#29

EVALUATORS AND DECISION-MAKERS: PERCEPTIONS OF THE EVALUATION PROCESS

by

Anne L. Schneider and Peter R. Schneider

January, 1977

INTRODUCTION

This paper reports on the needs assessment phase of the Model Evaluation Program in the State of Washington. The needs assessment was undertaken at the request of the Law and Justice Planning Office in order to assess the performance of the evaluation system and to recommend strategies for improving the evaluation process.
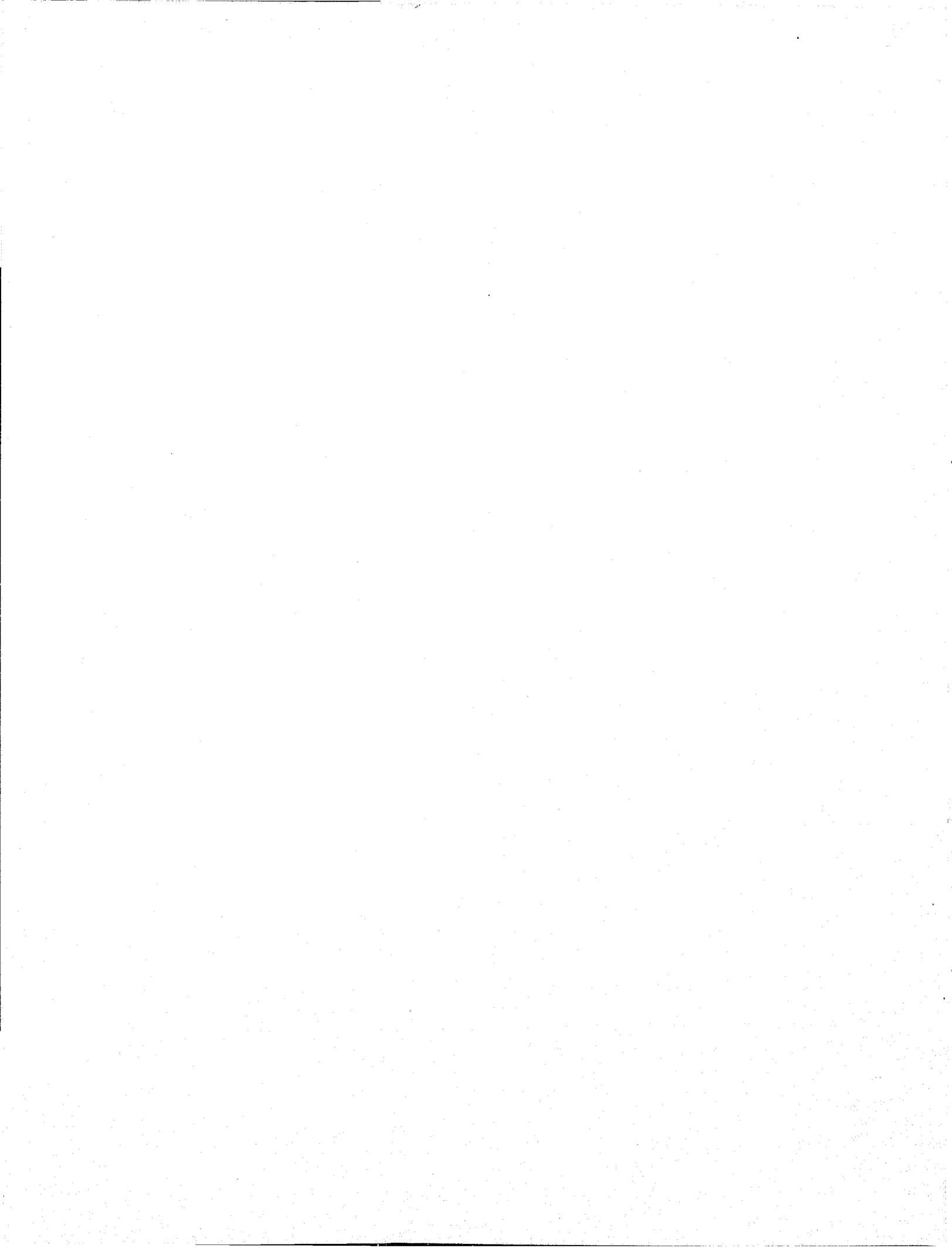
The major question which guided the eight-month study was as follows: What are the factors that constrain and/or facilitate the use of information from evaluation in the decision-making process? In order to answer this question we felt it was necessary first to visualize what an effective evaluation system would or should look like; second, to determine where, amidst the complex interrelated federal, state, regional and local agencies with responsibilities for law and justice planning and administration, the decisions are made that actually result in some action projects receiving funds and others not receiving them; third, to develop appropriate and adequate measures for the quality, amount of use and reliance on evaluation information; and fourth, to isolate specific problems evaluators face in their environment.

An overview of the methodology and results of the study are contained in this paper. It should be emphasized that the study was descriptive rather than hypothesis-testing, and that our results cannot be generalized to evaluation systems elsewhere. However, a number of propositions worthy of further consideration will be presented.

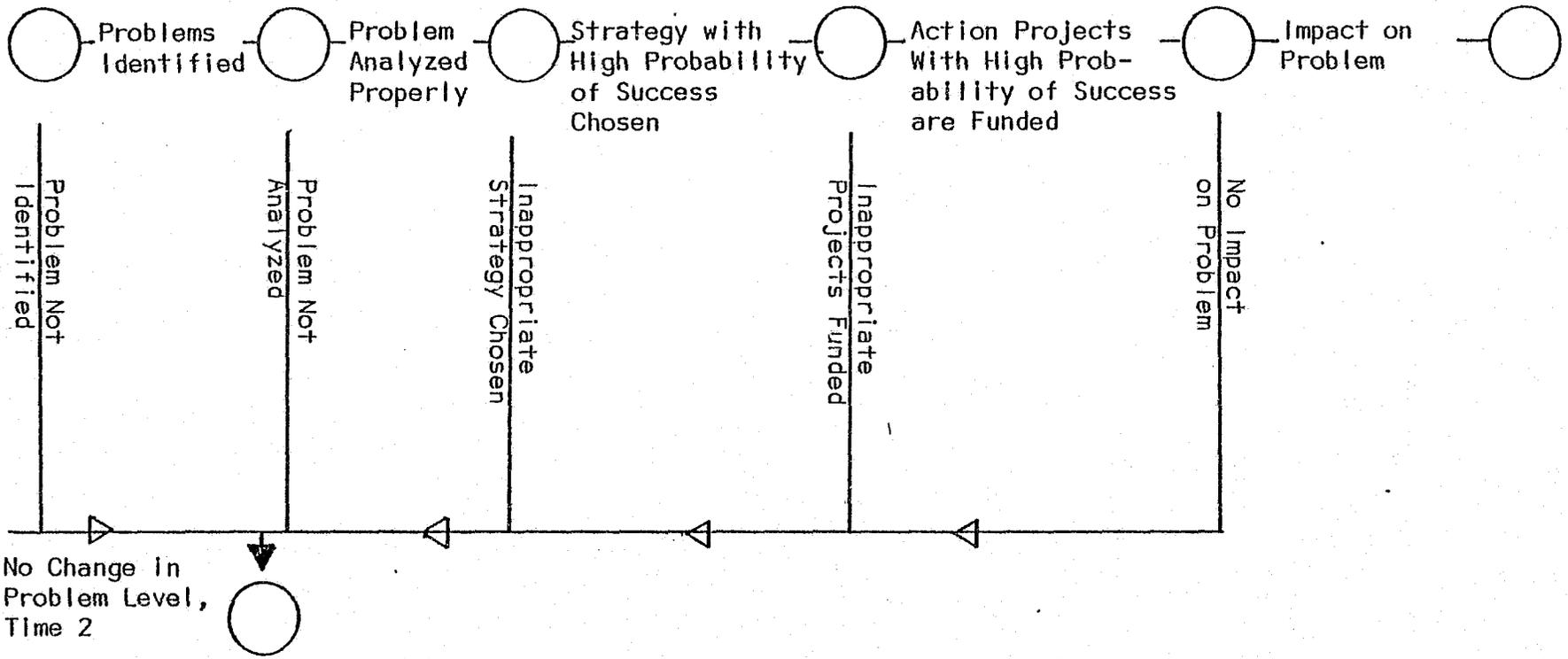## Role of Evaluation in Decision-Making: A Simple Model

Evaluation may be defined in terms of the type of question it seeks to answer and the procedures used in answering the question. The question asked by an evaluation often is phrased in the following way: What would (____) have been if (____) had been different, or if (____) had not existed at all? A common evaluation question in criminal justice, for example, is: What would the (crime/recidivism rate) have been if (some intervention) had been different, or if (some intervention) had not existed? The second part of the definition of evaluation is that the procedures used to answer the question(s) conform to generally recognized standards of social science research. Information produced by evaluation differs from many other types of information in that scientific procedures have been used to produce it and scientific standards are used to judge its accuracy.

A simple model of how the planning, decision-making and evaluation system should work is presented in Figure 1. The major role of evaluation is to measure the impact of action projects on problems in the community. Information from evaluations should include the estimated magnitude of impact on the problem, the probability that similar results would be achieved again, and the major factors that contributed to the success or failure of the project. Information from evaluations should then feed back into all parts of the system. These data are used to identify appropriate strategies and fund projects with a high probability of success in dealing with the problem. The system shown in the diagram focuses attention on the quality of evaluation reports as well as on the use of evaluations. Shortcomings in those areas increase the liklihood that inappropriate strategies will be selected or unsuccessful projects funded -- either of which means the problem is likely to continue unaffected.

Figure 1. Planning, Decision, Evaluation Process

Problems in
the Community,
Time 1

Fewer
Problems in
the Community,
Time 2

○—Problems   ○—Problem   ○—Strategy with   ○—Action Projects   ○—Impact on   ○
   Identified   Analyzed   High Probability   With High Prob-   Problem
                     Properly   of Success   ability of Success
                                      Chosen   are Funded

Problem Not Identified

Problem Not Analyzed

Inappropriate Strategy Chosen

Inappropriate Projects Funded

No Impact on Problem

No Change in
Problem Level,
Time 2

26

ω

The simplicity of the model evaluation system masks an enormously complex set of interlocking governmental units, agencies and individuals in the State of Washington who actually are responsible for identifying problems, choosing strategies, and allocating funds to specific projects. An additional task, therefore, was to determine where, among all participants, the effective decisions about project funding are made -- including who makes the decisions and what criteria are used to allocate scarce resources among the many persons and groups who request funding for specific projects.

THE DECISION-MAKING PROCESS

Information on decision-making procedures within the Washington law enforcement and criminal justice system was obtained through informal interviews with numerous persons in the state and intrastate regional agencies, documents prepared by the agencies, and a formal survey (mailed, with telephone follow-up) of major participants in the system. A complete description of the system would require far more space than may be allotted here; what follows is merely a summary.

As with all of LEAA's state planning agencies, the Law and Justice Planning Office is responsible for allocating LEAA funds to regional areas within the state and to other state agencies. The process through which funds are allocated is illustrated in Figure 2. Funds are received by the state on the basis of a crime population formula and, in turn, approximately 55 percent of these are allocated on a crime population formula basis to 21 regional areas. Over the last three years, approximately 45 percent of the funds have been allocated to other state agencies that have major responsibilities for criminal justice within the state.

Within the 21 regional areas, there are two types of funding procedures: Type A regions allocate funds directly to projects; Type B regions (a minority
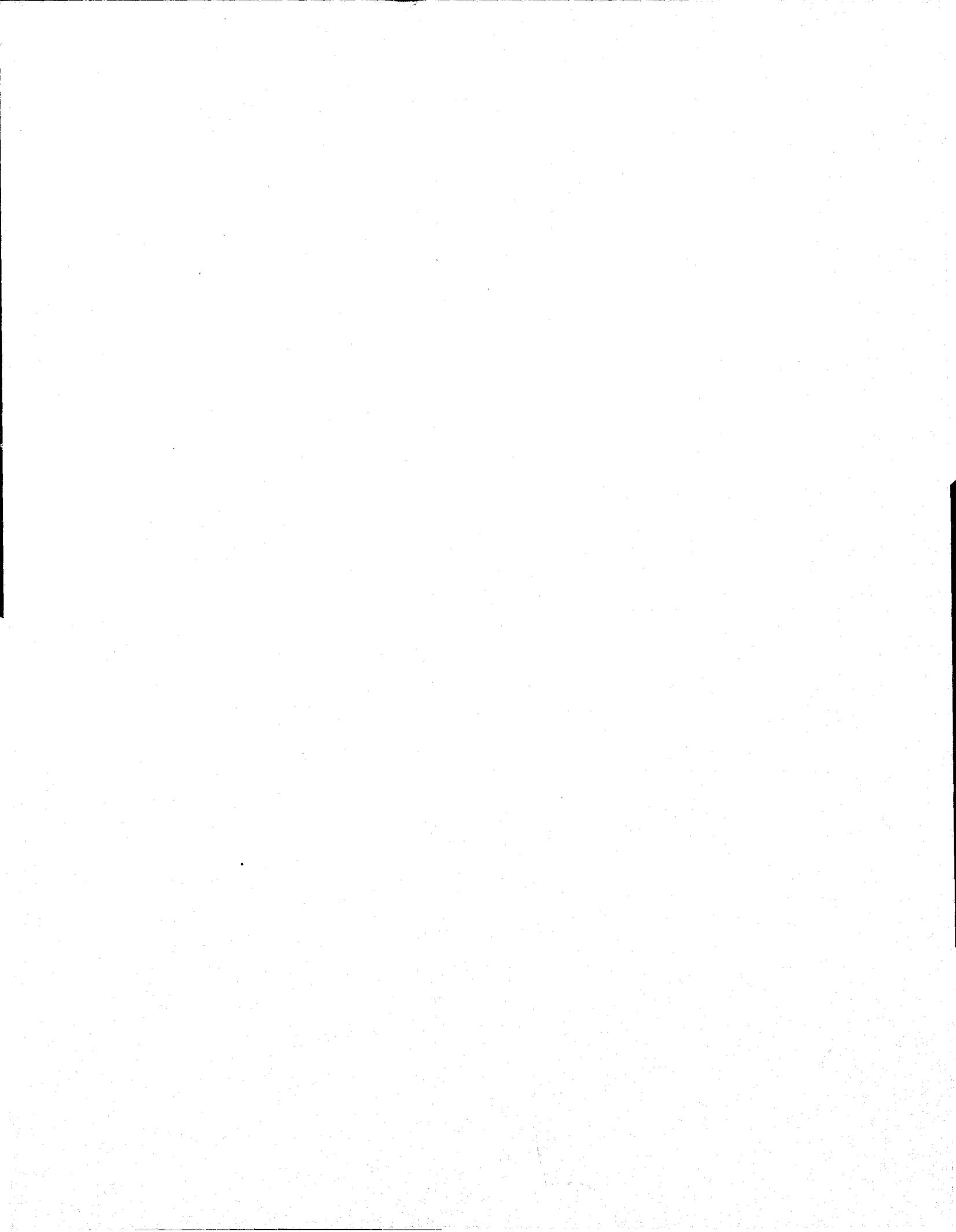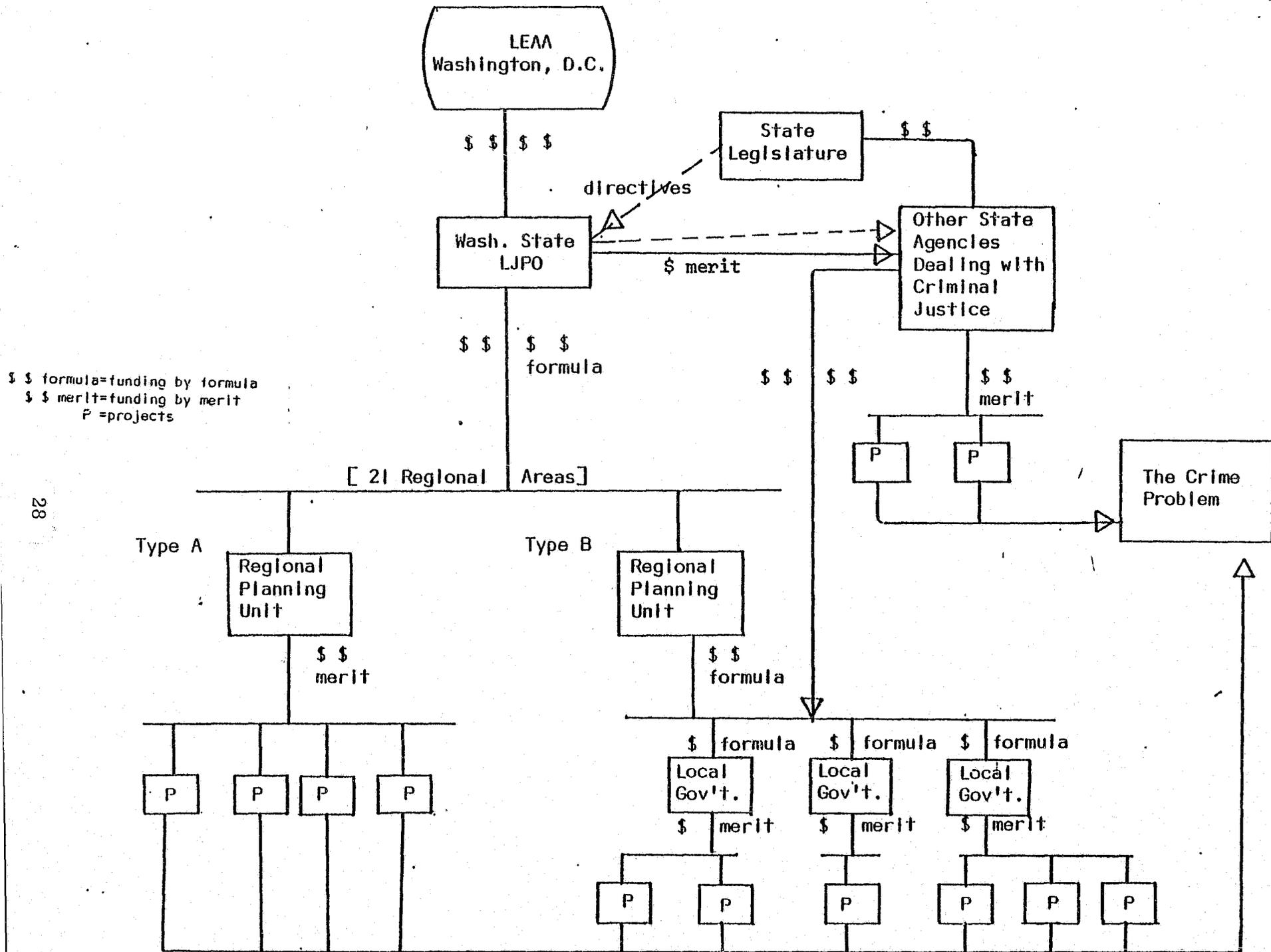
FIGURE 2

FIGURE 2

of the areas) allocate the money to local government units which, in turn, determine which projects will be funded and which will not. The planning guidelines specify how regional areas (and state agencies) are to make their decisions but they do not place many constraints on these decisions.

The guiding philosophy of the state planning office is that regional preferences should determine which projects will be funded, but that regional decisions should be reached on the basis of a highly rational analysis of problems, strategies, and projects. In order to measure the degree of decentralization in the system (and to identify the effective decision-making points) we obtained copies of the original requests for funds from each region and traced the review process through the entire state system.

Each region is informed, in advance, of how much money they can expect to receive. The planning guidelines specify that each region must perform a quantitative analysis of the crime problem, develop strategies for dealing with the problems, and then select the projects that have the greatest potential for ameliorating the most important problems. Each region is required to rank the projects it recommends in order of their priority. Recommendations from the regions are submitted to the state planning office for approval. The professional staff in the planning office conducts a project-by-project review and then submits its recommendations to the planning sub-committee of the Governor's committee. The Governor's committee is appointed by the Governor and is composed of criminal justice system officials, some elected officials, and some persons appointed from outside the system. After the sub-committee finishes its review, the full Governor's committee reviews the recommendations and makes a "final" recommendation concerning funding. Although this recommendation goes to LEAA and to the state legislature (for matching funds), there are almost no changes made beyond the Governor's committee.

Two measures were used to estimate the extent to which regional preferences prevailed in the final allocation of funds for 1977: The first is the extent to which projects that actually received funds were ranked above projects that did not receive money, and the second is the proportion of all regional projects which were judged "not acceptable," meaning they would not have been funded even if money were available.

Scores on the first indicator were developed by calculating the cut-off point for funded projects in the region and calculating the percentage of projects actually funded that were above that point. If the state, for example, funded three projects in a region and selected those ranked 1, and 3, then the score is 100 percent, indicating the state decisions complied perfectly with regional preferences as indicated by the ranks given to projects. If the state selected projects 1, 2, and 5 (or if they selected projects 2, 3, and 4), then in both instances they eliminated one preferred by the local area in favor of one lower-ranked project. The score would be 67 percent (two divided by three) since two of the three were within the top three ranked projects. From the data available, it does not appear as if any ranks were changed to eliminate a very expensive project ranked higher on the RPU list in order to fund several less expensive ones that were ranked lower.

Scores on the second indicator represent the proportion of projects that were listed as "not acceptable" for funding rather than as "not sufficient funds." The resulting score is an indication of the extent to which local preferences would have been overriden by the state if the LEAA money had been as high as the total requested amount from the regional planning units. Data from the analysis are presented in Table 1.

Of the projects eventually funded, 96.5 percent were in the same priority ranking recommended by the regional planning units. Conversely, one can say that less than four percent of the funded projects were ranked below a project that failed to receive funds. This analysis suggests that the original

TABLE 1: Regional Funding Requests and Determinations

Total Number of Projects Requested	225
Total Number of Projects Funded	116
Of Those Funded, Number that were ranked above all unfunded projects	112
Percent of adherences to Regional Preferences	96.5%
Total Number Judged "Not Acceptable" by Law and Justice Planning Office Staff	71
Percent Not Acceptable	32%
Number Judged "Not Acceptable" by Subcommittee	58
Percent Judged "Not Acceptable" by Subcommittee	26%
Number Judged "Not Acceptable" by Full Committee	58
Percent Judged "Not Acceptable" by Full Committee	26%

decisions concerning project ranks made at the regional level were the ones that eventually determined which projects would be funded and which would be eliminated. The highly decentralized nature of decisions pertaining to projects actually funded must be interpreted cautiously since the state decision-making procedures could have had considerably more impact if more funds had been available. The Law and Justice Planning Office staff rated 32 percent of the projects as unacceptable for funding even if the money had been available. Each of those decisions could have resulted in a lower-ranked project being funded. The planning sub-committee of the Governor's Committee changed thirteen of the "not acceptable" ratings to acceptable with conditions, but this still left 26 percent of the projects in the "not acceptable" category.

The evaluation system, initiated in 1974, also is decentralized. Provisions are made for fifteen positions: five at the state level, six in four urban/high crime regions, two in suburban regions, and one in a rural region. In the fourteen other regions, planners or project directors are responsible for evaluating LEAA-funded projects.

Most of the evaluation reports produced in 1976 came from the seven regions with evaluators. State evaluators did not produce any evaluation reports of regional or state projects. Their responsibilities, rather, are to coordinate evaluations, provide technical assistance -- especially in areas that do not have evaluators -- and synthesize information from the regional evaluators' reports. The last task is especially difficult given the variety of projects that were evaluated in the regions and the quality of the reports.

USE OF EVALUATION INFORMATION

Several types of data were collected in an effort to develop multiple estimates of the amount of use given to evaluation reports:

1. Surveys were conducted of regional and state planners, project directors, advisory planning committee members, and other professionals within the state LJPO. (See Appendix A for information on the sampling procedures, response rate, and other details of the methodology.) Each respondent was asked the following questions about the amount and type of use:

(a) The number of evaluation reports they became aware of during the 1976 planning process;

(b) The number of these which they reviewed, but did not use the information in their work; the number which contained information that they considered, and the number which played a major role in their work;

(c) The number of evaluation reports which the respondent used to argue for discontinuation of a project; lowered priority; continuation; or higher priority for a project;

(d) The number of evaluation reports that they used to argue for major changes in project operation.

Although these questions seem to ask for "factual" data, it should be recognized that responses are a mixture of facts and perceptions. Individuals cannot remember all instances of "use" and, especially at the higher levels of administration, an individual will have a difficult time recalling the source of different types of information. Nevertheless, there is no perfect way of reconstructing the "facts" concerning the amount of use by each of these respondents and the best recourse was to ask them.

2. In addition to asking the users of evaluation how many reports they used, the evaluators also were asked to estimate how many of their reports were used by each of the potential audiences: directors, planners (program managers and standards at the state level), project directors, and planning advisory committee members. These estimates by evaluators provide a second source of information on the extent to which evaluation reports were used by someone else in the system.

3. Another indicator of use is the perception of evaluators and all other persons in the survey concerning the importance of information from

evaluation compared with other sources and types of information. Respondents to the survey were asked to rate eight types of information on a scale ranging from one (least important source of information) to ten (most important source). Persons at the regional level were asked these questions in reference to the final decisions made by the planning advisory committee concerning which projects would be funded. Persons at the state level were asked to rate the importance of each type of information in final state-level decisions concerning which projects would be accepted or not accepted for funding.

4. The final indicators of use, available only at the state level, were the written reasons given for judging a project to be "not acceptable." These were summarized on the LJPO project review form. Since the major role of evaluation is to determine the effectiveness or efficiency of a project, reasons for rejection of a proposal were considered to be based on evaluation information if they referred to probable ineffectiveness, or evidence of prior poor performance/lack of impact as the reason for project rejection.

Recall and perceptions of the amount of use are presented in Tables 2 through 6. For these portions of the analysis, the respondents to the survey were divided in accordance with the type of position held and in relation to whether they worked at the state level, in a region which has an evaluation section within the planning agency (type 1 regions) or in regions that do not have an evaluation section. Because the number of respondents is quite small, inferences should not be made to a larger population although, in most instances, the respondents to the survey constitute a significant majority of all persons in the areas (see Appendix A).

Number of Reports

The data in Table 2 indicate that all agency officials in regions which have evaluators were aware of one or more evaluation reports whereas at the

TABLE 2   NUMBER OF EVALUATION REPORTS: USER ESTIMATES

	Total N	% Aware of one or more Reports	Avg. No. Reports, per person	Eval. Estimate of No. of Reports Available by by Area
State Planners, Directors, Others	11	74%	9	28
Type 1: Regional Planners, Directors, Others: (Eval.Regions)	11	100%	9.6	8
Type 2: Regional Planners, Directors, Others (Regions without evaluators)	9	56%	6	--
(Advisory) Planning Committee				
State (Gov.Committee)	9	11%	.22	28
Type 1 RPU	12	58%	4	8
Type 2 RPU	17	24%	1.6	--
Project Directors				
State	11	18%	.45	--
Type 1 Region	48	52%	.96	--
Type 2 Region	18	39%	1.0	--

TABLE 3   AMOUNT OF USE OF EVALUATION, AS RECALLED BY USERS[1]

User	No. Persons	PER CENT OF REPORTS THAT RESPONDENT SAID: were not used	were considered	played a major role
PLANNERS, DIRECTORS, OTHER PROFESSIONALS				
State LJPO	8	29%	45%	25%
Type 1 Regions (with evaluators)	11	11%	30%	59%
Type 2 Regions (no evaluators)	5	87%	5%	8%
PROJECT DIRECTORS				
State	(2)	(0)	(87)	(13)
Type 1 Regions	25	26%	43%	31%
Type 2 Regions	7	47	0	53
ADVISORY/PLANNING COMM.				
State, Gov.Committee	(1)	(0)	(0)	(100)
Type 1 Regions	7	0	6	94
Type 2 Regions	(4)	(25)	(40)	(35)

[1]Data are based only on the persons who indicated that they were aware of at least one evaluation report pertinent to their work. The number of cases is quite small in some instances because very few of the respondents were aware of any reports.
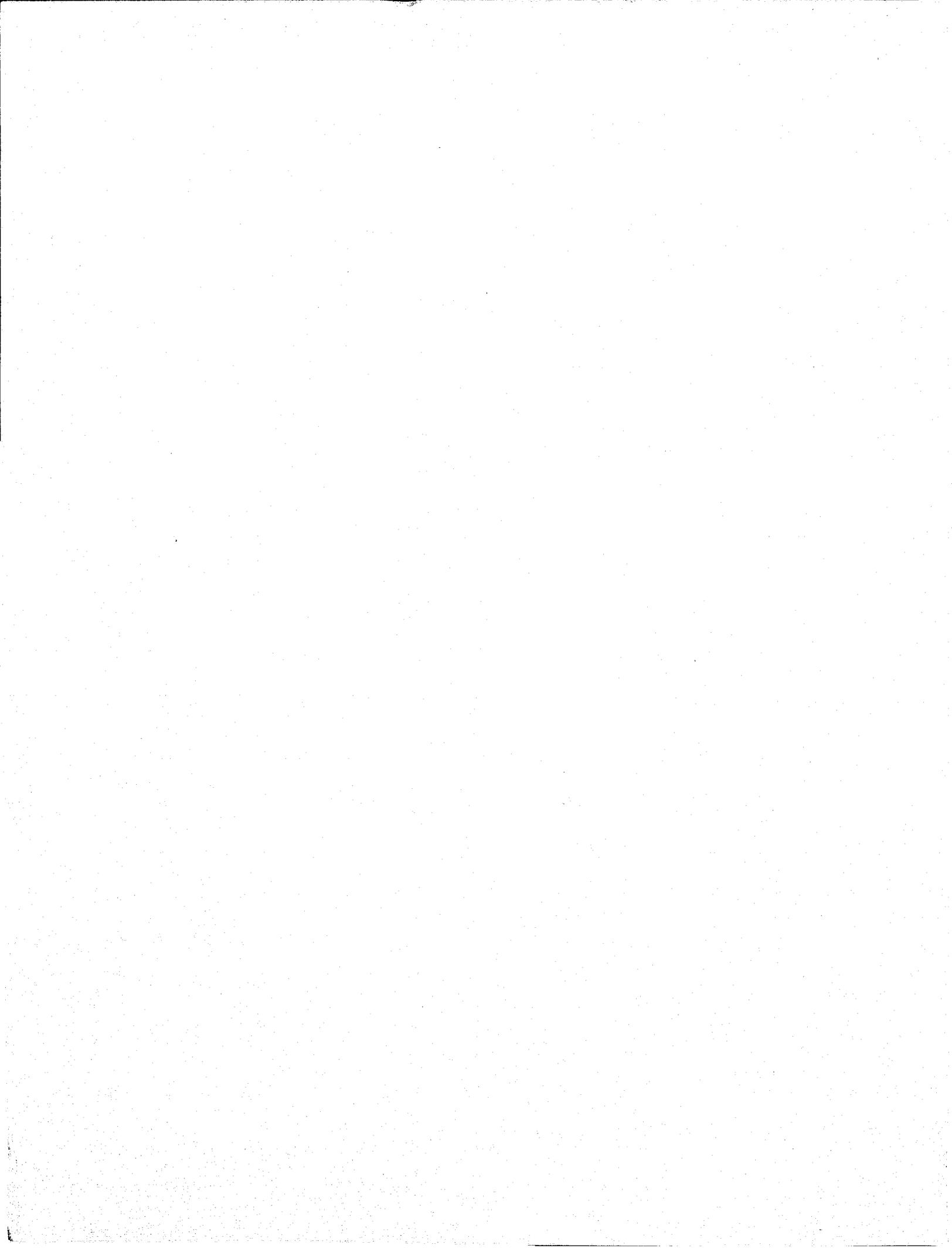
state level, 74 percent of the respondents said that they were aware of at least one report and 56 percent of the agency officials in regions without evaluators indicated they were aware of one or more reports. The same pattern of use is found for the advisory committees. In the regions with evaluators, 58 percent of the advisory committee members said they were aware of at least one report compared to 11 percent of the state-level committee respondents and 24 percent of the committee respondents from type 2 regions. Project directors indicate a similar pattern: 52 percent of those in the regions with evaluators said they reviewed one or more reports compared with 18 and 39 percent for the state level directors and type 2 regional project directors, respectively.

Use of Evaluation Information

Respondents who indicated that they had reviewed at least one evaluation report were asked to estimate the proportion of evaluation reports that were reviewed but not used; the proportion that were considered in their work or decisions; and the proportion of the reports that had a major role in their work or decisions. As shown in Table 3, agency officials in regions with evaluators estimated that a greater proportion of the reports had a major role than did either the state officials or the officials in regions without evaluators. There were only two state-level project directors and one person on the state Governor's Committee that had reviewed any evaluation reports. The degree of use estimates for these audiences are not a reliable indication for anyone except the small number of individuals involved.

Evaluator's Perception of Use

Evaluators at the state level estimated that 46 percent of the reports received by LJPO were reviewed by the LJPO professional staff, whereas the

evaluators in regions with an evaluation section estimated that 81 percent of their reports were reviewed by the regional planning agency (Table 4). In a similar way, state evaluators estimated that the Governor's advisory committee did not review or become aware of any reports whereas regional evaluators estimated that 62 percent of their reports were reviewed by the committee.

Table 4 also contains the number of reports, according to evaluator estimates, that were used by some potential audience and the extent to which these were utilized in decision making. State evaluators estimated that 67 percent of the reports which LJPO professional staff became aware of were not used, 28 percent were considered, and 6 percent played a major role in LJPO decisions. At the regional level, the estimates show that evaluators believed 31 percent of the reports which planners/directors became aware of were not used, 41 percent were considered and 29 percent played a major role.

Regional evaluators indicated that the degree of use of the advisory committees was about the same as for the planners/directors: 31 percent of the reports not used, 39 percent considered, and 31 percent having a major role. The majority of project directors, according to the evaluators, considered the reports, but only 17 percent played a major role in decision-making by the project and 19 percent were not used at all.

When the figures on extent of use (not used, considered, major role) are compared with the estimates of use recalled by the various audiences, it appears as if one of three things has happened:

(1) evaluators underestimate the extent of use, or

(2) evaluators are not aware of all the uses given to their reports, or

(3) users tended to overestimate the degree of use they gave to evaluation reports.

For example, the regional evaluators estimated that 29 percent of the reports played a major role in decisions made within the regional planning unit whereas
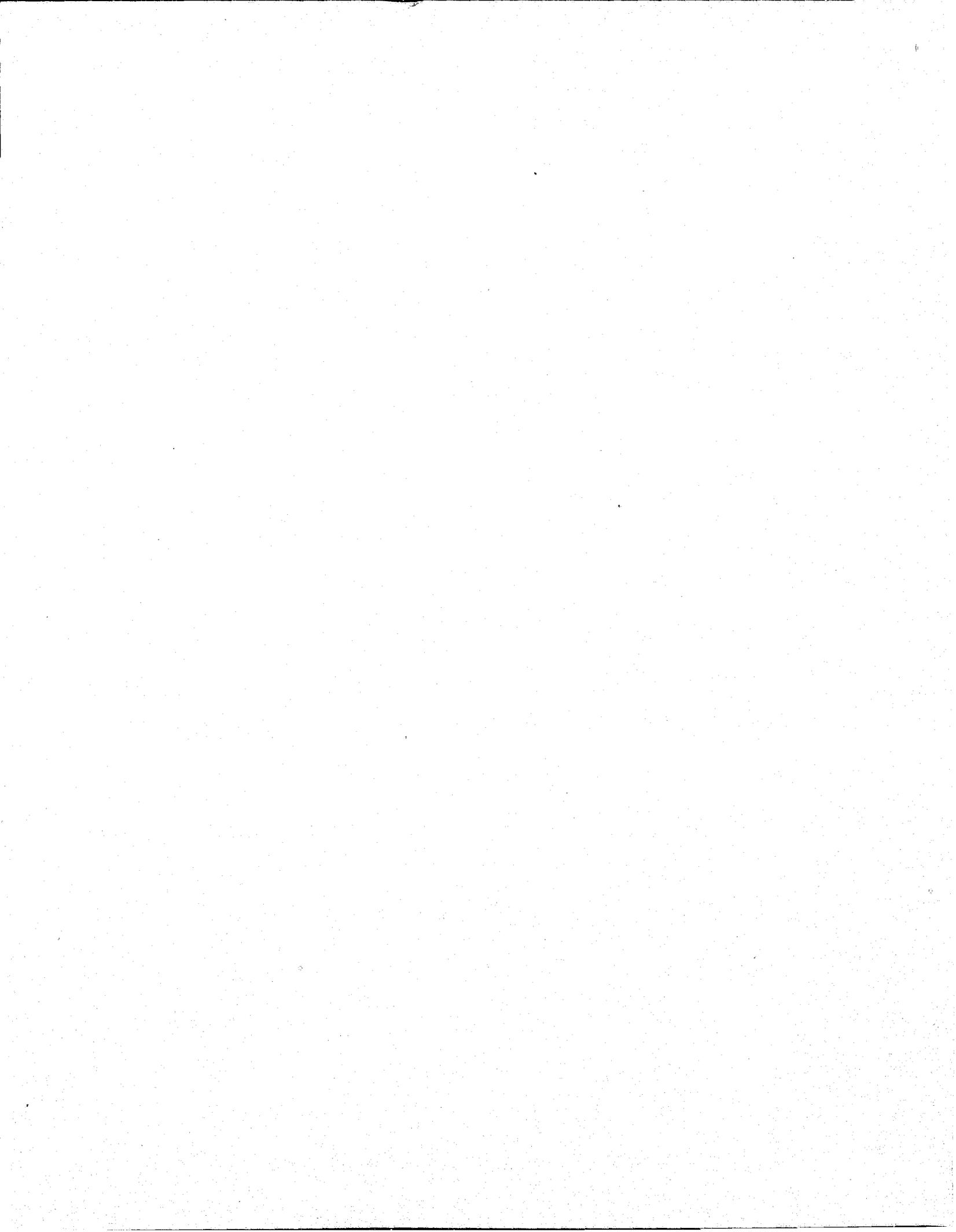
TABLE 4   AMOUNT OF USE OF EVALUATION INFORMATION, AS PERCEIVED BY EVALUATORS[1]

	No. areas N	No. resp. N	No. reprts. $\overline{X}$	Reviewed $\overline{X}$ %	Reviewed, not used $\overline{X}$ %	Considered $\overline{X}$ %	Major $\overline{X}$ %
State Evaluators							
Perception of use by LJPO planners/other professnsls.	1	5	28	46%	67%	28%	6%
Perception of use by Gov. committee	1	5	28	0	0	0	0
Regional Evaluators							
Perception of use by planners/ directors	7	8	56	81	31	41	29
Perception of use by advisory committee	6	7	45	62	31	39	31
Perception of use Project Directors	7	9	56	98	19	63	17

[1] The figures in the Table marked $\overline{X}$ % refer to the average proportion that evaluators indicated were used by each of the audiences. The percentages for not used, considered, and major role were calculated from a base figure that included only the reports the evaluators thought the groups had been aware of, thus the latter three categories total to 100 per cent. Perceptions of use by regional evaluators cannot be compared across categories except with considerable caution because different numbers of evaluators responded to the different questions. Seattle has no advisory committee and the evaluators in Seattle were excluded from that category reducing both the number of evaluators and the number of reports that were used as the base figure.

TABLE 5   TYPE OF USE OF EVALUATION AS RECALLED BY USERS[1]

	No. Resp.	Estimated Percentage of Evaluation Reports that were: used to argue for discontin- uance	used to argue for continuation	used to argue for significant change in project
Planners/Directors				
State	8	10%	12%	9%
Type 1 Region	11	14	65	47
Project Directors				
State	(2)	(0)	(75)	(50)
Type 1 Region	25	2	65	20
Type 2 Region	7	0	22	6
Planning Advisory Committees				
Gov. Committee	(1)	(0)	(100)	0
Type 1 Region	7	8	63	46
Type 2 Region	(4)	(21)	(36)	(12)
Evaluators				
State	5	7	43	4
Type 1 Region evaluators	9	18	41	19

[1] Number of respondents indicates the number that were aware of one or more evaluation reports. Percentages are based on these persons and on the number of reports they used in each of the different ways. Type 1 region refers to regions that have evaluation sections with at least one full-time evaluator.

planners/directors within the unit estimated that 59 percent of the reports played a major role.

Type of Use for Evaluations

Survey respondents were asked whether they had used evaluation reports to argue that a project should be discontinued, continued, changed in a significant way, and other similar questions. A summary of these results is presented in Table 5, along with evaluator's responses to the same questions.

A small percentage of the evaluation reports were used to argue that a project should be discontinued: estimates of 10 percent were given by state planners/other professionals; estimates of 14 percent from the regional planners; estimates of 2 percent from type 1 regional project directors, and estimates of "none used in this way" from project directors at the state level and in the regions without evaluation sections. Regional evaluators estimated that 18 percent of the reports were used to argue for discontinuation of a project and state evaluators estimated that 7 percent were used in this way. (These figures are based only on responses from persons who said they were aware of evaluation reports and the differences in percentages across audiences is based on a very small number of cases.)

Estimates of the proportion of reports used to argue that a program should be continued ranged from a low of 12 percent (estimated by the state LJPO staff) to a high of 65 percent -- an estimate given by the planners/directors in type 1 regions; project directors in type 1 regions, and the advisory committee in type 1 regions. Evaluators from those regions put the proportion at a lower level -- estimating that 41 percent of the reports were used to argue for a program continuation.

Evaluation reports also were used at times to argue for significant change in a project, such as change in its location, strategy of operation,

and so on. In the type 1 regions (regions with evaluation sections), the planners/directors estimated that they used 47 percent of the reports in this way; project directors estimated that 20 percent of the reports were used to argue for project change; the advisory committees, estimated that 46 percent were used for this type of argument, and the evaluators estimated 19 percent.

Use in LJPO Ratings

Another indicator concerning the extent of use given to evaluation reports could be obtained only at the state level. This was the written reasons given for judging a project to be "not acceptable" for funding and the written conditions stated for projects that were judged to be acceptable with conditions.

Of the 226 projects recommended by the regions for funding, there were 82 judged to be not acceptable, a total of 36 percent.

Of the reasons given for projects being not acceptable, the largest percentage (41%) were that the Submission 1 document from the region had not provided sufficient documentation of a need for the project. Twenty percent were judged not acceptable on the basis of policy decisions that had been made previously at the state level either by the Governor's Committee or the State Legislature. Two percent of those in the "not acceptable" category were there because the evaluation component was insufficient. There were other reasons also given for the project being not acceptable. Ten percent of the reasons for unacceptability involved indications that there was no evidence given that the project would have an impact on crime and/or that the project could reasonably be expected to achieve its goals. For only one of the cases in this category was it clearly said that data from the project indicated that it had no impact on its goals. In other words, there was one clear-cut instance in which evaluation information about a project was used as one of

several reasons for it being judged not acceptable. There were seven other instances in which "no evidence of effectiveness," or "unreasonable to expect project will be effective" or other similar types of evaluation-related statements were made. This provides additional documentation that evaluation information is not used much by the state planning office.

Comparative Importance of Evaluation Information

Respondents were asked to assess the importance of information from evaluation reports in comparison with other sources of information when final regional (or state) level funding decisions were made. Each person also was asked to indicate the "ideal" weights that they would attach to each type of information.

Regional respondents were asked to rate the importance of each of the sources of information in reaching final regional-level decisions about project funding or priority ranking. State-level respondents were asked to rate each in relation to final funding decisions by LJPO. The sources of information were to be scored on a ten point scale with 10 indicating the most important and one indicating the least important source of information. The same rating could be used more than one time. The sources of information to be compared were:

(1) Testimony or opinions of project directors or other project personnel based on experiences with the project;

(2) Testimony or opinions of program clients (excluding survey data contained in evaluation reports);

(3) Results of formal evaluations;

(4) Results of informal evaluations done by evaluators (monitoring, "trouble-shooting," etc.);

(5) Results of informal evaluations done by RPU planners (monitoring, "trouble-shooting," etc.);

(6) Testimony or opinions of community groups that are affected by the program;

(7) Testimony or opinions of the public, other than persons who are affected by the program;

(8) Newspaper or other media impressions of the project; editorial opinion about it.

As shown in Table 6, the regional planners/directors, evaluators and advisory committees in type 1 regions all gave scores on the current importance of evaluation that are above 5.0 which is the corrected average score. In these areas evaluation was ranked above all other sources of information by the planners/directors and the project directors. The advisory committee ranked it second. Individuals at the state level -- with the exception of state project directors -- gave evaluation a score that was below average. The state evaluators and the planners, program managers, and other professionals indicated its weight was three on the ten point scale. The Governor's Committee estimated the weight to be four. On the "ideal" weight, there obviously is no difference between the state and type 1 regions. Evaluators rate evaluation close to 9.0; planners in the type 1 regions gave it a weight of 8.5, and state-level professionals weighted evaluation at 8.0. The commit-tees provided "ideal" average weights of 7 and 7.5 on the ten point scale. Respondents from regions without evaluators weighed evaluation information higher in their "ideal" decision-making system than in the current one, but the scores are not as high as in the type 1 regions or at the state level.

Several comments are in order:

1. Regardless of which indicator one uses to estimate the extent to which planners and decision makers rely on evaluation information, it is apparent that the amount of use is greatest in the regional areas that have evaluation sections. The amount of use at the state level and within regions that have no evaluators attached to the planning agency is considerably less.

2. Even in regions that have the greatest amount of use (type 1 regions), the full potential of evaluation information has not been realized, as indicated

TABLE 6   IMPORTANCE OF EVALUATION

	Current Weight	Ideal Weight	Current Rank	Ideal Rank
State				
Planners, Prof	3.8	8.0	8	1
Adv.Committee	4.0	7.0	7	1
Evaluators	3.0	9.0	8	1
Proj.Directors	6.3	8.5	5	1
Type 1 Regions				
Planners	6.9	8.5	1	1
Adv.Committee	6.1	7.5	2	1
Evaluators	5.6	9	4	1
Proj.Directors	6.0	6.8	1	1
Type 2 Regions				
Planners	3.5	6.0	7	2
Adv.Committee	5.0	6.2	5	3
Proj.Directors	4.5	6.5	6	1

by the fact that not all evaluation reports play a major role in decisions; only a small percentage are used for the "hard decisions" such as arguing for discontinuation of a project, and by the fact that the current weight given to evaluation information is considerably below the "ideal" weight specified by respondents in the survey.

3. As will be shown in the subsequent section, however, the amount of use given to evaluation reports within the type 1 regions probably cannot increase until the technical quality of the evaluation reports improves.

Critiques of Evaluation Reports

Evaluations of LEAA-funded projects in 1976 were produced primarily by the regional evaluators. State-level evaluators did not, during 1976, conduct any in-house evaluations nor were there any state-level evaluation reports of on-going projects produced by outside contractors. Evaluations of several state-level projects were conducted by persons within the agency that administered the grant. There were so few of these available to us, however, that they were omitted from the critiques.

Although efforts were made to obtain copies of every evaluation report produced by the regional evaluators (and others), these reports were difficult to locate and the analysis in this section is based on 41 of the approximately 55 evaluation reports prepared by regional evaluators. All of the reports were critiqued by one of the authors of this paper and even though considerable efforts were made to insure that the rules used to judge the quality of the report were objective, there obviously is a subjective element to the conclusions that have been drawn.

Each report was critiqued and coded -- using a type of quantitative content analysis -- and the codes were keypunched for analysis. The more important types of data contained in the analysis were:

1. Type of project, targets/goals of the program, and strategy used
   by the project;

2. Number of months of project operation included in the evaluation;

3. The type of evaluation (outcome, process, monitoring against objectives, quantitative descriptions and problem analysis, non-quantitative descriptions);

4. Characteristics of the propositions and how each was tested; evaluators conclusion; reviewers conclusion;

5. Major problems that could bias the results; and

6. Other types of descriptive data included in the report but for which no propositions were tested.

The review procedure began by identifying an "analysis unit" which consists of one independent variable, one dependent variable, and one method of comparison or testing to determine the relationship between the independent and dependent variables. Each analysis unit in the report was to be specifically identified, reviewed, and then clustered (if needed) into groups of analysis units that tested the same proposition. A proposition is defined as having one criterion or dependent concept (even though there may be multiple measures of it); one independent or predictor concept (even though there may be multiple measures); and one or more methods of testing the relationship.

The original plan was to critique each analysis unit and code each of these onto a separate form. This proved to be a highly inefficient procedure and -- with some cost in terms of reliability -- the plan was changed so that each proposition was coded onto a single form and was treated as one "case" in the quantitative analysis.

In order to characterize each proposition and not lose too much of the richness in the data, codes were developed to show the strongest type of design that was used to test the proposition and the number of other quasi-experimental and non-experimental tests that were conducted. An additional test was "counted" if it assisted in ruling out at least one potential alternative explanation for the finding that was observed when the "best" test was

used. In addition to coding the number of different ways in which the same general proposition was tested, the coding form provides for a specification of how many different indicators were used to measure the same dependent concept (different types of crimes, for example) and the same independent concept (more than one very similar project, for example).

The most difficult parts of the critiquing were to determine exactly what the evaluator's conclusion was and what the reviewer's conclusion should be. Conclusions were coded into the categories shown in Table 7. At times, it was difficult to determine whether the evaluator was "positive" that the program had an impact or whether it should be coded as "probably." It also was difficult to code statements such as, "Recidivism rates were significantly lower for the project youths and the project (may have, seemingly, might have) contributed to the reduction." It was particularly difficult when the evaluator used different types of statements in different parts of the report, although this did not happen with much frequency.

The general rule for coding the reviewer's conclusion was that a definite conclusion could be drawn if there were no plausible alternative exaplanations that could account for the observed result -- including but no limited to whether the difference could be due to chance.

The following rules were used to determine the reviewer's conclusion:

Definite Conclusion (impact, no impact, harmful). A definite conclusion concerning positive impact, no effect, or harmful effect could be drawn if one or more of the following conditions was met in the evaluation report:

(a) Random assignment was used and there was no reason to doubt the integrity of the assignment; statistical and measurement procedures were appropriate.

(b) Quasi-experimental time series design with a comparison group or area (with at least ten pre-program time points); appropriate statistical

TABLE 7   CODING USED FOR CONCLUSIONS IN EVALUATION REPORTS

Outcome/Impact Evaulations in Which "The Project" is the Independent Variable:

01. Project had positive impact
02. Project "probably" had impact; other factors contributed
03. Impact occurred: project "may have" "might have" "could have" contributed
04. Mixed results, mainly positive
05. Project made no impact; no evidence of impact
06. Mixed Results; more negative than positive
07. Problem worsened: cannot ascertain project contribution
08. Problem worsened: project "probably" contributed
09. Problem worsened: project contributed

10. Data insufficient to draw any conclusion; design too weak to draw conclusion

Outcome/Impact Evaluations Testing Alternative Strategies or Comparing Projects

21. One strategy/project more effective than others
22. One strategy/project probably more effective
23. All strategies equally effective
24. Mixed results, mainly positive
25. No strategies are better or worse than others
26. Mixed results, mainly negative
27. One strategy probably less effective than others
28. One strategy is less effective than others

Monitoring Against Program Objectives

30. Project achieved objectives
31. Cannot determine; mixed results
32. Project did not achieve objectives

Monitoring/Process Evaluations Against Standard Set by Other Program, by Standards Established through Some Means Other than Project Goal Statements

40. Level of project activity is reasonable
41. Level of activity cannot be determined
42. Level of activity is below reasonable amount
43. Level of activity did not change over time (no learning, no increase in efficiency)

Cost Analysis/Process Evaluations

50. Program achieves cost savings/reduction
51. Program has no impact on cost
52. Program costs are more than the alternative method/project costs are greater than a reasonable standard

Other

60. Project not responsible for clients/project does not meet guidelines
70. Linkage between project activities and intermediate outcome does not exist making it virtually impossible to achieve goal or have impact.

and measurement procedures were used; results were not conflicting or uninterpretable; no serious alternative explanations exist for the results.

(c) Three or more alternative (and independent) quasi-experimental procedures were used to test the same proposition and each provided the same conclusion. Or, if more than three tests were used and if 80 percent or more of the tests resulted in the same conclusion, then the proper code was that a definite conclusion can be drawn.

Probable Conclusions Conclusions that were slightly less definite but strong enough to draw a judgment about the project could be made if the statistical and measurement procedures were appropriate and if one of the following conditions was met:

(a) Two different (and independent) quasi-experimental procedures were used to test the same proposition and each provided the same conclusion. Or, if more than two procedures, 65 to 79 percent yield the same conclusion.

(b) Quasi-experimental time series analysis without a comparison group; additional analysis reveals that other alternative explanations are not plausible.

(c) Comparison group design in which the analysis reveals that the experimental and comparison group did not differ much, if at all, prior to treatment; additional analysis reveals that the differences which do exist are implausible explanations for the results.

(d) Non-experimental designs would qualify for a "probably yes" only under exceptional circumstances. The analysis should make extensive tests to rule out alternative exaplanations for the findings. In addition, analysis should be undertaken to develop supporting evidence. For example, if the robbery rate for a street lighting project area has declined since the installation of the lights (according to a before and after analysis) then it also should be the case that night-time robberies should decline more than day-time

robberies. All crimes that presumably are deterred by lights should decline and the reduction should be primarily for the night-time occurrences. Similar before and after changes should not have occurred in areas without street lights.

If the analysis unit did not qualify for one of the above answers, then the reviewer's conclusion was that the data or design were too weak to arrive at a conclusion.

This differs to some extent from procedures used by other researchers in that it often is the case that conclusions of "no effect" are drawn when the data or design are too weak to show that the project was effective. In the review procedure we used, a judgment was not drawn concerning project ineffectiveness unless a conclusive statement -- based on strong design, statistical procedures and data -- could be made to show no effect.

Type of Information Produced

The data in Table 8 show the type of information that was produced by evaluation reports from regional evaluators. (Propositions tested in the reports written by persons other than LJPO regional evaluators were excluded from the Table). There were a total of sixty major propositions tested in the reports for an average of about two each. Data presented in reports that simply described the project or its environment were not coded as propositions. As noted before, a proposition involves at least one test or comparison designed to link the independent with the dependent variable.

Of the propositions tested, 42 percent were crime impact propositions; 15 percent were "system performance" propositions usually involving the impact of a project on arrest or conviction rates and seven of the propositions (12 percent) dealt with whether the project resulted in a cost savings to the system. An additional 12 percent were tests of propositions concerning client

TABLE 8 - CONCLUSIONS ABOUT PROJECT EFFECTIVENESS: EVALUATOR AND REVIEWER

	Crime Impact Propositions		System Perf. Propositions		Cost Analysis		Client Behavior, Attitude Prop.		Other Prop.		Totals		Percent of Total	
	Eval. N	Rev. N	Eval. N	Rev. N	Eval. N	Rev. N	Eval. N	Rev. N	Eval. N	Rev. N	Eval N	Rev N	Eval. %	Rev %
Project or Strategy Effective or Probably "	10	6	3	0	3	2			2	0	18	8	30	13
Impact Occurred: Project contribution Unknown									1	0	1	0	2	0
Project, Strategy Not Effective	9	8	2	0							11	8	18	13
Problem Worsened: Project Contribution Unknown														
Project or Strategy less effective than alternative	1	0			1	0					2	0	3	0
Data Insufficient to draw conclusion	3	11	0	7	1	3			1	5	5	26	8	43
Project Achieved Obj. Project Achievement Unk	2	0	2	0			2	2	3	3	9	5	15	8
Project did not Achieve Obj			1	1			2	2'	2	1	5	4	8	7
Activity Level Reasonable			1	0			2	2	2	2	5	4	8	7
Activity Level not known			0	1					0	1	0	2	0	3
Activity Level too low									2	1	2	1	3	2
Inappropriate clients							1	1			1	1	2	2
Linkage Missing									1	1	1	1	2	2
TOTAL NO. PROP.	25	25	9	9	5	5	7	7	14	14	60	60		

[1]Data in the table refer to conclusions drawn on each type of proposition by the evaluator and the reviewer. These propositions were tested in evaluation reports produced by the regional evaluators in 1975 and 1976. Not all reports were available, but our estimate is that approximately 75 percent of the 1976 reports were included in the critiques.

behavior, such as client attitudes, "successful exit" from a project, and so on. A variety of different kinds of propositions are grouped together into the "other" category. Several of these concerned victim reporting rates.

Comparisons of the evaluator's conclusions and the reviewer's conclusion should be made with an understanding that no claim will be put forth that the reviewer's critiques are entirely reliable and it may be the case that some of the evaluators would disagree with the conclusions drawn by the reviewer. Examination of the final columns in the table shows that the evaluators tested 18 propositions in which they concluded that the project was effective (30 percent) whereas the reviwer drew this conclusion for only eight of the propositions (13 percent). Evaluators tested 11 propositions in which the conclusion was that the project was not effective whereas the reviewer drew this conclusion in eight instances (13 percent).

Similar types of drop-offs can be observed in the conclusions concerning whether the project was harmful, achieved its objectives, did not achieve objectives, and so on. Most of the shifts by the reviewer were to the conclusion that there were highly plausible alternative explanations for the observed relationship and that one could not conclude that the project was effective or ineffective. The data imply that the reviewer simply was using more conservative standards (or more rigid criteria); this is not the case in relation to more than half of the evaluators with whom there was very little -- if any -- disagreement between the conclusions drawn by the reviewer and the evaluators. There were a total of 19 "shifts" made by the reviewer for the 60 propositions included in the analysis. Thus, there was disagreement concerning the most appropriate conclusion for 32 percent of the propositions and agreement on 68 percent.

The major type of disagreement concerned the extent of evidence needed before drawing a definite conclusion. Of the 19 shifts, 16 involved a change

from a definite conclusion to a conclusion that the data were insufficient to determine whether the project was effective, ineffective, or harmful.

Disagreement by the reviewer with conclusions drawn by the evaluators was not distributed evenly across different evaluators. Seven of the shifts (37 percent) concerned propositions tested by one evaluators and three other evaluators drew 13 conclusions that were shifted by the reviewer.

Discussion of Utility of the Information

The information produced by the regional evaluators would be of greatest value to planners and decision makers who decide on a project-by-project basis whether or not to re-fund the project. This does not mean that the information produced in evaluations is ideally suited to the task. If we consider the "ideal" information from evaluation to consist of the probability that a particular project contributed a specified number of "units" to the achievement of a particular goal, for a particular cost, then the information falls far short of the ideal. If the more reasonable standard is defined as providing scientific information on whether the project had an impact on a problem, then the reviewer's conclusions would suggest that 16 propositions involving approximately 12 different projects could be judged as effective or ineffective.

Using the "reasonable" standard defined above and the reviewer's assessment of the reliability of the conclusions, this would mean that there were only about 12 of the more than 200 LEAA projects for which evaluation information was sufficiently reliable to argue for continuation or discontinuation. If the evaluator's assessments were used as the standard, there were 24 of the 200 projects that could have been judged on the basis of outcome-oriented evaluation results. It should be recalled that the proportion of reports actually used to argue for continuation or discontinuation at the regional level (59 percent, according to the evaluators) exceeded this number and at

the state level was approximately 50 percent. Both of these exceed the pro-
portion of reports that were judged to be sufficiently reliable for this
purpose.

The point to be made is that there were not very many evaluation reports
which were reliable enough (in terms of a project's effectiveness vis a vis
outcome or impact measures) to be used as a conclusive guide to decision-
making.

Project directors may have used the reports in a variety of ways, but
most of the reports were not designed to tell the project director which of
the alternative services or strategies being used were most effective or which
cost the least.

Most of the evaluation reports produced at the regional level were not
designed to provide the type of information that state-level planners or
other officials would find most useful. If state level officials attempted
to use an evaluation report to justify a decision not to fund a particular
project within a region, they would be acting inconsistently with the general
principle of decentralized decision-making. Most of the reports do not compare
different strategies or different projects. It often is the case that one does
not know what type(s) of "treatment" the non-treatment control group received
and, therefore, it is difficult to determine what the project was tested
against. Furthermore, the evaluations generally made no attempt to determine
why a project was effective or ineffective -- the theoretical linkages between
activities and outcomes were seldom tested. This makes it difficult for
state-level officials to know whether the project is transferrable to another
area or another situation.

Although one might argue that state-level evaluators should "synthesize"
the evaluations in order to produce policy recommendations for state officials,
this would not be an easy task if one goes beyond the simple summarizing of

reports. Even though several similar projects were evaluated, the procedures used to collect data generally differed; the types of statistical analysis differed; the types of comparisons or controls that were used differed; and so on. In order to produce high quality policy papers, the state evaluators would need the raw data and would need to conduct secondary analysis of it.

OTHER PROBLEMS IN THE EVALUATION PROCESS

As discussed previously in this report, there are marked differences between the state planning process and the regional process concerning the use of evaluation reports and the degree of involvement in decision-making. These differences also exist in relation to a number of other attitudes and perceptions about the evaluation system.

Satisfaction and Support for Evaluation

As shown in Table 9, the degree of satisfaction with the quality of regional evaluation reports differs substantially between the state and regional areas. Regional evaluators are more satisfied with the quality of their evaluations than are the state evaluators and regional agency offi-cials, such as planners and administrators also are more satisfied with the evaluation reports than are their counterparts at the state level.

The same type of distinction exists for evaluators and agency personnel in terms of their satisfaction with the amount of use given to evaluation information (Table 10). State-level individuals are less satisfied than persons within the regions. A similar type of distinction exists with regard to attitudes toward regional evaluation efforts and state evaluation (Table 11). Each respondent was asked whether regional evaluations are worth what they cost, cost more than they are worth, or contribute more than their cost. The same question was asked about state-level evaluations. Regional evaluation

TABLE 9 . SATISFACTION WITH QUALITY OF REGIONAL EVALUATIONS

GROUP		Very Satisfied %	Satisfied %	Dissatisfied %	Very Dissat. %
EVALUATORS					
Regional	N=8	12.5	75	12.5	0
State	N=5	0	40	60	
	[Tau B = .50; (p) = .04]				
AGENCY					
Region, Type 1	N=11	55	36	9	0
State	N=9	0	67	11	22
	[Tau B = .55; (p) = .005]				

TABLE 10 . SATISFACTION WITH USE OF EVALUATIONS

		Very Satisfied %	Satisfied %	Dissatisfied %	Very Dissat. %	
EVALUATORS						
Regional	N = 8	25	50	25	0	100%
State	N = 5	0	0	80	20	100%
	[Tau B = -.68; (p) = .006]					
AGENCY						
Regional	N = 11	9	64	27	0	100%
State	N = 11	0	18	45	37	100%
	[Tau B = -.57; (p) = .002]					

TABLE 11 DIFFERENCE IN ATTITUDES TOWARD REGIONAL AND STATE EVALUATION

	Regional Evaluations worth the cost, or worth more than cost	State Evaluations worth the cost, worth more
EVALUATORS		
Regional N = 9	89%	0%
State	80%	80%
AGENCY		
Type 1 Regional N= 11	100%	50%
State N= 9	56%	70%
Type 2 Region N= 6	100%	( )
COMMITTEES		
Type 1 Region N= 14	86%	13%
State N= 6	100%	33%
Type 2 Region N= 17	71%	20%
PROJECT DIRECTORS		
Type 1 Region N= 40	82%	43%
State N= 9	66%	90%
Type 2 Region N= 17	71%	62%

efforts are given more support than state evaluations by persons within the regions whereas state-level officials (with the exception of members of the advisory committee) tend to have higher support for state evaluations.

Occupational Satisfaction

The surveys included a variety of questions which, when viewed in conjunction with one another, may be used as indicators of evaluators' "occupational satisfaction." Two of these indicators are the number of years an evaluator wishes to remain in his present position and a comparison by the evaluator of the _actual_ and _ideal_ amount of time spent on different tasks. Other indicators are based on the evaluators' _perceptions_: perceptions of the amount of influence they exert in their own offices, and perceptions of how their work and its importance is viewed by their professional colleagues in the criminal justice system.

Perhaps the most astonishing feature of these indices (at least to the outsider) are the large and consistent differences between the state and regional evaluators. While neither group perceives its situation to be completely harmonious, the state evaluators, by comparison, appear far more dissatisfied. Of the ten regional evaluators surveyed, for example, half expressed a desire to remain in their jobs for two years or more, and four for one to two years (the tenth responded "don't know"). Of the five state evaluators surveyed, one left during the needs assessment, another was transferred to a different section, and the other three indicated intentions to leave from "as soon as possible" to "no later than six months." At least in part, intentions to leave may be attributed to desires for promotion, advancement, and greater security. Evaluators' positions are usually grants awarded to regional planning units on a year-to-year basis and subject to the vagaries of funding decisions; thus there would be an understandable tendency among persons in

such situations to seek regular positions either in or out of state government. In addition, there is very little discernible opportunity for advancement within the "profession" of criminal justice evaluator.

Another index to occupational satisfaction may be obtained by taking a closer look at the data presented in Table 12. By adding up the differences between the actual and ideal proportions of time allocated to the various tasks, and averaging those differences across each group of evaluators, "deviation scores" of 22 for the eight RPU evaluators and 52 for the five LJPO evaluators are obtained. The upper limit of these scores is 200, which would occur if an evaluator spent 100 percent of his/her time on a task for which the ideal allocation was zero percent, and zero percent of his/her time on a task for which the ideal allocation was 100 percent. The scores may be interpreted as meaning that the RPU evaluators would, on the average, allocate only 11 percent (22/200) of their time differenly, while the state evaluators would allocate 26 percent (52-200) of their time differently. (If only the three state evaluators who were still in their positions as of November 1, 1976, are counted, the latter figure jumps to 35 percent.) In general, the regional evaluators would prefer to spend more time doing evaluations and less time collecting data for the state plan, while the state evaluators would prefer to spend more time doing evaluations and less time with administrative tasks.

Indications of the extent to which the evaluators perceive that others in the criminal justice system understand and support evaluation may be seen in Table 13. In almost every instance, the perceptions of understanding and support for evaluation among colleagues is greater for the regional evaluators than for the state evaluators. It is illuminating to note that when all non-evaluators surveyed were asked to respond to the same question, they, and in particular the state level personnel, indicated considerable understanding.

TABLE 12 ' EVALUATORS' ALLOCATION OF TIME

	N	Doing Evaluations Current %	Ideal %	Disseminating Information Current %	Ideal %	Coordinating Current %	Ideal %	Participation in Decision Making Current %	Ideal %	Collecting Data for plan Current %	Ideal %	Project Monitoring Curr. %	Ideal %	Other, including Administration Current %	Ideal %
Local $\overline{X}$	8	55	60	6.5	6.4	6	8	7	7	14	9	7	7	4	2
State $\overline{X}$	5	9	28	19	21	31	25	6	7	2	3	1	0	30	14
Whatcom N	1	50	75	5	1	5	10	1	1	25	10	9	2	5	1
King (8)	2	65	65	5	5	5	10	7.5	15	12.5	12.5	2.5	2.5	3	6
Seattle	1	33	60	7	10	15	15	5	5	25	0	5	10	10	0
Pierce	1	60	60	10	10	5	5	5	5	10	10	10	10	0	0
Thurston	1	80	80	5	5	0	2	5	3	8	2	2	8	0	0
Yakima	1	40	25	10	10	5	10	15	20	10	15	15	15	5	5
Spokane	1	50	50	5	5	10	10	10	10	10	10	10	10	5	5

TABLE 13
PERCEPTIONS OF
UNDERSTANDING AND SUPPORT FOR EVALUATIONS

Question: How well did _____ understand what evaluation should and could contribute to the over-all performance of the criminal justice system?

Audience	Level of Understanding	N
DIRECTORS	1.7	13
Regional	1.13	8
LJPO	2.6	5
PLANNERS	1.7	15
Regional	1.4	10
LJPO	2.2	5
ADVISORY COMMITTEE	2.8	11
Regional	2.3	6
Governor's Comm.	3.4	5
PROJECT DIRECTORS	2.4	15
Regional	2.4	10
State	2.5	5
EXTERNAL POLICYMAKERS	2.7	15
Regional	2.7	10
State	2.6	5
EVALUATORS	1.2	15
Regional	1.3	10
State	1.0	5
LJPO PROGRAM MANAGERS	2.2	5
LJPO STANDARDS & GOALS	2.0	3

Levels of Understanding:
1. Understood very well
2. Understood somewhat
3. Understood very little
4. Did not understand at all

NON EVALUATORS' SELF
ASSESSMENT OF UNDERSTANDING

	Level	
All Non Evaluators	1.57	120
Regional	1.77	96
State	1.52	24

and support of evaluation (the average for all non-evaluators was 1.57; for non-evaluators at the state level the average was 1.52).

An almost identical pattern, but with more striking differences, was observed when each evaluator was asked his perception of how important his/her colleagues thought evaluation was to the over-all effectiveness of the state planning agency or the regional planning unit (see Table 14). Again, regional evaluators believed to a greater extent than the state evaluators that other in the criminal justice system think their work is important. Regional evaluators were, on the average, quite accurate in their perceptions of how their colleagues view the importance of their work. State evaluators, on the other hand, greatly under-estimated the importance of evaluation to their co-workers.

Perceptions among evaluators of the impact of their work on the understanding and assessment of importance of evaluation by other persons and groups in criminal justice may be determined from Table 15. The perceptions of both groups of evaluators for external policy makers and planners were about the same; otherwise, the regional evaluators, on the average, perceived a slight increase in the understanding and assessment of importance of evaluation among their colleagues, while the state evaluators perceived a slight decrease. The non-evaluators themselves, at both the state and local level, indicated that their understanding and assessment of the importance of evaluation had slightly increased during the past year.

The pattern of differences between state and regional evaluators repeats itself with respect to their perceptions concerning the attitudes of others toward evaluation in general (Table 16). In addition, state evaluators indicate that their effectiveness was hampered by a lack of understanding and support for evaluation to a greater extent than regional evaluators (Table 17). The evaluators appear quite accurate in their perceptions of the amount of influence

TABLE 14
PERCEPTIONS OF NON-EVALUATORS
IMPORTANCE OF EVALUATIONS

Question: How important did_____think evaluations were?

Audience	Level of Importance	N
DIRECTORS	2.2	13
Regional	1.75	8
LJPO	3.0	5
PLANNERS	2.1	15
Regional	1.7	10
LJPO	2.8	5
ADVISORY COMMITTEE	2.9	11
Regional	2.2	6
Governor's Comm.	3.8	5
PROJECT DIRECTORS	2.3	15
Regional	2.2	10
State	2.6	5
EXTERNAL POLICYMAKERS	2.6	13
Regional	2.6	8
State	3.3	5
EVALUATORS	2.4	14
Regional	1.9	9
LJPO	3.2	5
LJPO PROGRAM MANAGERS	2.6	5
LJPO STANDARDS & GOALS	2.4	5

Levels of Importance:
1. Very important
2. Somewhat important
3. Somewhat unimportant
4. Very unimportant

NON EVALUATORS' SELF
ASSESSMENT OF IMPORTANCE
OF EVALUATIONS

	Level	N
All Non Evaluators	2.25	120
Regional	2.31	96
State	2.00	24

TABLE 15
PERCEPTIONS OF CHANGE IN
UNDERSTANDING AND IMPORTANCE OF EVALUATIONS

Question: Have you observed any marked change by_____in terms of how well they understand what evaluations should contribute and in terms of their assessment of how important evaluation is to the effectiveness of the law and justice system?

Audience	Amount of Understanding	N	Change Importance	N
DIRECTORS	2.6	13	3.0	13
Regional	2.3	8	2.6	8
LJPO	3.2	5	3.6	5
PLANNERS	2.2	14	2.6	14
Regional	2.2	9	2.8	9
LJPO	2.2	5	3.0	5
ADVISORY COMMITTEE	2.6	9	2.5	10
Regional	2.2	5	2.2	5
Governor's Comm.	3.0	4	3.8	5
PROJECT DIRECTORS	2.0	12	2.33	12
Regional	1.9	9	2.1	9
State	2.33	3	3.0	3
EXTERNAL POLICYMAKERS	2.6	14	2.7	14
Regional	2.7	9	2.8	9
State	2.4	5	2.6	5
EVALUATORS	1.9	14	2.8	14
Regional	1.8	9	2.2	9
LJPO	2.2	5	3.8	5
LJPO PROGRAM MANAGERS	2.6	5	3.4	5
LJPO STANDARDS & GOALS	2.4	5	2.8	5

Amount of change:
1. Increased greatly
2. Increased some
3. Stayed the same
4. Decreased some
5. Decreased greatly

NON EVALUATORS' SELF
ASSESSMENT OF CHANGE

	Underst.	Impt.	N
All Non Evaluators	2.3	2.5	120
Regional	2.4	2.26	96
LJPO	2.13	2.19	24

TABLE 16
PERCEPTIONS OF
ATTITUDES TOWARD EVALUATION

Question: How would you describe the attitude of_____toward evaluation?

Audience	Attitude	N
DIRECTORS	1.8	14
Regional	1.22	9
LJPO	2.8	5
PLANNERS	1.6	15
Regional	1.3	10
LJPO	2.2	5
ADVISORY COMMITTEE	2.3	12
Regional	1.9	7
Governor's Comm.	3.0	5
PROJECT DIRECTORS	2.0	13
Regional	2.0	10
State	2.0	3
EXTERNAL POLICYMAKERS	2.0	14
Regional	2.0	9
State	2.0	5
EVALUATORS	1.1	15
Regional	1.1	10
LJPO	1.0	5
LJPO PROGRAM MANAGERS	2.6	5
LJPO STANDARDS & GOALS	1.8	5

Attitudes:
1. Very supportive
2. Somewhat supportive
3. Somewhat unsupportive
4. Very unsupportive

NON EVALUATORS' SELF
ASSESSMENT OF ATTITUDES

	Attitude	N
All Non Evaluators	1.44	120
Regional	1.50	96
LJPO	1.29	24

TABLE 17

PERCEPTIONS OF PROBLEMS

Question: To what extent did a lack of understanding and support for evaluation hamper your effectiveness as an evaluator?

	Lack of Underst. Problem Level	Lack of Support Problem Level	N
ALL EVALUATORS	1.6	1.8	15
REGIONAL	1.8	1.9	10
LJPO	1.4	1.2	5

PROBLEM LEVEL

1. Major problem
2. Minor problem
3. No problem at all

they exercise on decision-making in comparison with other groups -- which, in general, is equal or less (Table 18). Although there are some slight over-estimations and under-estimations, the regional and state evaluators' perceptions of their influence consistently matches the non-evaluators' perceptions at both levels. The differences between the groups are that regional evaluators perceive greater influence on their directors and advisory commissions while state evaluators perceive greater influence on other groups.

Skills, Experience, and Turnover

Evaluation is essentially a research activity and, consequently, demands training and experience in the conduct of research, familiarization with different types of designs and methodologies, possession of the appropriate skills in data collection and analysis, and access to the necessary research facilities. In an effort to isolate specific problems and needs of the evaluators with respect to their research activities (as opposed to the quality of the evaluation reports, use of evaluation in decision-making, and so forth) all evaluators in the state planning office and regional planning units were asked to make a self-assessment of their research skills and respond to questions concerning their educational backgrounds and training, salaries, experience, and access to data analysis equipment. The response rate of the fifteen evaluators surveyed was 100 percent; data pertaining to degrees earned, experience, salaries, number of semester credit hours in statistics and research, and self-assessed research skills are summarized in Table 19.

Lack of experience is one of the most revealing characteristics of the evaluators, particularly when coupled with an unusually high turnover rate. Only three of the evaluators reported having as much as three years' professional experience; another eight had from 1.5 to 2.5 years, and the remaining four had only one year or less. Moreover, two-thirds of the evaluators (ten)

TABLE 18
PERCEPTIONS OF THE INFLUENCE
OF EVALUATIONS SECTION ON DECISION MAKING

Question: How much influence did the evaluation section have on decisions made by_____?

	Eval.Percept. Degree of Influence	N	Non-Eval.Percept. Degree of Influence	N
DIRECTORS	2.16	14	2.08	120
Regional	1.8	9	2.0	96
LJPO	2.8	5	2.33	24
INTERNAL AGENCY PLANNING SESSION	1.92	13	2.23	120
Regional	2.0	8	2.25	96
LJPO	1.8	5	2.17	24
ADVISORY COMM.	2.65	12	2.19	120
Regional	2.4	7	2.06	96
Governor'sComm.	3.0	5	2.71	24
PROJECT DIRECTOR	2.17	13	2.40	140
Regional	2.4	8	2.44	96
State	1.8	5	2.21	24
EXTERNAL POLICY MAKERS	2.38	13	2.46	120
Regional	2.5	8	2.50	96
State	2.2	5	2.33	24

Degree of Influence:
1. More influence than most other sections
2. Equal influence
3. Less influence than most other sections

TABLE 19

DEGREE	EXPERIENCE		SALARY		STAT., RES.		SKILLS		N
	Avg.	Range	Avg.	Range	Avg.	Range	Avg.	Range	
Ph.D.[1]	2.67	2 - 3	18,767	15,300-21,000	17.67	4 - 25	1.35	1 - 1.91	3
MASTERS[2]	1.67	1 - 3	14,058	9,348-17,000	19.17	12 - 25	2.20	1.8-3.09	6
B.A.[3]	1.62	.7- 2.5	12,472*	10,200-15,000	7.67	0 - 21	2.59	1.66-3.37	6

1. Includes two state evaluators

2. Includes one state evaluators

3. Includes two state evaluators

* Average based on five respondents
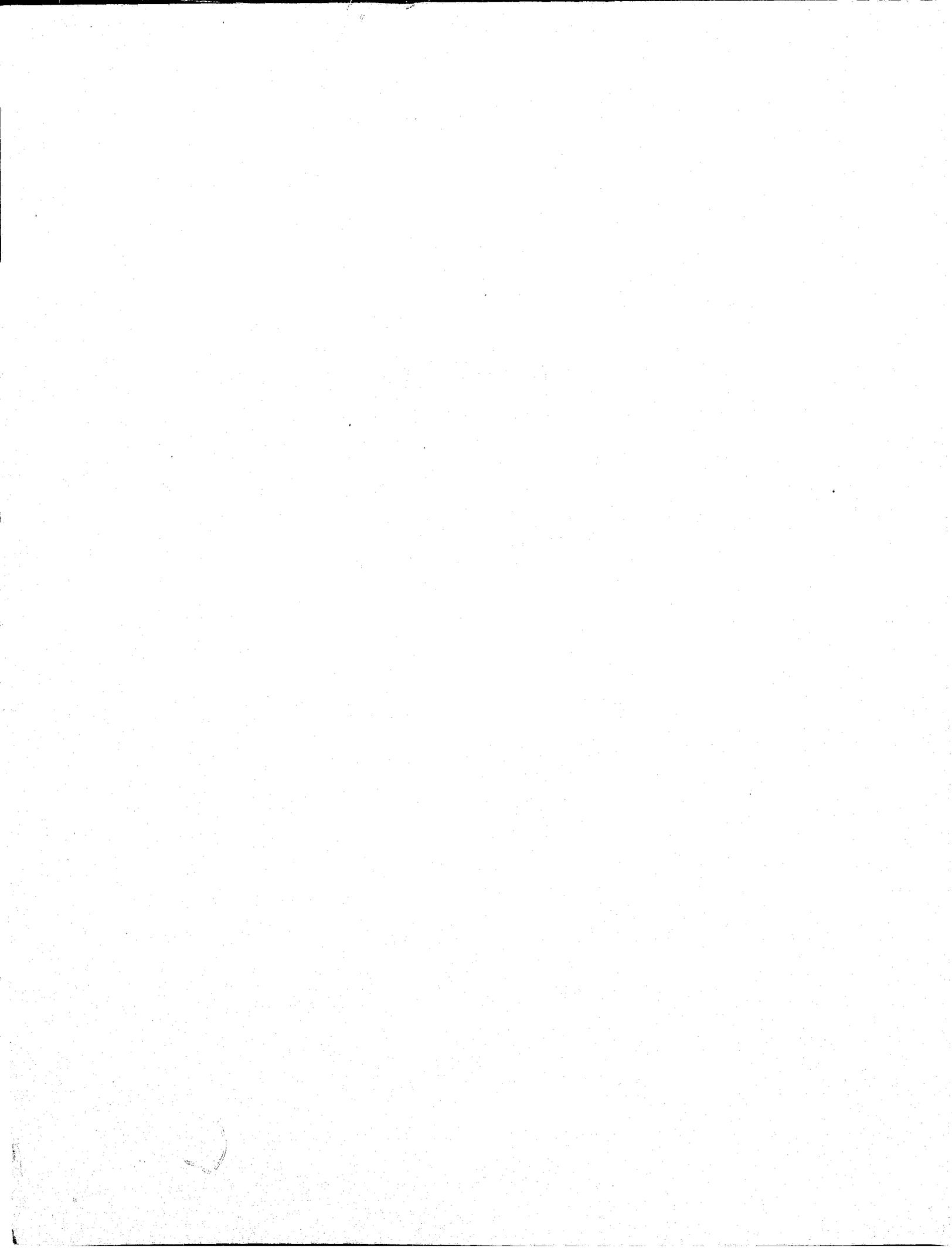
reported having gained all of their experience in their present positions, implying no prior experience at the time they were employed. Unless the practice of hiring inexperienced evaluators is modified, the problem is likely to become a chronic one. Only three of the evaluators hired as re-. cently as two years ago are still in their positions. Assuming a base of fifteen evaluators (as called for in the 1974 State Plan) this reflects a turnover rate of 80 percent. While some of the turnover might be attributed to normal processes of adjustment in a new organization, there is no evidence to indicate that the turnover rate will, as one evaluator suggested, "taper off" in the future. In fact, only three of the evaluators expressed a desire to remain in their positions for more than another two years. If the others act upon their intentions to leave within the next two years, then two years from now the turnover rate will still be 80 percent.

The research skills of the evaluators are difficult to assess quantitatively, primarily because there is no standard -- either absolute or contained in the job descriptions -- against which they could be compared. In the critiques of the evaluation reports, the comparison was drawn between what could have been done with available data and what actually was done. The critiques suggest there is considerable room for improvement, but it is certainly conceivable that the reports might have been of higher quality had the evaluators had more time and/or access to high-speed data processing equipment. Other sources of information concerning skill levels include self-assessment of research skills, educational attainment, and number of semester hours of statistics and research methodology taken in college.

The questionnaires mailed to the evaluators contained a 35-item checklist of skills and techniques applicable to the conduct of evaluation research. Included were items pertaining to the ability to phrase research questions, perform statistical operations of varying degrees of difficulty, and cast the

propositions and findings into theoretical frameworks. The respondents were asked to rank themselves on each item in accordance with a four-point scale in which 1 was "expert," 2 was "advanced," 3 was "learner," and 4 was "beginner." The results, broken down by highest degree earned, are presented in the table. One person, a Ph.D. with three years' experience as an evaluator and 25 semester-hours of statistics and research courses, rated himself "expert" on every item. Two other persons, each with BAs and no course work in statistics or methodology, rated themselves about halfway between "beginner" and "learner." The scores are only slightly skewed toward "expert" and otherwise are fairly normally distributed about a mean of 2.18 (slightly below the "advanced" level). As would be expected, the scores are highly consistent with educational attainment, experience, salary level, and course work in statistics and research. It should be noted, at least in passing, that the evaluators' self-assessed research skills are somewhat greater than would be suggested by the critiques of their written reports. However, it bears repeating that the quality of the evaluation reports may be due to factors other than the skills of the evaluator.

The evaluators are very well-educated as a group, with 60 percent holding doctorates (3) or masters(6). However, two-thirds of those with advanced degrees are located in either the State Planning Agency or Seattle-King County. While one should not over-emphasize "credentialism" or under-rate the value of a bachelor's degree, it is important to make the distinction because research skills are rarely taught at the undergraduate level. This is borne out by the fact that the average number of semester-hours in statistics and research taken by those with bachelor's degrees is less than half those taken by evaluators with advanced degrees.

Computer Facilities

Another problem frequently cited by regional evaluators is the difficulty

of gaining access to computers equipped with the necessary software for doing statistical analyses. In one of the surveys conducted as part of the needs assessment, only four of the nine responding regions in which evaluations were done reported access to computers and a statistical package such as SPSS. Three of the four regions indicated at least partial reliance on hand calculators for data analysis, and the remainder reported total reliance on hand calculators. While the state of the art in calculators has progressed remarkably in recent years, the basic problems of doing data analysis with such equipment remain: it is tedious, time-consuming, and prone to error in direct proportion to the number of cases being analyzed.

The need for computers in the conduct of evaluation research may actually be greater than that expressed by the evaluators. Our critiques of 41 evaluation reports completed over the past two years -- most of which were done with hand calculators -- suggests that reliance upon such equipment for data analysis greatly restricts the quality of the evaluation. Evaluations done with calculators are characterized, for example, by univariate as opposed to multivariate analyses, aggregation of cases across time and disparate groups, and reliance upon statistics which are easily calculated rather than those most appropriate for the analysis. Consequently, alternative hypotheses based on the influence of intervening variables are not examined, information is wasted and data are consistently under-analyzed. Moreover, efficiency is reduced by the inability to maintain a data file which can be updated by the addition of new cases or variables. In short, the use of computers not only increases the quantity of data which may be analyzed, but also the quality of the analysis itself.

SUMMARY AND CONCLUSIONS

This study was designed to describe the evaluation process in the State of Washington and develop intervention procedures that would improve it. The

intent was not to test propositions or hypotheses about the conditions that are necessary or sufficient for evaluations to have an important role in planning and decision-making. Nevertheless, a number of conclusions concerning the conduct of evaluation would seem to be justified on the basis of the study.

1. Evaluations need to be available and used at the point(s) in the decision-making process where the decisions are made which eventually result in some projects being funded and others not being funded.

The analysis indicated that decisions which eventually result in the funding of regional projects are made, for the most part, at the regional level or even at a lower level of government within the region. The evaluation system also is generally decentralized. There is considerable evidence that the evaluation, planning, decision-making process is most effective in those regions that have one or more full-time evaluators. It also is the case, however, that the decentralized nature of the system has resulted in problems for the state-level evaluators in that their role is not clearly defined.

2. Information from evaluation reports is most likely to be used by the audiences for whom the evaluation was designed and produced; the evaluation report will not be of as much value for other persons in the system.

A variety of different indicators were developed to estimate the extent to which evaluation reports produced by the regional evaluators were used within the regional planning agency. Regardless of which indicator is used, there is strong evidence that regions which have evaluators use evaluations in developing recommendations for project rankings and funding. Regions without evaluators do not have many evaluation reports upon which they could rely. Most of the evaluations available to the state planning office were those produced by the regional evaluators and these were not used to any great extent in the state review process. Possible reasons for under-utilization

at the state level can be summarized as follows:

(a) The evaluations were not designed for use by state-level officials. The information (generally) is tied specifically to one project (rather than to a comparison of projects) and seldom contains enough detail about alternative procedures used within the project to provide guidelines for state officials concerning standards of project operation.

(b) Individuals within the state planning office are less trusting of the reliability and quality of regional evaluations than are persons within the region where the evaluation was produced.

(c) Although the state planning office reviews each plan and each project recommended for funding, the general philosophy of LJPO is that regional preferences should prevail. The state planning office does not view its role as one which imposes strategies on the regional areas nor as one in which it determines the specific projects to be funded. Thus, LJPO has no clearly identified role for use of evaluation reports produced by regional evaluators. If the state office wished to use these to argue for raising or lowering the priority given to a project, then this would be contrary to the general philosophy of decentralized decision-making. If they wished to use the regional evaluation reports to argue that a particular type of project or strategy (such as diversion of juveniles) has not been shown effective by the evaluations, then this too would constitute a decision of a nature that the state planning agency does not normally wish to make.

3. Information from evaluation reports is most likely to be used by persons who are supportive of evaluation and who wish to rely on it heavily, even if other sources of information also are used in decision-making.

The data indicate that officials in the regions with evaluators and within the state planning office would like to rely heavily on evaluation reports in making decisions about project funding. State-level planners,

program managers, standards and goals personnel, evaluators, administrators, and members of the Governor's Committee indicated that (on the average) they would rank information from formal evaluation reports first in terms of the type of information they would like to use in decision-making. This information outranked (in terms of the "ideal" decision-making system) seven other information sources including informal evaluations, opinions of community groups that are affected by the projects, clients of the projects, testimony of project directors, informal assessments done by planners, opinions of the unaffected public, and media commentary about projects.

The state-level officials ranked information from evaluation either last or next to last in terms of the current decision-making procedures. Most persons from regions which have evaluators ranked information from the evaluations either first or second in terms of the current decision-making procedure. The exception were the evaluators who, on the average, indicated that evaluation information ranked fourth in terms of current decision-making. All groups in regions with evaluators ranked evaluation first in terms of their "ideal" decision-making system. Regions without evaluators tended to rank information from formal evaluations low in relation to current decision-making and either first or second on the "ideal" decision-making scale.

The data suggest that there is a generally high level of support for evaluation and a strong desire to use it.

4. Information from evaluation reports cannot be relied upon extensively for "hard" decisions unless the information is scientifically reliable and can withstand strong challenge from persons who disagree with the conclusions drawn in the evaluation report.

Critiques of evaluations produced at the regional level suggest that although many of them are of high technical quality, there are not many reports that are crime-oriented and reliable enough methodologically to withstand

challenges. The data from the critiques indicate that there cannot be any dramatic increase in the use of evaluations to make "hard" decisions until the number of scientifically reliable evaluations increases. From a more speculative standpoint, an argument could be made that the degree of use is at its maximum unless the reliability of the reports increases and there is even some indication that evaluations in some regions may be used too much -- given the quality of the evaluations.

5. The production of high quality evaluations is facilitated by a supportive agency, evaluators with the technical skills to conduct evaluative research, and evaluators who have the facilities to conduct sophisticated data analysis.

The information contained in the report suggests that "support" is not a major problem for regional evaluators, but there is a problem for many regional evaluators either in terms of their technical skills or their facilities.

At the state level, there is generally high support for evaluation information, but there also is considerable dissatisfaction with the way evaluation has been handled at the state planning office. From the evaluator's point of view, one would conclude -- based on data presented earlier in this report -- that there is a lack of organizational support for their efforts.

From the point of view of others in the state planning office, there is not a lack of support for evaluation, but a lack of satisfaction with the performance at the state level. Many of the evaluation reports from the regions do not contain conclusive information about the effectiveness of projects; there have been very few (if any) state-level evaluation reports pertaining to topics about which LJPO is in a position to make decision. In short, the role of evaluation at the state level is not clear.

The turnover rate among evaluators was an alarmingly high 80 percent over the past two years, and data from our surveys indicate that this rate will continue through 1979. There is comparatively little opportunity for upward mobility within the "evaluation profession" in the State of Washington, and so at least part of this turnover may be attributed to desires for advancement and promotion. However, there is evidence to suggest that regional evaluators are frustrated by a lack of time and resources to do their jobs properly. Evaluators at both levels -- but state evaluators to a greater extent -- indicate further frustration due to a perceived lack of influence and perceptions that others in the criminal justice system do not understand evaluation or consider it important.

A P P E N D I X   A:. <u>SURVEY METHODOLOGY</u>

Note: The analysis and information in this section were prepared primarily

by

Colleen Cleary

APPENDIX A: METHODOLOGY AND DATA

The Sample

The survey consisted of local and state evaluators, planners, administrators, members of advisory committees, and project directors. The population included all State Law and Justice Planning Office (LJPO) evaluators, planners, program managers, standards and goals personnel and all local Regional Planning Unit (RPU) directors, planners, evaluators employed as of October 1976. All members of the 1976 Governor's Committee and the Juvenile Justice Advisory Committee were surveyed. A 20 per cent sample of each RPU advisory committee was drawn using a random numbers table. A preliminary screening for the sample of project directors was first conducted by identifying the total number of projects current as of October 1976. Excluded from this total were projects listed as equipment, minor personnel, or evaluation. A 40 per cent sample of project directors was then drawn using the table of random numbers. Excluded from this sample were any additional projects found (according to RPU information) to be no longer current.

The Instruments

The survey instruments consisted of one basic questionnaire with three versions: one version for evaluators, one for non-evaluators (planners, directors, program managers, standards & goals personnel and advisory committee members) and one survey for project directors. Within each version different forms were developed so questions referred specifically to the audience surveyed. There were six forms of the basic questionnaire.

Evaluators

In the questionnaire evaluators were asked to recall specific information concerning the amount of use given to evaluation reports by each of several users--planners, directors, project directors, and advisory committees. Direct (open-ended) and additional (closed) questions were asked about the specific ways evaluation reports were used. Evaluators were also asked whether the reports contained the information that each of the users wanted. A series of attitudinal/perceptual questions about evaluation provided information on the perceived importance and understanding of evaluation. Another series of questions were asked concerning the actual and ideal weight of several alternative sources of information for decision making--formal evaluation reports, informal reports, testimony. Evaluators were also asked to estimate how they spend their time, some background questions and a long series of questions concerning their own assessment of their evaluative skills. Two forms of the evaluator questionnaire were developed: one form for state (LJPO) evaluators and one form for local (RPU) evaluators.

Non-Evaluators

This questionnaire was identical to the evaluators except for the following differences:

1. As <u>users</u> of evaluation reports, non-evaluators were asked how many reports they personally used and in what way they were used, replacing the evaluators' questions on the estimated amount and type of use made of evaluation reports.

2. Additional questions were phrased to tap personal experience with evaluation reports rather than perceptions of how others use evaluation.

3. Only the evaluators were asked to assess their skills for doing evaluations.

Three forms of the non-evaluator questionnaire were developed: one form for state (LJPO) planners and program managers, one form for local (RPU) directors and planners, and one form for all state and local advisory committees.

Project Directors

This questionnaire was identical to the non-evaluators except for the addi-

tional questions concerning specific information about evaluation reports on their project, i.e., how many, who did them and how were they used. The same form of this questionnaire was sent to both state (LJPO) and local (RPU) project directors.

One additional questionnaire--the Case Study Protocol--was sent to one person in each RPU to obtain factual data about the evaluation process at the local level during 1976.

ADMINISTERING THE INSTRUMENTS

An advance letter designed to notify each person about the survey was sent one week prior to the questionnaire. The survey instruments were administered in two different fashions, mailed questionnaire and telephone interviews.

Mailed Questionnaire

All RPU evaluators, planners, directors and selected project directors and advisory committee members were mailed appropriate versions of the questionnaire. All members of the state advisory committees and the selected state project directors also received mailed questionnaires. Those who had not returned the questionnaire in 10 days were telephoned and responses were obtained over the telephone.

Telephone Interviews

All state (LJPO) evaluators, program managers, planners and standards & goals personnel were mailed the questionnaire for their review one week prior to the telephone interview. During the interview, each person referred to their copy.

Confidentiality of the data was maintained by assigning unique code numbers to each questionnaire. As the questionnaires were returned or as the interview was conducted, this number was destroyed and a new number assigned.

CODING

All questionnaire responses were coded by a team-editing procedure, i.e., each survey was read by a "first reader" coding consistently the open-ended

questionnaire with the developed codes and noting logical inconsistencies or problems in the responses. A second person then read and checked the codes of the first coder. All coders were trained in the editing procedures.

SURVEY RESPONSE

Seventy-five per cent of the questionnaires were returned by mail or through telephone interviews. Of those questionnaires returned 71 per cent were completed. Table A-1 summarizes the total number of questionnaires administered, returned and completed. Also given are the number of respondent returns for each survey form.

Of the 29 per cent of questionnaires which were not completed, exactly half of the respondents returned the questionnaire or a letter indicating reasons for non-completion (see Table A-2). During the telephone follow-up of those questionnaires which had not been returned in 10 days, respondents indicated reasons for non-return. In most cases, the respondents either lacked information about evaluation use because they were too new or inactive, or they considered themselves no longer a member of the audience surveyed due to retirement or resignation. Very few respondents (two per cent of the returned questionnaires) indicated negative reasons, such as, evaluations or survey questionnaires were a waste of time. A number of questionnaires were received too late to be included in the data for the preliminary report. Other reasons of non-completion were inappropriateness of the questionnaire to a particular project, insufficient time to complete the questionnaire or respondent's inability to locate the questionnaire. In some cases, no reason was given.

Non-Response of Survey

Of those individuals who never returned the questionnaires, 53 per cent had been contacted by telephone. Eighty per cent indicated they would mail the questionnaire within a week and 20 per cent indicated they had already returned the questionnaire. Forty-seven per cent of the non-respondents were never contacted.

TABLE A-1. RESPONSE RATE[1]

	Survey	N Sent	Total Returned	% of N. Returned	N Completed	% of returns Completed
Evaluators	1. State	5	5	100	5	100%
	2. Local	13	13	100	10	77%
Non-Evaluators	3. Regional	30	24	80	20	83%
	4. State	16	13	81	11	85%
	5. Adv.Comm.	149	96	64	51	53%
Projects	6. Project Directors	112	95	85	75	79%
Case Study Protocol		18	11	65	11	100%
TOTAL		343	257	75%	183	71%

[1]Figures in Column Three represent the proportion of persons in the original sample who returned the questionnaire or who were interviewed by telephone. Figures in Column Five indicate the proportion of questionnaires returned that were complete enough to use in the analysis. Some questionnaires were sent back with very few responses on them.

TABLE A-2 REASONS FOR NON-COMPLETION OF QUESTIONNAIRE

Reason for Incomplete Questionnaires	Questionnaire/letter returned with reason N	Questionnaire not returned. Reason given over telephone N
1. Lack of information, too new, inactive	12	11
2. No longer member of audience surveyed--resigned, retired	4	6
3. Evaluations/Questionnaires a waste of time	1	7
4. Questionnaire received too late for preliminary draft	7	-
5. Questionnaire inappropriate	6	1
6. Insufficient time to complete questionnaire	1	3
7. No reason given	2	3
8. Question did not reach person; could not recall	4	6
Totals	37	37

The majority of these individuals were the selected members of local RPU advisory committees; however, some project directors were not contacted as shown in Table A-3. Unavailable telephone information or unsuccessful attempts to contact the individual during business hours prevented obtaining the responses by telephone.

The analysis and interpretation of data in the main body of the report probably were not biased much, if at all, by the non-response to the survey. A test of the type of bias introduced by non-respondents was made by correlating the number of days before respondents returned the questionnaire with selected questions on use of evaluation reports and attitudes toward evaluation. The assumption underlying the test is that attitudes and/or behaviors of non-respondents are a linear projection of "slow" respondents. For example, one could propose that persons who responded to the questionnaire were more apt to use evaluation reports, more positive toward evaluation, had a greater understanding of it, and so on than did persons who never responded at all. If so, it is reasonable to believe that persons who responded only after several telephone contacts or who gave their responses over the telephone several weeks after they received the questionnaire would resemble the non-respondents to a greater extent than they resemble the "quick" respondents. Thus, if there is a relationship between the number of days before a person responded and their use of evaluations or their attitudes about evaluation, we would conclude that failure to obtain responses from about 25 per cent of the original sample biased the results.

As shown in Table A-4, there were almost no significant correlations between slowness of response and any of the behavioral or attitudinal characteristics of respondents.

Slow respondents did not indicate greater use of evaluations than quick respondents, and did not differ in terms of the "ideal" weight given to evaluation reports. Slow respondents were somewhat less apt to indicate a high level

TABLE A-3   CONTACTS MADE WITH NON-RESPONDENTS

| | Survey Form | Contacted by Telephone | | Total Contacted | Never Contacted |
		Had Mailed	Will Mail		
Evaluators	1. State	N --	N --	N --	N --
	2. Local	--	--	--	--
Non-Evaluators	3. Regional	1	5	6	--
	4. State	1	2	3	--
	5. Adv. Comm.	3	15	18	35
Projects	6. Project Directors	3	9	12	5
Case Study Protocol		1	6	7	--
TOTALS		9	37	46	40
Proportions (N=86)		10%	43%	53%	47%

TABLE A-4

SUMMARY TABLE: Correlations between number of days to return and selected questions: All respondents

Survey Question	Correlation (R)	Slope (B)	Significance	N
Current weight of use of formal evaluations	-.008	-.003	.47	112
Ideal weight of use of formal evaluations	-.075	-.019	.21	117
Number of reports reviewed	-.027	.039	.38	121
Per cent of reports which were				
reviewed, but not used	-.068	-.001	.30	61
considered	.219	.009	.04	61
played a major role in decision making	-.031	.001	.40	61
Understanding of evaluation reports	.135	.012	.07	116
Quality satisfaction of evaluation reports	.12	.001	.45	102
Use satisfaction of evaluation reports	-.133	-.013	.08	115

of understanding of evaluation and somewhat more likely to be satisfied with the extent to which evaluation reports were used. Even though both of these correlations are close to significance at the .05 level, the substantive importance of the relationship is trivial--as indicated by regression coefficients of .01 and -.01 respectively.

The majority of non-respondents were project directors or advisory committee members. The major reason accounting for the lower response rate of advisory committee members is that we had limited resources to use in contacting persons who did not return the questionnaires and most of our efforts were to obtain complete responses from evaluators, agency personnel, and project directors (in that order). Thus, members of the advisory committees were not contacted as frequently nor with as much persistence as other persons.

In order to test for response bias within the group of project directors and advisory committee members, correlations were calculated for each of these groups separately (Tables A-5 and A-6). Project directors who were slow to respond do not appear to be signficantly different from those who responded quickly and the same pattern is shown for members of the advisory committees.
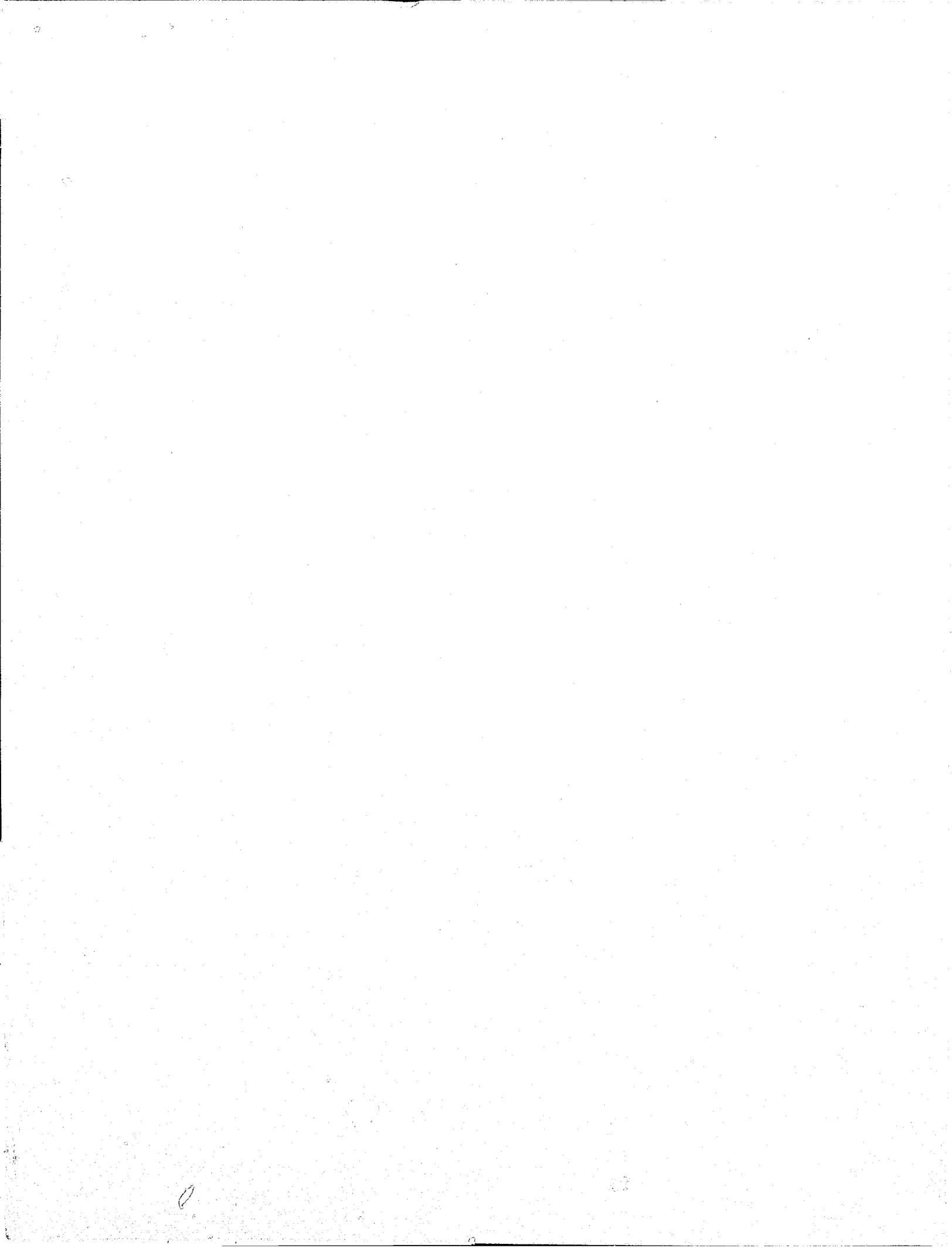
TABLE A-5

SUMMARY TABLE: Correlations between support and use questions and number of days
to return using only project directors responses

Survey Question	Correlation (R)	Slope (B)	Significance	N
Current weight of use of formal evaluations	-.229	.089	.07	43
Ideal weight of use of formal evaluations	-.067	-.018	.33	46
Number of reports reviewed	.0003	.00008	.49	53
Per cent of reports which were				
reviewed, but not used	.011	.0002	.48	23
considered	.246	.015	.13	23
played a major role in decision making	-.01	-.0005	.48	23
Understanding of evaluations reports	.15	.012	.14	53
Quality satisfaction of evaluation reports	.01	.001	.47	43
Use satisfaction of evaluation reports	-.057	.006	.35	50

TABLE A-6

SUMMARY TABLE:

Survey Question	Correlation (R)	Slope (B)	Significance	N
Current weight of use of formal evaluations	.14	.037	.20	39
Ideal weight of use of formal evaluations	.044	.01	.39	41
Number of reports reviewed	.147	.298	.18	40
Per cent of reports which were				
reviewed, but not used	-.364	-.002	.09	15
considered	.457	.013	.04	15
played a major role in decision making	-.188	-.006	.25	15
Understanding of evaluations reports	.172	.015	.13	43
Quality satisfaction of evaluation reports	-.154	.012	.19	36
Use satisfaction of evaluation reports	-.35	.030	.013	39

END