approaches for gathering evaluative information on project outcomes. Portland with a mean score of 2.7 was considered to have used the most applicable evaluation approaches, but it must be noted that this figure is based on documentation from only six reviewed projects. Atlanta and Denver are remarkably similar, with mean scores (2.2 and 2.1 respectively) reflecting a reliance upon evaluation approaches which typically included at least two comparable data points in their assessments of project success. Denver, however, presented documentation for 14 projects as opposed to Atlanta's 10. Among the remaining cities, Newark stands out with a mean of only 1.0--representing a failure to develop a data base needed to compare observed project outcomes.

Due to the relationship observed between evaluation approach applicability and project focus (see Section 5.1.4, page 43), it is necessary to look more closely at the mean scores calculated for each



TABLE XIX  
EVALUATION APPROACH APPLICABILITY BY  
IMPACT CITY

IMPACT CITY	N	EVALUATION APPROACH APPLICABILITY PERCENT IN EACH LEVEL			MEAN SCORE $\bar{X}$
		1	2	3	
ATLANTA	10	10% (1)	60% (6)	30% (3)	2.2
BALTIMORE	20	75% (15)	20% (4)	5% (1)	1.3
CLEVELAND	30	70% (21)	27% (8)	3% (1)	1.3
DALLAS	11	55% (6)	36% (4)	9% (1)	1.5
DENVER	14	14% (2)	58% (8)	28% (4)	2.1
NEWARK	9	100% (9)	0% (0)	0% (0)	1.0
PORTLAND	6	16% (1)	0% (0)	84% (5)	2.7
ST LOUIS	19	47% (9)	43% (8)	10% (2)	1.6
TOTAL	119	54% (64)	32% (38)	14% (17)	1.6

city. It may be that these city differences are, in actuality, a reflection of the type of projects for which documentation was reviewed rather than a reflection of real city differences. To determine the actual source of these observed variations, mean scores were calculated for each city, adjusting for the effects of project focus. These scores represent what the expected city performance would have been if project focus bore no relationship to the applicability of the evaluation approach used. City expectations are arrived at by taking the percent of a city's reviewed projects of a particular focus and weighting this percentage by the program-wide mean for that focus (see Table XV, page 45). An adjusted mean score for each city is obtained by summing these three weighted figures. The formula for this calculation is:

$$e\bar{X}_i = \%R_i \bar{X}_R + \%CR_i \bar{X}_{CR} + \%SO_i \bar{X}_{SO}$$

where:

subscript: i indicates the Impact city for which calculations are made;

$e\bar{X}$  = expected mean score;

$\bar{X}$  = program-wide mean;

R = recidivism-reduction projects;

CR = crime-reduction projects;

SO = systems/other projects.

The results of these calculations for each Impact city are presented in Table XX.

Table XX also compares actual mean scores for each Impact city to the calculated expected mean score for each city (adjusting for the effects of project focus). Portland, Denver and Atlanta all performed substantially better than would have been expected given the particular mix of projects for which evaluation documentation was reviewed. On the other hand, three cities (Baltimore, Cleveland, and Newark) performed somewhat less well than might be expected in light of the type

TABLE XX

COMPARISON OF EXPECTED AND ACTUAL MEAN APPROACH  
APPLICABILITY SCORES BY IMPACT CITY

IMPACT CITY	N	PROJECT FOCUS			EXPECTED X SCORE	ACTUAL X SCORE	ACTUAL DIFFER- ENCE
		PERCENT CRIME REDUC- TION	PERCENT RECIDI- VISM RE- DUCTION	PERCENT SYSTEMS/ OTHER			
ATLANTA	10	40%	50%	10%	1.70	2.2	+.50
BALTIMORE	20	15%	50%	35%	1.58	1.3	-.28
CLEVELAND	30	10%	63%	27%	1.59	1.33	-.26
DALLAS	11	18%	9%	73%	1.49	1.54	+.05
DENVER	14	14%	72%	14%	1.63	2.14	+.51
NEWARK	9	22%	67%	11%	1.66	1.0	-.66
PORTLAND	6	33%	33%	33%	1.61	2.68	+1.07
ST. LOUIS	19	26%	42%	32%	1.61	1.63	+.02
TOTAL	119	19.3%	51.3%	29.4%			

NOTE: The expected mean score is calculated by multiplying column entries by the program-wide mean for each of the 3 project foci (crime-reduction multiplier = 1.87, recidivism-reduction = 1.63, systems/other = 1.38 as shown in Table XV, page 45).

(i.e., focus) of the projects for which documents were reviewed. Again, Newark fell lower than the other seven cities, even when adjustments were made for project focus. Finally, St. Louis and Dallas were noteworthy for the consistency observed between actual and expected mean scores.

City differences between expected and actual mean scores, both in direction and magnitude, suggest that the influence of project focus on approach applicability does not override inherent city differences in the selection and use of project evaluation approaches. City rankings based on expected mean scores are virtually identical to those based on actual mean scores. Only two sets of cities (Denver and Atlanta, St. Louis and Dallas) shifted in their relative standings. In both of these cases, these shifts reflected slight differences in mean scores (.03 in both cases).

The preceding analysis suggests that the Impact cities differed greatly in their handling of Impact project evaluation responsibilities. Similarly, variations were observed when projects of differing foci were examined in terms of the type of evaluation approach used to identify project outcomes. In the next section, several key elements of the Impact program concept will be examined to provide a context for viewing the results presented in the preceding sections. Following this discussion, two additional aspects of evaluation reporting will be examined, again with an eye toward providing an understanding of the nature of the Impact program project evaluation effort.

6.0 ANALYSIS OF STAGES OF THE EVALUATION PROCESS

Project-level evaluation in the Impact program is being performed as part of what has come to be called the COPIE-cycle. Key elements in this cycle--Crime-Oriented Planning, Implementation and Evaluation--were envisioned as operating in conjunction with one another, working to foster a more rational and comprehensive approach for solving the crime problems facing urban areas. The extent to which the COPIE-cycle could actually be implemented and could reap anticipated benefits is therefore a major focal point in MITRE's overall assessment of the Impact program.

At this point in time, data are presently available from this assessment of Impact project evaluation efforts to examine interrelationships among several key elements in the cycle. Specifically, this section presents the results of an analysis of the interrelationships observed, on a project basis, among the two aspects of the COPIE-cycle dealing specifically with project-level evaluation: (a) evaluation planning and (b) the conduct of project-level evaluation, particularly the quality of the evaluation reporting effort and the applicability of selected evaluation approaches.

Before looking at the relationship between evaluation planning and evaluation reporting/approach applicability, it is important to understand the relationship between these two later aspects of the evaluative process which have been the focus of this document. Table XXI facilitates this understanding by presenting relevant percentage distributions and mean quality scores. These figures indicate that evaluation reporting quality and approach applicability are clearly related to one another in a positive, symmetrical fashion. That is, project documents that qualified for the highest levels of reporting quality typically used the most applicable evaluation approaches. Thus, documents presenting "substantiated information"

TABLE XXI  
EVALUATION REPORTING QUALITY BY EVALUATION  
APPROACH APPLICABILITY

EVALUATION REPORTING QUALITY	N	EVALUATION APPROACH QUALITY			MEAN QUALITY $\bar{X}$
		1	2	3	
1	28	93% (26)	7% (2)	0% (0)	1.1
2	49	57% (28)	41% (20)	2% (1)	1.5
3	32	31% (10)	41% (13)	28% (9)	2.0
4	10	0% (0)	30% (3)	70% (7)	2.7
TOTAL	119	54% (64)	32% (38)	14% (17)	1.6
MEAN QUALITY $\bar{X}$		1.8	2.5	3.4	2.2

(level 4 in the evaluation reporting typology) have a mean approach applicability score of 2.7 out of a possible 3.0. At the same time reviewed documents containing the most rigorous evaluation approaches (level 3) were typically viewed to be the highest quality evaluation reporting effort, with a mean of 3.4 out of a possible 4.0.

Next, relationships between evaluation planning quality and both evaluation reporting and approach applicability are examined. The quality of project-level evaluation plans prepared by Impact evaluators was assessed and documented by MITRE staff in February 1975.<sup>5</sup> Guiding this assessment process was a model of the project-level evaluation planning process. This model detailed key steps in developing a plan for collecting and analyzing data needed to assess project achievements. Drawing upon the individual steps in the model, a typology was developed to characterize various levels of evaluation planning quality. Four such levels were defined, ranging from no overall plan for subsequent evaluation, to a plan which would encourage an evaluation effort capable of providing outcome information which could be linked back to the activities of the project.

For the purposes of the present analysis, this typology has been collapsed to represent two general types of project evaluation plans. The first type represents only minimal planning. Included in this group are those plans which failed to lay out a comprehensive approach for collecting and analyzing data to assess project success. While some of these plans presented a general idea of what would be looked at in the evaluation, they nonetheless stopped short of operationally defining major outcome measures and specifying the intended evaluation approach.

<sup>5</sup> See MITRE Technical Report MTR-6981, "An Analysis of Project-Level Evaluation Plans," dated April 1975

The second type is constituted by those evaluation plans which presented a clear idea of how changes anticipated in the targeted crime problem would be identified and possibly linked to the project's activities. These plans are viewed in this analysis as representing a viable strategy for subsequent project evaluation.

Additionally, eleven documented projects which did not have reviewed evaluation plans were included in this analysis. These projects have necessarily been evaluated in the absence of a documented, previously conceived plan for assessing the success of the project in meeting its activity and outcome objectives. (The existence of such documentation is not to be entirely ruled out, however, simply because it was not submitted to the LEAA or forwarded to The MITRE Corporation.)

Tables XXII and XXIII provide insights into the overall relationship between project-level evaluation planning and the quality of subsequent evaluation efforts. Both evaluation reporting quality and evaluation approach applicability appear to be positively associated with more comprehensive evaluation planning. Looking more closely at Table XXII, it appears that the likelihood of using a more rigorous evaluation approach is greater for those projects which had comprehensive evaluation plans than for those projects with little or no evaluation planning. Specifically, 59 of 68 (87 percent) of those projects documented in the absence of an evaluation approach were also characterized by minimal or no evaluation planning. This does not mean, however, that rigorous evaluation could not be conducted in the absence of such plans. Seven of the 16 projects (43 percent) which were evaluated using what was considered a rigorous evaluation approach (approach level 3 in Table XXII) did not have what were considered to be comprehensive evaluation plans ("no plan submitted" or "minimal plan"). In fact, Table XXII indicates that at least one project was considered to be rigorously evaluated without having submitted documentation outlining plans for subsequent project evaluation.

TABLE XXII  
EVALUATION PLANNING QUALITY BY EVALUATION  
APPROACH APPLICABILITY

EVALUATION PLANNING QUALITY	N	EVALUATION APPROACH APPLICABILITY			MEAN QUALITY $\bar{X}$
		1	2	3	
NO PLAN SUBMITTED	11	82% (9)	9% (1)	9% (1)	1.3
MINIMAL PLAN	78	64% (50)	28% (22)	8% (6)	1.5
COMPREHEN- SIVE PLAN	30	30% (9)	40% (12)	40% (9)	2.0
TOTAL	119	57% (68)	29% (35)	13% (16)	1.6

Looking at Table XXIII, it is clear that good evaluation reporting is related to comprehensive evaluation planning. Specifically, 90 percent (9 out of 10) of the reviewed projects with documentation classified as providing substantiated information (level 4 in the evaluation reporting typology) were also viewed as having comprehensive evaluation plans. The single exception to this was a Portland project which fell into this highest reporting category in the absence of a documented evaluation plan. Similarly, documentation from 3 projects were classified as providing explanatory information (level 3 in the evaluation reporting typology) in the absence of such plans, resulting in a higher mean reporting quality score for projects without documented and reviewed evaluation plans than for those projects with only minimal evaluation plans ( $\bar{X} = 2.4$  and  $\bar{X} = 2.0$ , respectively).

These findings suggest that while the two aspects of the COPIE-cycle dealing with project-level evaluation are highly related to one another, a host of other factors such as the caliber and turnover of the evaluation staff and the management of the evaluation effort, are operative, and may be more directly related in some instances to the actual evaluation performance than is the use of previous planning documents.

TABLE XXIII  
EVALUATION PLANNING QUALITY AND EVALUATION  
REPORTING QUALITY

EVALUATION PLANNING QUALITY	N	EVALUATION REPORTING QUALITY				MEAN QUALITY ( $\bar{X}$ )
		1	2	3	4	
NO PLAN SUBMITTED	11	9% (1)	55% (6)	27% (3)	9% (1)	2.4
MINIMAL PLAN	78	28% (22)	49% (38)	23% (18)	0% (0)	2.0
COMPREHEN- SIVE PLAN	30	17% (5)	17% (5)	37% (11)	30% (9)	3.1
TOTAL	119	23.5% (28)	41% (49)	27% (32)	8.4% (10)	2.2

## 7.0 ADDITIONAL ASPECTS OF EVALUATION REPORTING

### 7.1 Presence of Limitations or Qualification of Findings

Evaluation in the Impact program was intended to provide new insights and knowledge regarding project activities and their effectiveness in dealing with targeted crime problems. While precise knowledge about the relationship between project activities and observed outcomes was typically not attained, evaluation, nonetheless, informed decision-makers on what specific projects were doing, why, at what cost, and with what measurable benefits.

Providing this information in a responsible fashion requires, however, the candid reporting of limitations encountered in the process of identifying and assessing relationships between observed outcomes and project activities. For instance, using a simple before/after approach to gauge project success inherently leaves room for differing interpretations regarding the degree to which observed changes are due to the operation of the project. Other limitations which are not inherent to specific evaluation approaches but are, nonetheless, crucial to a proper interpretation of reported findings also exist. Studies often have deficiencies in conceptualization, measurement, analytical techniques, and data, or they may have complexities which limit their applicability, especially to decision-makers. The reporting of these limitations, like those inherent in the use of less rigorous evaluation approaches, is essential in aiding decision-makers to make choices on a more responsible basis. From the evaluation documentation reviewed to date, it appears that Impact project evaluators effectively see evaluation reporting as including the responsibility to explicitly state limitations so that reported results may be properly interpreted and applied.



## 7.2 Description of Reported Limitations

As indicated earlier, the provision of limitations or qualifiers to help the reader of an evaluation report interpret the significance of project findings is an important aspect of the evaluation reporting process. The types of limitations contained in the evaluation materials reviewed for Impact projects are presented below. This presentation provides an overall sense of what project-level evaluators felt necessary to include in evaluation documentation, and hence, a sense of the difficulties and problems encountered in the evaluation effort. Unfortunately, without more detailed information about project-level activities, little more can be said concerning the decision-making process involved.

In the course of the review process, reports were read and limitations explicitly stated in these documents were recorded. Unlike the previous analysis, the recording of these data was not contingent upon the presence or absence of evaluative findings in the reviewed documents. In fact, of the 34 projects previously mentioned as not providing findings in terms of the major objectives of the project, 19 nevertheless explicitly reported in evaluation documentation limitations (problems) which were encountered in the evaluation process. For the subsequent analysis, limitations reported in documents received from these 19 projects will be included with those 74 projects for which findings were provided and for which limitations were reported. For each project reviewed, a maximum of 3 reported limitations were coded for this analysis. In cases where more than 3 limitations were recorded, a judgment was made as to which limitations were most significant with respect to the evaluation effort.

Of the 119 projects reviewed, 34 projects produced only status or progress type reports and therefore did not pose possible interpretive problems (i.e., findings were not addressed). This left 85 projects with documentation presenting preliminary or full-fledged evaluation findings which could be interpreted and used by practitioners, policy-makers, and researchers interested in discerning project success. Of these 85 projects, 74 (87 percent) illuminated findings by mentioning or accounting for limitations. Additionally, 19 of the 34 projects reviewed which did not provide findings in terms of outcome (or intermediate) objectives also took the opportunity to report limitations concerning evaluation efforts. Thus, a total of 93 projects, or 78 percent of the 119 projects reviewed reported limitations.

Some differences were noted when documentation was examined by project focus. All of the crime-reduction projects for which findings were reported also reported important limitations to the interpretation of findings. Only 80 percent of the recidivism-reduction and 89 percent of the systems/other projects provided similar information with which to temper results presented. Additionally, it should be noted that systems/other focused projects had the highest incidence of projects (46 percent) with documentation not providing findings. Since many of these projects did not provide for evaluation beyond the achievement of activity objectives, the applicability of the reported limitations question is moot for many of these cases.

A look across the Impact cities (Table XXIV) shows that cities differed in the proportion of reviewed projects with documentation that did not provide findings as well as in the proportion of projects citing limitations. However, 80 percent or more of the projects which provided findings in each Impact city, included within the evaluation documentation explicit limitations with respect to interpretation of the findings.

TABLE XXIV  
 PRESENCE OF LIMITATIONS IN IMPACT EVALUATION  
 BY CITY DOCUMENTATION

CITY	TOTAL RECEIVED	FINDINGS PRESENT	LIMITA- TIONS PRESENT	PERCENT TOTAL	PERCENT OF THOSE PROJECTS WITH FINDINGS
ATLANTA	10	10	10	100%	(100%)
BALTIMORE	20	13	11	55%	(85%)
CLEVELAND	30	21	17	57%	(81%)
DALLAS	11	7	6	55%	(86%)
DENVER	14	12	11	79%	(92%)
NEWARK	9	2	2	22%	(100%)
PORTLAND	6	5	5	83%	(100%)
ST. LOUIS	19	15	12	63%	(80%)
TOTAL	119	85	74	62%	(87%)

Table XXV provides an accounting of the limitations reported in materials received from the 119 reviewed projects (93 of which reported some evaluation limitations). Not surprisingly, 54 percent of all limitations reported in the evaluation documentation involved data constraints. Typically this included statements in the text of evaluation documents concerning the availability or quantity of data with which to measure the achievement of a major objective of a project. Of the data limitations reported, over 50 percent (58 of 110) related to this predicament--one common to evaluators. The prevalence of this problem is not surprising considering the large number of Impact projects for which evaluation is conducted by some group other than project staff.

Another common data problem/limitation recorded in the MITRE review related to the reliability of data. The reporting of data reliability difficulties commonly concerned problems with the use of data from what the evaluators felt to be unreliable sources (for instance, the use of notes from a project director's notebook to indicate whether clients participating in a rehabilitation program had been rearrested). Thirty-nine of all 110 data limitations (19 percent of all limitations) reported were of this nature.

The final general category of data limitations mentioned in the reviewed reports related to problems encountered with the unit of data utilized in measuring project achievement. Commonly these limitations dealt with data being too general rather than applying to client- or area-specific measurement. Thirteen of all 110 data limitations reported were of this type.

Problems with the design approach accounted for 33 percent (67 of 208) of all the limitations reported. Nine of the 67 design problems related to the lack of a comparison base. Other design-related limitations frequently found in the reviewed documents included

TABLE XXV  
BREAKDOWN OF REPORTED LIMITATIONS

TYPE OF LIMITATION	TYPE OF FOCUS			
	TOTAL	CRIME REDUCTION	RECIDIVISM	SYSTEMS/ OTHER
DATA PROBLEMS	53% (110)	44% (22)	57% (60)	54% (28)
AVAILABILITY OR QUANTITY	28% (58)	24% (12)	31% (33)	25% (13)
RELIABILITY	19% (39)	14% (7)	19% (20)	23% (12)
UNIT OF MEASURE- MENT OR ANALYSIS	6% (13)	6% (3)	7% (7)	6% (3)
DESIGN APPROACH	33% (67)	53% (27)	28% (29)	22% (11)
NO BASIS FOR COMPARISON	4% (9)	----	7% (7)	4% (2)
COMPARISON, CON- TROL GROUP/AREA DIFFICULTIES	14% (28)	20% (10)	13% (14)	8% (4)
CONFOUNDING EFFECTS PROJECT ATTRIBUTION TRENDS MATURA- TION, ETC.	15% (30)	33% (17)	8% (8)	10% (5)
MEASURES	8% (18)	2% (2)	8% (8)	15% (8)
OBJECTIVES	4% (9)	1% (1)	3% (3)	10% (5)
MEASURES	4% (9)	1% (1)	5% (5)	6% (3)
PROJECT	6% (13)	1% (1)	7% (7)	10% (5)
TOTAL LIMITATIONS REPORTED	100% (208)	100% (52)	100% (104)	100% (52)
TOTAL PROJECTS REVIEWED	119	23	61	35

inadequacies in the comparative aspects of the approach (28 of 67) and inability to attribute results in terms of project outcome to project activities, due to other influences or confounding effects (30 of 67).

Difficulties concerning project objectives and measures selected to bridge the gap between an objective, and the data required to determine achievement of that objective, accounted for only 8 percent (18 of 208) of the reported limitations. Reporting of limitations concerning objectives (9 of 18) always referred to the fact that the objectives cited for the project were either inappropriate, realistically unachievable or arbitrarily quantified. Likewise, limitations reported concerning measures (9 of 18) typically dealt with the fact that they (a) lacked precise operational definition, or (b) were not sensitive in revealing the true nature and extent of changes which the project may have affected.

"Project" limitations were typically related to the nature of the project (including problems of implementation), making an evaluation design infeasible. Also included in this category are those reports which stated that it was too early in the project's life to present any evaluation findings. Six percent or 13 of all 208 reported limitations were of this type.

Project focus, as indicated in Table XXV, played a role in the nature of the limitations reported in the reviewed evaluation documentation. The prevalence of "design approach" limitations among crime-reduction projects (53 percent) compared to recidivism-reduction (28 percent) and systems/other (22 percent) projects is particularly notable. The greatest part of this difference can be explained by the fact that 33 percent (67 of 208) of those limitations cited in the documentation reviewed for crime-reduction projects concerned the tempering of findings with the admission that attribution of changes to project

activities was not possible or within the scope of project evaluations. On the other hand, limitations reported for recidivism-reduction and systems/other focused projects rarely involved issues of outcome attribution (8 percent and 10 percent, respectively) focusing instead upon data problems. This, in part, reflects differences in the existence and accessibility of data needed to evaluate projects of the three foci.

Crime-reduction projects could typically rely upon data from established sources (i.e., police department records, FBI Uniform Crime Reports) to measure changes in the occurrence of specific types of crime over time. Additionally, projects of this type were often implemented by police departments, thus reducing the need to create additional lines of communication between project operators and data sources. As a result, evaluators of crime-reduction projects were generally not confronted with massive data gaps and were thus in a position to concentrate upon the meaning of observed measurements and their comparison.

This was not the case for recidivism-reduction and systems/other projects. These types of projects reported limitations dealing with data more frequently than crime-reduction projects (57 percent and 54 percent, as compared to 44 percent for crime-reduction projects). The dependency of projects of these foci upon data other than previously established crime rates appears to be the major stumbling block for evaluators responsible for these projects. In the case of recidivism-reduction projects, a multiplicity of data problems existed ranging from definitional difficulties to information gaps attributable to staff turnover. Determining what constitutes recidivism, how long it should be measured, and who should be responsible for collecting client follow-up data, all contributed to the difficulties encountered in evaluating recidivism-reduction projects. For instance, a project

established to provide intensive supervision to youthful offenders in most instances had to assume responsibility for the collection of information concerning the behavior of its clients while they participated in the project. Lacking the previous experience and time needed to coordinate possible data sources and collect data, the limitations reported by such a project thus reflected the inability to focus energies on the measurement and meaning of project outcomes. These problems might be further exacerbated if project evaluators sought to gauge the more important long-term impact of project activities through the collection of client follow-up data.

A relatively high frequency of limitations concerning data (54 percent of all reported) was also evident in the systems/other projects reviewed. Like recidivism-reduction projects, these projects also could not generally rely upon established data sources for the raw data needed for project-level evaluation.

Preoccupation with data problems/limitations in recidivism-reduction and systems/other projects may then account for the relatively infrequent mention of limitations dealing specifically with the design approaches used to evaluate projects. This is further substantiated (as shown in Table XXIV) by the fact that none of the reviewed crime-reduction projects reported the absence of a basis for comparison.

A look across cities failed to show any discernible patterns. In all but one city, data limitations constituted the most frequently reported limitation with reported design limitations always following in frequency. Only in Portland, where only six projects were reviewed, was this order reversed.

This review of the Impact evaluation reports, in terms of reported limitations, reveals that evaluators frequently reported with

**CONTINUED**

**1 OF 2**

activities was not possible or within the scope of project evaluations. On the other hand, limitations reported for recidivism-reduction and systems/other focused projects rarely involved issues of outcome attribution (8 percent and 10 percent, respectively) focusing instead upon data problems. This, in part, reflects differences in the existence and accessibility of data needed to evaluate projects of the three foci.

Crime-reduction projects could typically rely upon data from established sources (i.e., police department records, FBI Uniform Crime Reports) to measure changes in the occurrence of specific types of crime over time. Additionally, projects of this type were often implemented by police departments, thus reducing the need to create additional lines of communication between project operators and data sources. As a result, evaluators of crime-reduction projects were generally not confronted with massive data gaps and were thus in a position to concentrate upon the meaning of observed measurements and their comparison.

This was not the case for recidivism-reduction and systems/other projects. These types of projects reported limitations dealing with data more frequently than crime-reduction projects (57 percent and 54 percent, as compared to 44 percent for crime-reduction projects). The dependency of projects of these foci upon data other than previously established crime rates appears to be the major stumbling block for evaluators responsible for these projects. In the case of recidivism-reduction projects, a multiplicity of data problems existed ranging from definitional difficulties to information gaps attributable to staff turnover. Determining what constitutes recidivism, how long it should be measured, and who should be responsible for collecting client follow-up data, all contributed to the difficulties encountered in evaluating recidivism-reduction projects. For instance, a project

established to provide intensive supervision to youthful offenders in most instances had to assume responsibility for the collection of information concerning the behavior of its clients while they participated in the project. Lacking the previous experience and time needed to coordinate possible data sources and collect data, the limitations reported by such a project thus reflected the inability to focus energies on the measurement and meaning of project outcomes. These problems might be further exacerbated if project evaluators sought to gauge the more important long-term impact of project activities through the collection of client follow-up data.

A relatively high frequency of limitations concerning data (54 percent of all reported) was also evident in the systems/other projects reviewed. Like recidivism-reduction projects, these projects also could not generally rely upon established data sources for the raw data needed for project-level evaluation.

Preoccupation with data problems/limitations in recidivism-reduction and systems/other projects may then account for the relatively infrequent mention of limitations dealing specifically with the design approaches used to evaluate projects. This is further substantiated (as shown in Table XXIV) by the fact that none of the reviewed crime-reduction projects reported the absence of a basis for comparison.

A look across cities failed to show any discernible patterns. In all but one city, data limitations constituted the most frequently reported limitation with reported design limitations always following in frequency. Only in Portland, where only six projects were reviewed, was this order reversed.

This review of the Impact evaluation reports, in terms of reported limitations, reveals that evaluators frequently reported with



candor limitations which might affect findings and that the reported limitations most often concerned either the data or design aspects of the evaluation process.

### 7.3 Operational Problems and Recommendations

In addition to serving as a vehicle for the reporting of findings, evaluation documentation may be used to report operational problems and recommendations for improving projects. When this is done, evaluation reports can assist decision-makers by providing them with a range of information on project operations, effectiveness and costs. Additionally, the citing of operational problems and associated recommendations increases the utility of the evaluation report for project directors interested in replicating or expanding the project. The remainder of this section will be devoted to a discussion, therefore, of the problems and recommendations cited in Impact project evaluation documentation.

#### 7.3.1 Reporting of Operational Problems

Impact project evaluation documents were examined to determine the extent to which this evaluative function was fulfilled. An examination of reports submitted on the 119 assessed Impact projects indicated that 79 percent (94) of the projects contain information regarding operational problems. While it had been expected that evaluation documentation prepared by project staffs would communicate operational problems to a greater extent than would reports prepared by other agencies (Crime Analysis Team, State Planning Agency, outside consultant, Crime Analysis Team/Project, Crime Analysis Team/Consultant, and other agency), this was not the case. The lack of difference between the two groups of evaluators in using reports as a vehicle to pinpoint problems indicates that evaluators generally tended to communicate the operational problems encountered by projects to outside audiences regardless of allegiance or affiliation.

#### 7.3.2 Description of Reported Problems

What then were the problems project-level evaluators considered worth mentioning? For each project reviewed, a maximum of two reported operational problems have been coded for this analysis. In cases where more than two problems were reported, judgments were made as to the relative importance of these problems in affecting project performance.

Table XXVI provides a distribution of operational problems according to the frequency with which they were mentioned by projects in their reports. Two types of problems predominate, accounting for over half of the 149 operational problems cited in reviewed documentation. Problems encountered in staffing the project or training staff members to perform project activities were most frequently mentioned (35 percent or 52 of all 149 problems cited in reviewed documentation). Such problems included difficulties encountered in filling staff vacancies due to civil service requirements; the inability to attract qualified personnel due to low salaries, and conflicts experienced in providing extensive staff training in conjunction with the delivery of project services. This finding suggests that debilitating staffing problems may surface as a result of the failure to adequately consider personnel requirements in advance of project implementation.

Developing lines of communication essential to the operation of projects relying upon support and assistance from other organizations frequently proved to be a challenge to project operators. Twenty-three percent (34 of 149) of the operational problems reported dealt with issues of this type. Included in this category were problems encountered in obtaining a sufficient flow of client referrals from outside sources, satisfying criteria guiding project participation, and building community support essential for the successful operation of crime prevention projects. This finding suggests a strong awareness on the part of project operators and evaluators of the need for and consequences of greater coordination among projects, outside agencies and the community.

TABLE XXVI  
OPERATIONAL PROBLEMS REPORTED

TYPES OF PROBLEMS	FREQUENCY OF MENTION	PERCENT OF TOTAL PROBLEMS PRESENTED
Funding; Excessive Bureaucracy	5	3%
Acquisition of Site and Equipment	22	15%
Personnel; Staffing and Training	52	35%
Intergovernmental; Community Coordinating; and Client Referral	34	23%
Internal Management	28	19%
Other	8	5%
TOTAL	149	100%

Evaluation documents rarely contained reference to operational problems stemming from funding procedures or delays resulting from bureaucratic excesses (5 of 149, or 3 percent). While such problems were reported as impeding project implementation,<sup>6</sup> they were apparently not considered to be appropriate topics for inclusion in reviewed documentation. This is not surprising since problems of this type are, for the most part, resolved before project evaluation can take place.

Problems more frequently reported dealt with the acquisition of project equipment or sites (15 percent). Difficulties specifically related to the procurement of equipment apparently plagued projects beyond the early implementation phase. This problem, which predominated in an earlier analysis of reasons for project implementation delays, was also of concern to evaluators. Finally, internal management problems, as could be expected in the operation of any organization, also received attention in reviewed documentation. Twenty-eight of the 149 problems (19 percent) recorded in the reviewed evaluation documentation cited such general management concerns.

### 7.3.3 Reporting of Operational Recommendations

Evaluators of 55 percent of the 119 assessed projects utilized reports as a vehicle for presenting recommendations for improvement of project operations as opposed to 79 percent presenting operational problems. Thus, more evaluators tended to present problems than to provide recommendations addressing those problems. Table XXVII shows little variation in the frequency of reporting recommendations by city. All but two cities (Baltimore and Portland) provided recommendations for between 56 and 68 percent of their projects. Baltimore, however,

<sup>6</sup>See MITRE Technical Report MTR-6961, "An Assessment of the Implementation Quality of High Impact Anti-Crime Projects", August 1975.

presented recommendations in only 30 percent of its projects whereas Portland gave recommendations for only 17 percent (1 of 6) of its projects.

Recommendations were provided to a somewhat greater extent in reports of crime-reduction projects than recidivism-reduction and systems/other projects. As shown in Table XXVIII, 70 percent of crime-reduction projects presented recommendations in their reports in contrast to 52 percent of recidivism-reduction projects and 49 percent of systems/other projects.

As differentiated from the documentation of operational problems, reports written by project staffs on project performance did present recommendations with greater frequency than those prepared by other agencies. Two-thirds of the projects (14 of 21) that were self-evaluated presented recommendations in their reports in contrast to 52 percent (42 of the 82) of those projects evaluated by outside agencies.<sup>7</sup> Project staffs are involved in the day-to-day operations of projects, and can be expected to be more aware than outside evaluators of needed improvements in project operations, and have a greater incentive to report and push for such improvements. (This involvement, however, may be accompanied by a lessening of objectivity.)

Finally, an examination of the type of operational recommendations reported revealed that the substance of these recommendations typically corresponded to reported operational problems or consisted solely of general statements urging project refunding. For example, suggestions regarding new or streamlined client intake procedures were prevalent in those documents citing client referral problems. Likewise, projects

<sup>7</sup> These do not add to 119 since 16 projects had either shared responsibility with an outside agency or the arrangements were not reported.

TABLE XXVII  
PRESENCE OF RECOMMENDATIONS IN REPORTS OF  
ASSESSED PROJECTS BY CITY

	ATLANTA		BALTIMORE		CLEVELAND		DALLAS		DENVER		NEWARK		PORTLAND		ST. LOUIS	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
RECOMMENDATIONS REPORTED	6	60	6	30	17	57	7	64	9	64	5	56	1	17	13	68
RECOMMENDATIONS NOT REPORTED	4	40	14	70	13	43	4	36	5	36	4	44	5	83	6	32
Total	10	100	20	100	30	100	11	100	14	100	9	100	6	100	19	100

N = 119 Projects

TABLE XXVIII

PRESENCE OF RECOMMENDATIONS IN  
PROJECT REPORTS BY PROJECT FOCUS

	TOTAL		CRIME REDUCTION		RECIDIVISM REDUCTION		SYSTEM/ OTHER	
	N	%	N	%	N	%	N	%
RECOMMENDATIONS REPORTED	65	55	16	70	32	52	17	49
RECOMMENDATIONS NOT REPORTED	54	45	7	30	29	48	18	51
Total	119	100	23	100	61	100	35	100

encountering staffing delays due to civil service requirements often recommended that these requirements be modified. While operational recommendations were presented for about half of the projects reviewed, these recommendations were generally not a logical extension of the evaluation results nor specific enough to adequately inform those responsible for resolving project difficulties.

In sum, the majority of reviewed project documentation addressed questions dealing with operational problems and recommendations (79 percent and 50 percent respectively) as well as with issues related to the proper interpretation of reported findings (62 percent). The inclusion of such information by project evaluators suggests an awareness of and sensitivity to the diversity of needs which may be fulfilled by project evaluation. In short, these efforts provide us with a body of information which adds to and refines existing knowledge in the area of criminal justice.

#### 8.0 SUMMARY AND CONCLUSIONS: PROJECT-LEVEL EVALUATION IN THE IMPACT PROGRAM

Understanding what may realistically be expected from project-level evaluation efforts, both in terms of quantity and quality, is vital to the formulation and implementation of realistic evaluation policies and guidelines. In previous sections of this document, the results of an analysis of evaluation documentation reviewed for 119 Impact program projects were presented in an effort to further this understanding. Specifically, results of these analyses focus upon the number, type, and quality of Impact project evaluation efforts. Further, variations in evaluation reporting quality and approach applicability were detailed for each of the eight Impact cities and for the differing types of projects designed and implemented as part of the program. While projects in the Impact program are still on-going and thus subject to further evaluation, the results of these analyses, nonetheless, provide several useful insights.

First, outcome-oriented project evaluation is both a realistic expectation and feasible within the context of an action-oriented program such as Impact. Documentation specifically earmarked for evaluative purposes was received for 140 or 60 percent of the projects funded with Impact monies as of July 1, 1975. Additionally, questions dealing with project outcomes were posed and at least partially answered for 72 percent of the 119 documented projects that were reviewed.

The manner in which this information was gathered and presented precluded, in about half the cases, critical assessments of the reliability and validity of reported findings. In these cases, documentation failed to cite limitations regarding the interpretation of findings and to include data needed to assess the face validity of such findings. Similarly, shortcomings were noted in the approaches used to gather reported information for 54 percent of the projects. These had been evaluated in the absence of some standard or comparison base against

which the direction and magnitude of observed outcomes could be interpreted. Additionally, rigorous evaluation was conducted for only 14 percent of the projects included in this analysis.

Weaknesses in the evaluation approaches used to gauge project outcomes necessarily cast doubts upon the credibility of reported findings. Where evaluation documentation explicitly provided findings concerning the attainment of project outcome objectives (85 of 119 projects) the credibility of these claims, regardless of their nature, was viewed as questionable for almost half of the projects. This finding further reflects the tremendous difficulty Impact evaluators had in providing findings about project outcomes based on the type of evaluation approaches used and data presented in reviewed documentation.

This lack of face validity suggests that project evaluations performed in the Impact program were typically not powerful enough to permit the attribution of observed outcomes to project activities. In light of this, the successes and failures observed among Impact projects must be critically reviewed on a project-by-project basis before applying the results in other contexts or locales. It remains to be seen whether or not the extent and type of documentation produced and disseminated via project evaluation reports changes as projects progress and approach the termination of Impact funding. Significant changes seem doubtful, however, given the fact that 75 percent of the projects not documented and available for MITRE review have already been funded for over one year and thus could have been expected to provide at least preliminary evaluative information. Additionally, staffing and management problems which have impeded evaluation efforts in several cities are not likely to be resolved in the remaining days of the Impact program.

Our findings suggest that it is somewhat misleading to treat the Impact program project-level evaluation effort as a single entity. Little homogeneity was found among the eight Impact cities in terms of either the amount, type or quality of project evaluation conducted. For instance, the type of information presented regarding project activities and outcomes varied considerably among the cities. For four of the cities (Atlanta, Denver, Portland, and St. Louis) better than 75 percent of the projects reviewed had documentation classified as full-fledged evaluation reports. In contrast, no full-fledged evaluation reports were received from Newark.

Variations in evaluation reporting quality and approach applicability were also noted. Documentation from three cities (Atlanta, Denver, and Portland) was consistently viewed as reflecting good evaluation reporting and as including outcome information generated using relatively rigorous evaluation approaches. On the other hand, evaluation documentation from Newark, Cleveland and Baltimore was less impressive when viewed in terms of the manner in which information was gathered and reported.

The heterogeneity observed among evaluation documents produced by the eight Impact cities has implications for evaluation policy. In the absence of standard procedures for organizing, staffing, performing and documenting the evaluation of project-level activities it is unrealistic to expect uniformity in either the quantity or quality of project evaluation efforts. While the LEAA provided some guidelines and examples for use in the development of project-level evaluation plans, similar guidelines were not disseminated regarding the actual collection and reporting of evaluative information. As a consequence, opportunities to encourage greater uniformity (and therefore comparability) in the type of project information collected were not used to full advantage in the Impact program.

The type of project to be evaluated also appears to have colored several aspects of the Impact project evaluation effort. While the conduct of rigorous evaluation was evidently feasible for all three types of projects examined (e.g., crime-reduction, recidivism-reduction and systems/other) the identification and use of bona fide evaluation approaches was less prevalent among projects designed to reduce recidivism or improve the functioning of the criminal justice system. This finding suggests that in the absence of a standardized, regularly-updated data base such as that available to measure changes in reported crime rates (almost exclusively used in crime-reduction focused project evaluations) project evaluation efforts are less likely to produce information addressing changes occurring in conjunction with and/or due to project activities. Thus it needs to be understood (and integrated into evaluation plans) that recidivism-reduction and systems/other focused projects must create their own baseline data (where none exists) or establish a new data base in order to measure project effects.

Variations in the type of information provided in reviewed documentation also varied by project focus. The majority of crime-reduction focused projects were viewed as having full-fledged evaluation reports (90 percent were either full-fledged or preliminary evaluation reports). Documentation for recidivism-reduction and systems/other projects were typically less oriented toward the provision of project outcome information, with a greater proportion of reports being status or progress reports. Again, data availability and ease of collection may be partly responsible for these differences. In effect (and unsurprisingly), data problems were more frequently cited as limiting the interpretation of findings in reports for recidivism-reduction and systems/other projects than for crime-reduction projects.

Insights were also obtained regarding interrelationships among several aspects of the evaluative process. As expected, evaluation



reporting quality and approach applicability are positively related to one another. Additionally, findings presented reveal that those projects which prepared and disseminated comprehensive evaluation planning documents had a greater likelihood of being evaluated using a rigorous evaluation approach. Similarly, there was a greater likelihood that the information gathered for these projects would be transmitted via good evaluation reports. While these findings do not mean that sound evaluation did not occur in the absence of such planning efforts, it does suggest the need for comprehensive project evaluation planning and the corresponding utility of encouraging such planning in similar evaluation efforts.

Finally, our findings suggest that the majority of evaluation documents reviewed contained essential ingredients for use as an aid to decision-making and to the improvement of project operations. Limitations crucial to the interpretation of reported findings were reported in documentation for a full 62 percent of all the reviewed projects; this includes those projects (34 of 119) for which findings in terms of outcome (or intermediate) objectives were not provided. Additionally, 79 percent of the project documents reviewed discussed at least one project operational problem, while recommendations designed to improve project operations were cited in 54 percent of the reviewed documents. Thus, Impact project evaluations were viewed, and perhaps even utilized, as vehicles for the improvement of project operation.

## APPENDIX I

PROJECT-LEVEL EVALUATION TECHNICAL REVIEW  
DATA COLLECTION FORMPROJECT CODE 

PROJECT NAME

CITY

PROJECT OVERVIEW

I. PROJECT DESCRIPTION

PROJECT FOCUS

PROJECT OBJECTIVES

Outcome

(1)

(2)

Intermediate

(1)

(2)

Activity

(1)

(2)

GENERAL DESCRIPTION OF AREA SERVED

DESCRIPTION OF PROJECT CLIENTS

CITY

1 = Crime Reduction

2 = Recidivism

3 = Systems/Other

☐

INDIVIDUAL OR AGENCY INITIATING PROJECT

AGENCY RESPONSIBLE

PROJECT DIRECTOR

II. FUNDING AND IMPLEMENTATION

RESOURCES

Requested	<input type="text"/>
Received	<input type="text"/>
Time Period Covered (Months)	<input type="text"/>
Allocation of Resources	
% of Personnel	<input type="text"/>
% of Equipment	<input type="text"/>

DATES

Submission of grant application	<input type="text"/>
Hiring of Project Director	<input type="text"/>
Date of award	<input type="text"/>
Initial provision of services (e.g., first client received or first deployment of manpower)	<input type="text"/>
Refunding award date	<input type="text"/>
End of refund award period	<input type="text"/>
Suggested changes for improved implementation.	<input type="text"/>

Extent and Scope of CAT,  
SPA, RO Assistance

Yes = 1  
No = 0

CAT	<input type="text"/>
SPA	<input type="text"/>
RO	<input type="text"/>

IMPLEMENTATION PROBLEMS

CODE

Primary	<input type="text"/>
Secondary	<input type="text"/>

TURNOVER

Project Director      Yes = 1  
Supervising Staff      No = 0

Professional Staff	<input type="text"/>
Support Staff	<input type="text"/>

CRIME-ORIENTED PLANNING

0 = No Data To Support Problem  
1 = Data Alluded to But Not Cited  
2 = Data Substantiated But is General  
3 = Data Specific To Area  
4 = Data Links Activities To Problem Solution

<input type="text"/>
----------------------

III. EVALUATION PLANNING

A. Provisions made for conducting evaluation.

- (1) Automated/manual data collection and management system.

Yes = 1

☐

No = 0
- (2) Standardized forms.

☐
- (3) Reporting periods.

☐

Frequency (Most frequent) Not Specified = 0

Monthly = 1

Quarterly = 2

Semi-Annually = 3

Annually = 4

☐
- (4) Number Evaluation Personnel

☐
- (5) Evaluation Responsibility

☐

- B. Evaluation Component

No Plan = 1

CODE

What = 2

How = 3

Linkages = 4

☐

- C. Evaluation Design

CODE

1 Baseline

2 Comparison

3 Control Area

4 Control Group

5 Comparison Area

6 Projection

12 Baseline/Comparison Group

13 Baseline/Comparison Area

14 Baseline/Control Group

15 Baseline/Control Area

16 Baseline/Projection

26 Comparison Group/Projection

88 Other

99 Not Applicable

☐

IV. PROJECT FINDINGS

JUDGE CODE

☐

Does the Evaluation Report provide findings in the terms of outcome/intermediate objectives.

Yes = 1

No = 0

☐

Did any of the following change from the Evaluation Component Review Form?

- Outcome

Yes = 1

No = 0

Measures = 8

N/A = 9

(1)

☐

(2)

☐
- Intermediate

(1)

☐

(2)

☐
- Activity

(1)

☐

(2)

☐

If yes discuss the nature of the change.

What are the major findings?

Does the report indicate either explicitly or implicitly that the project met its outcome objective(s)?

Yes = 2

Partly Met = 1

No = 0

If no, or partly met provide reasons.

P

S

Additional (side) Benefits/Drawbacks

V. EVALUATION REPORTING

Does the Evaluation Report provide activity data (what services the project provided) in specific enough terms to indicate what the project is about.

Yes = 1  
No = 0

Are the data upon which the findings are based presented in the report (i.e., is the reader in the position to validate)?

Yes = 1  
No = 0

Is the report readable and logically presented?

Unacceptable = 1  
Acceptable = 2  
Good = 3  
Excellent = 4

Is this report an:

Interim = 1  
Final = 2  
Phased = 3

Does the Evaluation Report account for limitations in the interpretation of findings?

Yes = 1  
No = 0

List important limitations.

Mentioned	Accounted For	P	S
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>		
<input type="checkbox"/>	<input type="checkbox"/>		
<input type="checkbox"/>	<input type="checkbox"/>		
<input type="checkbox"/>	<input type="checkbox"/>		

List any major limitations which should have been accounted for in the report (if applicable).

P ☐  
S ☐

Overall measure of Evaluation Report ☐



VI. EVALUATION APPROACH

What type of research design was used in the Evaluation Report?

- Before/After = (Time Frame)
- Projection = (Base)
- Comparison Group = (Specify)
- Comparison Area = (Specify)
- Control Group = (Specify Selection)
- Control Area = (Specify)
- Other

Specification

CODE

On a scale, rate the design approach in the context of the limitations of this specific project.

1  2  3  4  5

Low

High

In this context, what do you see as the major drawbacks of the approach which prevented you from giving the report a higher rating?

Given the drawbacks, do the findings based on outcome/intermediate objective(s) in the report appear to be justified?

Yes = 1   
No = 0

Operational Problem

P

S

Recommendations Reported

- Operational Recommendations

P

S

VII. DOES THE PROJECT APPEAR TO BE INNOVATIVE?

If yes, why?

Yes = 1 ☐  
No = 0

VIII. MITRE COMMENTS

APPENDIX II

IMPACT CITY	DESCRIPTION OF ALLOCATION OF IMPACT PROJECT-LEVEL EVALUATION RESPONSIBILITY
Atlanta	Crime Analysis Team (CAT) has responsibility for project-level evaluation; relies upon outside, university-based team of consultants for quantitative analysis of data and drafting of evaluation reports at regular intervals (e.g., six month and annual reports).
Baltimore	Evaluation responsibility split between Crime Analysis Team and State Planning Agency (SPA);* SPA has responsibility for the evaluation of projects implemented by state agencies; to date, both organizations have relied extensively upon inputs from operating projects in the preparation of evaluation documentation designed primarily for use in the yearly project refunding sessions.
Cleveland	Crime Analysis Team has responsibility for project-level evaluation; fulfilled through use of outside consultants located in the CAT office; project documentation produced en mass on a yearly or bi-yearly basis regardless of the length of time individual projects have been operational.
Dallas	Crime Analysis Team has responsibility for project evaluation; fulfilled through heavy reliance on quarterly reports submitted by operating projects.
Denver	Crime Analysis Team has responsibility for project evaluation; fulfilled through a reliance on CAT monitoring visits and reports generated by project personnel and evaluators contracted for by individual projects; interdependence between CAT and project reflected in CAT generated nine month reports and project generated annual reports.
Newark	Evaluation is a CAT responsibility with reliance upon project inputs for the preparation of evaluation reports.
Portland	The State Planning Agency** has sole responsibility for project evaluation; in-house evaluation capability handles data collection, analysis, and report writing; documentation not provided at regular nor within pre-determined time-frames.
St. Louis	Crime Analysis Team has primary responsibility for evaluation; in-house evaluation unit supplemented by creation and operation of evaluation unit within St. Louis Police Department; reports generated in accord with phased project funding periods (prior to extensive turnover among in-house personnel).

\* In Maryland, the State Planning Agency is formally called the Governor's Commission on Law Enforcement and the Administration of Justice.

\*\* The SPA in Oregon is formally referred to as the Oregon Law Enforcement Council (OLEC).

**END**

*7-22-55/111111*