

UC Berkeley

IURD Working Paper Series

Title

Survey of State Program Evaluation Practices

Permalink

<https://escholarship.org/uc/item/5hm1b55c>

Author

Ridge, Susan Shea

Publication Date

1973-04-01

SURVEY OF STATE PROGRAM EVALUATION PRACTICES

Susan Shea Ridge

April 1973

Working Paper No. 209/RS20

Frederick C. Collignon, Project Director - Michael B. Teitz, Principal Investigator

Project for Cost Benefit Analysis and Evaluation of Rehabilitation Services

The research reported here is being supported by a grant from the Rehabilitation Services Administration of the Social and Rehabilitation Service, U.S. Department of Health, Education and Welfare.

## FOREWORD

Program evaluation is an activity which is receiving increased attention within rehabilitation programs. Evaluation not only provides feedback on how well program objectives are being achieved, but also can enable agencies to identify where and how program performance can be improved, and to anticipate problems which may arise in the future. At a time when the mandate of rehabilitation agencies is continuing to expand yet additional resources are not available, program evaluation becomes critical if agencies are to learn how to allocate and manage limited resources to achieve maximum effect.

In this study, Susan Shea Ridge, a research economist on the staff of the Institute of Urban and Regional Development, reviews evaluation practice in state agencies, using the results of a national questionnaire survey. This survey, which was initiated at the request of the Rehabilitation Services Administration of the U.S. Department of Health, Education, and Welfare, documents existing state practice. The survey also served the important function of identifying for R.S.A. for the first time those individuals and agency units across the country which had been assigned the principal responsibility for program evaluation in state agencies for vocational rehabilitation and services to the blind. The information provided by the survey has subsequently proved useful in planning for regional and national training workshops for state agency evaluation staff.

Because of the rapid and necessary development of evaluation activity within state agencies, the information concerning existing agency practice reported in this study will hopefully become quickly dated. We have already discovered a rapid turnover across time in the individuals and units assigned to program evaluation, in part because of the increasing emphasis which state agencies are placing upon evaluation and the resultant upgrading of the evaluation function within agency organizations. We would also put forward one reservation concerning the use of the survey results reported in this study.

It seems clear that not all of the responses of the state agency personnel reported in this study can be taken at face value. Program evaluation is a new activity in the rehabilitation field, and many state practitioners are not fully familiar with all the language and methodologies of evaluation. Although we attempted to define terms carefully in the questionnaire and pretested it in the California state agency, we perceive many inconsistencies in state agency responses and some probable misinterpretations of the intent or meaning of questions. The copies of studies and reports sent by many agencies with the completed questionnaires often did not reflect the degree of sophistication which the questionnaires seemed to indicate. The example of the discrepancy between reported follow-up studies and actual follow-up studies is reported in our summary (p. 11). It is also true that many agency staff differ in their understanding of the scope of program evaluation. State agency studies and routine reports revealed few state agencies

which were actually analyzing the impact of rehabilitation services upon clients, other than through the use of status 26 and other closure definitions.

We have attempted elsewhere\* to review the state of evaluation practice within the rehabilitation field. In this review, we drew extensively upon these questionnaires, but also upon our reading of numerous state studies, budgetary and P.P.B.S. documents, other reports by researchers and R.S.A. contractors, and upon our extensive field experience and discussions with agency staff in many states.

Frederick C. Collignon  
and Michael B. Teitz

---

\*Frederick C. Collignon, "An Overview of Program Evaluation Activity in Rehabilitation Services Programs: Current Status and the Problems Ahead," Institute of Urban and Regional Development, University of California, Berkeley, Working Paper No. 207/RS019, February, 1973. This paper was also delivered at a symposium on program evaluation in the rehabilitation field, conducted by A.R.C.A. at their San Diego national convention in 1973.

## ACKNOWLEDGMENTS

In reporting this study, acknowledgments should be made of assistance in the design and review of the questionnaire by Les Cole, John Noble, and Ronald Conley of the Rehabilitation Services Administration. Betty Dieckmann and Paul Mueller of the Department of Rehabilitation of the State of California gave generously of their time in pretesting the questionnaire, and contributed substantially to its development. Mr. Cole and Milton Schurr also provided invaluable assistance in helping the project achieve a higher rate of state agency response to the questionnaire. I would also like to acknowledge the support of the survey effort provided by the Institute on Rehabilitation Services and, in particular, the members of the Prime Study Group on Program Evaluation.

Within our project, invaluable contributions were made by Joan Weinberg, who assisted in data analysis and in the writing of sections of the survey summary, and by Barbara Thompson, who was responsible for the overall supervision of questionnaire processing and the numerous mailings and phone calls which proved necessary to achieve the high rate of state agency response.

Susan Shea Ridge

## TABLE OF CONTENTS

	page
SUMMARY OF STATE PROGRAM EVALUATION PRACTICES	1
APPENDIX I PROBLEMS AND PROCESS OF QUESTIONNAIRE DESIGN	23
APPENDIX II THE QUESTIONNAIRE	34

## SUMMARY OF STATE PROGRAM EVALUATION PRACTICES

The survey of State Program Evaluation Practices questionnaire was mailed to 79 state agencies for vocational rehabilitation. Every state or territory has a VR agency, but some have a separate agency for dealing with the visually handicapped. Thus, 51 of these agencies are "general" VR agencies dealing with a range of disabilities (orthopedic, blind, behavioral problems, etc.). The other 28 are agencies for the blind and visually handicapped only. The Survey was designed to determine how many states had established a program for evaluation of their activities, what type and level of evaluation was going on, and what the needs of the agencies might be in order to increase or improve their evaluation activity.

Of the 79 questionnaires sent out, 72 (or about 91%) were returned. Ninety-six percent (49) of the general agencies returned the questionnaire, compared with 82% (23) of the agencies for the blind. Two of the questionnaires were returned without being completed, but one was received too late for inclusion in this draft analysis. Thus, the statistics and information given below are based on 69 returned and completed questionnaires: 46 general agencies, and 23 blind agencies.

Of those who replied, 61% indicated that there was a component within the VR agency whose primary responsibility is VR program evaluation. Seventy-two percent of the general agencies had such components, but only 39% of the agencies for the blind. Additional



agencies indicated they were in the process of establishing such a component. Typical titles for the components were: Office of Program Planning and Evaluation; Program Development and Evaluation; Planning and Research; Research and Evaluation. This was a much larger number of evaluation units than we had expected and seems to indicate that the need for formal program evaluation is something of which most agencies are aware.

Looking at the agencies by size (measured by FY69 Section 2 expenditures), it might be concluded that the larger the agency, the more likely it is to have an evaluation unit. However, it should be noted that our sample of the larger agencies is quite small.

Size of agency Section 2, in thousands	# Responding *	# with Unit	# with Unit (of those re- sponding)
under \$10,000	31	20	65%
\$10 - 19,000	12	10	83%
\$20,000 plus	3	3	100%

\* general agencies only

If we break down the data by region<sup>1</sup>, we find that 100% of the agencies in region VII have evaluation units. In regions II,

---

<sup>1</sup>At the time of this writing, the most recent data available on agency size and region was from Fiscal Year '69, State Data Book: Federal State Program of Vocational Rehabilitation. At that time, there were nine regions and our data is based on this breakdown. It should be noted that as of February 1973, State agencies have been divided into 10 regions.

III, IV, V, and IX a relatively high percentage of agencies have evaluation units, while in regions I, VI, and VIII a low percentage have them.

Region of agency	# Responding	# with Unit	% with unit of those responding
I	10	3	30%
II	5	4	80%
III	7	5	71%
IV	5	4	80%
V	5	4	80%
VI	12	4	33%
VII	7	7	100%
VIII	6	2	33%
IX	12	9	75%

Another measure of the amount of evaluation activity in the state agencies is the number of man-years spent in planned evaluation. We specified that this should not include clerical help, data processors and other persons indirectly involved with evaluation. The breakdown was as follows:

Man-years spent	% of Agencies	
	All Agencies	General only
0	4.3%	2.2%
less than 1	30.4%	26.1%
1-3	27.5%	39.1%
4-6	10.1%	13.0%
7-9	1.4%	2.2%
10	1.4%	2.2%
Don't know/NA	23.2%	15.2%

If it can be assumed that those agencies which have not established evaluation components spend less than a man-year on evaluation, then more than half of all general agencies allocate less than one man-year to evaluation activity. Looking at the general agencies alone, the distribution was similar, but a larger proportion (39%) was in the 1-3 year category and a smaller proportion (only 2.2%) spent no man-years at all, or didn't know the number. The blind agencies fell most frequently into the "less than 1" or "1-3" categories and none spent more than 6 man-years. This and responses to other questions, indicates that the general agencies are more involved with evaluation than the agencies for the blind.

Again, when we break down the responses of general agencies by agency size, a possible trend for the larger agencies to be more active shows up. However, the small number of responses in the larger agency category precludes any definite conclusions.

Man-Years Spent on Evaluation, by Size of Agency Budget

# of Man-years spent	Number of Agencies*		
	Size of Agency Budget (in thousands)		
	Under \$10,000	\$10-19,000	\$20,000 +
0	1	0	0
less than 1	10	2	0
1-3	12	6	0
4-6	3	2	1
7-9	0	0	1
10+	1	0	0

\*general agencies only

If we break down the responses of all agencies by region we find that region IX agencies are more likely to spend more (over 4) man-years on program evaluation than are the other region agencies. Notably, all agencies which responded in region III spent approximately equal man-years on program evaluation. This may reflect either a formal or informal regional policy regarding man-year allotment to program evaluation.

<u>Region of Agency</u>	<u>Man-Years Spent on Program Evaluation</u>					
	<u>0</u>	<u>&lt;1</u>	<u>1-3</u>	<u>4-6</u>	<u>7-9</u>	<u>10+</u>
I	1	3	1	1	0	0
II	0	1	1	2	0	0
III	0	0	6	0	0	0
IV	0	1	1	2	0	0
V	0	2	2	0	0	0
VI	1	3	5	0	0	0
VII	0	1	4	1	0	0
VIII	1	3	1	0	0	0
IX	0	4	1	1	1	1

Backgrounds of evaluation personnel will be determining factors in the way program evaluation activity is undertaken. In some 45% of the agencies, personnel responsible for evaluation have had direct training in this area. Somewhat more of this training was obtained on-the-job (33%) than in special workshops or training programs in evaluation (28%). Examining the general backgrounds of evaluation personnel, one discovers that the great preponderance came out of rehabilitation counseling. Some 45 agencies' evaluation personnel had this background; 27 agencies had personnel with

professional background in education. Very few agencies reported personnel with background in disciplines such as economics or statistics. Law and Business administration were mentioned by a few agencies and "research" by some. Evaluation requires both substantive knowledge of the field being studied and methodological competence and familiarity with good research and experimental design methods. Neither can substitute for the other, but our results indicate that if the agencies are short on anything, it is the methodological background.

Most evaluations of state VR programs occur at the state central office level. Some 45% of the respondents indicated that district and local offices do perform evaluations either at the request of the state office or on their own initiative. However, only eight agencies reported that two or more man-years were spent in planned evaluation at the regional, district, or local levels within the state. Only three agencies reported regular employment of consultants and 43% said they were not used at all.

The agencies were asked whether various groups (e.g. supervisors, counselors, clients) were consulted during the course of an evaluation. They were also asked to break down their responses in terms of the stage in the evaluation process at which the consultation was made. The following indicates the number of positive responses.

<u>Phase of evaluation</u>	<u>Group Consulted</u>		
	<u>Regional/district Administration</u>	<u>Supervisors</u>	<u>Counselors</u>
establish priorities	36 (52%)	43 (62%)	31 (45%)
re: planning design	30 (43%)	31	14 (20%)
re: conducting evaluation	29	40 (58%)	25 (36%)
re: interpreting findings	30	38 (55%)	24
re: implementation	35 (51%)	42	28 (41%)

Administrators, supervisors, and counselors can be expected to have different perspectives from state central office personnel and from each other. Consultation with persons at these levels will provide unique inputs into possible factors to be considered in an evaluation, insight into actual workings of a program, and other valuable information. Whether or not these persons are consulted during the course of an evaluation will likely have implications for the potential implementation of changes suggested by the evaluation results. Familiarity with the evaluation and the confidence that all important factors were considered on the part of these field personnel will increase their confidence in the results and their willingness to implement any findings. The above table indicates that regional or district administrators, supervisors, and counselors are routinely consulted by, at most, half and often a smaller proportion of state agencies. Supervisors are consulted more often in almost all phases. Both supervisors and counselors are consulted least often with respect to planning of evaluation design -- but this is probably a phase in which methodological competence surpasses field knowledge in importance.

When asked the same question about client groups or representatives (the disabled themselves), only 15 agencies could say clients were consulted in any of the above ways. Interestingly, 10 of these 15 were agencies for the blind and only five were general agencies. This may reflect the fact that the blind, as a population are more organized than other disabled groups and that they have associations which "legitimate" the desire for participation in VR planning and the evaluation process.

Looking at the responses to this question according to size, for general agencies only, yielded no significant information. However, if we break down responses by region, we find that no agencies in region IV or region V consult clients in the evaluation process, and in none of the regions do more than one third of the agencies consult clients in the evaluation process.

<u>Region of agency</u>	<u># in region</u>	<u># consulting clients</u>	<u>% consulting clients inregion</u>
I	10	2	20%
II	5	1	20%
III	7	1	14%
IV	5	0	0%
V	5	0	0%
VI	12	4	33%
VII	7	2	29%
VIII	6	2	33%
IX	12	3	25%

An evaluator might well conceive that the disabled VR clients themselves can best decide whether the VR process has fulfilled

their needs. Yet from the limited response to this question, it is apparent that few agencies either share this conception, or at least have acted in this area.

A number of questions were asked to determine what (or who) is evaluated, and how the topics for evaluations are chosen. Twenty-three (33%) of the agencies had specific priorities for programs to be evaluated. Among the 30 or so specific projects mentioned, there was no clear-cut pattern. A handful involved evaluation of specific disability groups, such as the mentally ill or spinal disorders. Another half dozen involved the area of field services and case procedures. A few more referred to specific programs such as cooperative programs, training programs, or vending stand programs. Others varied from developing evaluation designs to following-up closed cases to studying "high producing" counselors.

There was more concurrence on the reason certain projects were given priority. The four most often cited reasons were as follows:

<u>Reason for priority</u>	<u>Number of agencies</u>	<u>Number specifying reason as the one most frequently used criteria</u>
scale	24	7
disability type served	19	3
purpose is of special interest	20	3
degree of problem with program	17	5

The scale of the program, making it more important relative to other programs, was both the most often cited criteria, and the one most often specified as of primary importance. Problems with a program or a crisis with a particular program ranked only fourth as a cited criterion for priority, but is ranked higher among the



criteria cited as most frequently used. Routine evaluation of all programs is not apparently undertaken by most rehabilitation agencies.

In response to questions about particular types of studies, about 43% of the agencies reported they had rarely or never conducted follow-up studies of clients after case closure. The other agencies reported that they had done follow-ups at least for a few closure groups. Most of the studies were reported as having been done for successfully closed clients only (26's). Only a very few agencies reported that they had ever conducted follow-up studies for all closed clients -- and most of these were agencies for the blind. A few of these studies had been conducted only for particular groups of interest -- the mentally retarded, public assistance recipients, alcoholics, clients rehabilitated into the home. In a few cases, the follow-up procedure was quite routinized and so could be repeated fairly easily at regular intervals. For instance, one agency for the blind routinely sent a questionnaire to each counselor one year after closure of each client. This was completed by the counselor in conjunction with the former client. In another case, public assistance recipients were followed-up every six months through a routine check with the welfare department. Obviously, there are different levels of evaluation involved here -- some agencies are employing just a check on whether or not a client is still employed, other agencies are studying a broader range of aspects relating to the client's rehabilitation. The fact that most follow-ups are done only for successful clients suggests that there is little attempt to assess the factors responsible for the difference between a successful and an unsuccessful closure, or to assess the impact

of the rehabilitation process on the employment and adjustments of a client who was officially closed "not rehabilitated." If unsuccessfully closed clients (08, 28, 30) were followed-up and found to be employed, this might suggest either that rehabilitation services yield delayed benefits or that, in some cases, it was not rehab services which were responsible for the 26-closures' improved employability. These questions are important ones for which follow-ups might provide some insight but our data indicates that little of this type of effort is being made.

One of the most obvious reasons for follow-up is to determine whether clients remain employed for longer periods than the relatively short 90-day period required for closure in the 26 status. Many of the follow-up checks were reported as conducted even sooner than 90 days after closure, and very few studies were conducted more than two years later. The breakdown of studies reported is as follows:

	<u>Time between Closure and Follow-Up</u>				
	<u>less than 3 mo.</u>	<u>3-6 mo.</u>	<u>7 mo. - 1 yr.</u>	<u>13 mo. - 2 yr.</u>	<u>2 yr+</u>
<u># studies</u>	26	31	42	32	8

While the largest number of studies fell into the "7 months - 1 year" category, too many seem to be conducted after only very short periods, and very few after periods long enough to provide a strong test of the stability of the employment of clients whose cases were closed as successful.

A closer look at follow-up studies seems to indicate that the reports of follow-up activity by state agencies are exaggerated. As part of an independent effort to assist R.S.A. in collecting

data for a benefit-cost study, a letter was sent to 44 state agencies which stated on the questionnaire that they had recently completed follow-up studies for periods longer than six months after closure. Of the 44 agencies, over half (some 23) of the agencies did not respond. Phone calls suggested that in many cases, no studies existed in printed form which could be distributed outside the agency even informally. That is, whatever follow-up studies were conducted were highly informal, consisting of sporadic phone calls by counselors, the ad hoc gathering of client testimonials, etc. Of the 21 agencies which did respond, 13 agencies said that they could not be of help for assorted reasons: the follow-up study proved to be for a period of time less than seven months, studies were "not available," studies did not actually probe employment and earnings status after closure, studies were of the informal kind described above, etc. An additional four agencies had not yet completed the study. Thus, ultimately only four of the 44 agencies were actually able to return follow-up studies which contained information on the post-closure earnings and employment of rehabilitants. This experience in following up the statements made on the questionnaire suggests that state agencies differ widely in their understanding of evaluation terms or alternatively in their perceptions of what constitutes the acceptable completion of an evaluation study.

Information on the types of statistics and measures which are routinely generated by state information systems gives further insight into what is evaluated. A number of measures were listed in the questionnaire, and respondents asked to indicate whether each was used in his agency, and whether routinely or occasionally.

<u>Measure or statistic</u>	<u>number of agencies using routinely</u>	<u>number of agencies using occasionally</u>	<u>Total</u>
Clients/counselor or man/year	44	7	51
Rehabs/counselor or man/year	48	6	54
Cost/rehab	53	11	64
% applications accepted	46	14	60
successful closures/ total closures	59	4	63
time from referral to plan	43	14	57
time from plan to closure	35	15	50
cost-benefit ratios	15	3	18
payback period	8	7	15
payback rate of return	9	11	20
other	5	5	10

As can be seen the first seven measures were reported as very widely used. Cost/rehab and the ratio of successful closures to total closures, for example, are generated in over 91% of the agencies. The latter is generated routinely by 74% of the blind agencies, and 91% of the general agencies.

If we break down the data by size of agency (general agencies only), we find that among the larger agencies (\$20,000 plus), the last three measures are used almost as often as the first seven measures. We are, however, limited in the conclusions which can be drawn from this because of the small number (3) of larger agencies responding.

If we look at the responses with respect to region we find that again the first seven measures are widely used, and that the percentage of agencies in a region employing a given measure decreases substantially for the last three measures, except in region II. Four out of five region II agencies used the eighth measure -- "cost benefit ratios." Region II agencies are consistent with other region agencies in their percentage drop for "payback period" and "payback rate of return," the ninth and tenth measures, respectively.

Most of these measures do not really provide information evaluating the rehabilitation process. The first measure, clients per counselor, is simply a monitoring of counselor work-load. While the most commonly used measure, successful closures per counselor, does provide an index of success, it is clearly only a first step toward a true evaluation of whether the program's objectives are being met in the most efficient and effective way possible. Measures which provide more information on the impact of rehabilitation services programs upon clients, the taxpayer, and the economy, are much less often used. Some of the "other" measures began to get more deeply into the evaluation of the program in meeting objectives -- (e.g., one agency measured the "% accomplishment of objectives") or to get at some possible reasons for differences between successes and failures (e.g., an agency routinely generated information on client characteristics).

About 37% of the general agencies report that they routinely break down their measures for each disability type and 24% more do this occasionally. Such a breakdown begins to provide more information on factors involved in success versus failure.

Evaluation of the non-monetary and non-employment-related benefits of the rehabilitation process is of interest to more and more people at all levels of the process, including the clients themselves. Because the measurement of such benefits is difficult and a relatively new area of interest in rehabilitation, it was not expected that many of the agencies would have developed measures of these benefits. Six of the agencies<sup>2</sup> said they had developed measures for assessing such non-monetary benefits as unpaid family work (housekeeping, childcare); improvement in functional ability, such as self-care or recreation; and emotional-psychological adjustment. Another 24 agencies,<sup>3</sup> or 35%, indicated they were working to develop measures in at least one of these areas. If we examine the responses with respect to size of agency, we find no significant correlation between agency size and present or planned development of measures of non-monetary benefits. Looking at regional data, the only significant statement which can be made is that region IV is the only region in which no agencies either have developed or have plans to develop measures for non-monetary benefits. Hopefully, increased coordination and communication among these agencies and other groups can lead to fruitful developments along this line.

A very important aspect of program evaluation -- in addition to who conducts the evaluation and what is evaluated -- is the data

---

<sup>2</sup>Arkansas, West Virginia, Massachusetts, New York, Virginia, Nebraska (blind).

<sup>3</sup>Alabama, California, Delaware, Kansas, Maryland, North Carolina, Oklahoma, Oregon (blind), Louisiana, Oregon, South Dakota, Texas, Arizona (blind), Michigan, New York, North Dakota, Ohio, Rhode Island, Vermont, Virginia, Wyoming, Arizona, Nebraska, Nevada (blind).

and methodology used. The strength of the evaluation design and methodology will determine how reliable the findings are.

Similarly, the better the design, the more likely it should be that the results will be accepted and implemented.

The data sources reported as most frequently used were: the federally-required R-300 reporting form, other routinely collected program data (such as some of the measures discussed above), and data collected by the agency in special studies. Other sources used were an additional state-required reporting form similar to the R-300, data made available by other parties (such as federal studies), or data from other agencies, such as workshops.

The use of control groups was very rare in state agency evaluations. Only two agencies reported routine use of control groups, and nine more agencies reported occasional use.<sup>4</sup> The importance of a control group in an evaluation should not be underestimated. It was suggested above that if follow-up studies of clients closed as not-rehabilitated and receiving few services were to show many of these clients employed, this might cast doubt as to whether successful completion of the rehabilitation plan is a major factor in the employment success of clients closed in status 26. Similarly, groups of clients who go through the rehabilitation process might usefully be compared with a group of persons similar in disability and demographic characteristics who had not been through this process to examine the extent to which the former groups' improvement is uniquely attributable to services received.

---

<sup>4</sup>Routinely used by Florida, occasionally used by Montana, Oklahoma, Washington, West Virginia, Wisconsin (blind), Virginia, Massachusetts, Ohio, Wyoming.

Replicative studies were also a fairly infrequent occurrence. Three agencies reported that they routinely (and 18 agencies reported that they occasionally) replicated studies, which would enable a comparison of a program at two or more different times or places. This type of replication would enable conclusions to be drawn as to whether a system that works in one type of area would work as well in another area. There are a number of reasons why a program might work differently at different times or in different places. There may be shifts in the employment level of the national economy over time. There may be major variations across local economies (e.g., a farm area versus a growing industrial area, areas of heavy and light unemployment). There may be major differences across areas in the availability and quality of supporting services from non-rehabilitation agencies. The following table shows that some agencies report that their studies do indeed take such factors into account, but that a significant portion of the agencies do not consider such factors.

In evaluating programs, has there been explicit consideration of any of the "situational" factors which might affect a program's outcome, but are beyond program control, such as conditions in the community and economy?

	<u>Yes-routine</u>	<u>Yes-occasional</u>	<u>Rarely</u>	<u>Never</u>	<u>Don't Know</u>
Conditions in national economy	11	14	7	19	6
Local Economy	16	18	6	15	5
Community attitudes	9	18	5	18	6
Other agencies in community	10	22	4	17	5
Other	3	2	2	3	1



The usefulness of evaluation results and the ability to implement those results must be considered in any discussion of evaluation activity. Hopefully, during the planning of the evaluation study, the evaluator thinks ahead to the time when implementation will occur. Attempts can be made to ensure that the results will be meaningful and that implementation will be feasible.

A first step in this direction would be to attempt to assess or "evaluate" the probable effectiveness of programs when they are only in the proposal stages. This would involve specifying explicitly what the objectives of the program are and how these objectives might be measured. It would involve careful analysis of the program to determine its critical elements, how they might be adjusted or adapted if necessary, and whether a particular program fits the needs of the particular area or client group for whom it is being considered. When asked whether the agency was involved in the formal evaluation of proposed programs -- that is, evaluation of programs not yet started, 39% of the agencies said they were. To what degree their evaluations covered the points outlined above was not determined.

Similarly, the expansion of a given program requires special information. Expansion will often involve serving more severely disabled or less easily rehabilitated persons -- thus making incremental costs and effects different from average costs and effects. When asked if evaluations had ever projected the effects of a program expansion, almost 40% of the agencies had done this, although only half of these had explicitly considered the differences between incremental and average costs and effects.

One of the problems with evaluation has been that courses of action which appear to be the logical result of evaluation findings are not taken. This problem of "no results" from evaluation may be due to unanticipated negative findings, the failure to plan ahead for the possible changes which might be warranted after the evaluation, or to a variety of other reasons. Some program evaluators suggest that a type of "contingency analysis" would mitigate against this problem. Prior to an evaluation, an explicit attempt should be made to identify the alternative courses of action to be taken for each possible outcome of the evaluation. A very small number of rehabilitation agencies has apparently done this; some 70% of the agencies either report that they had not done such contingency analysis or do not answer the question. However, 48% of the agencies said there had been informal planning along these lines, in addition to the 25% who do perform such analyses, at least occasionally.

As to results of evaluations that had already been performed in the states, 42 of the agencies were able to give examples of program changes due to evaluation. Examples of this were changes in counselor assignments; subdivision of a metropolitan area office into three offices; improved case recording; establishment of uniform counselor procedures throughout the state; or establishment of a training program for welfare department personnel focusing on the referral process. Twenty-two states gave examples of program changes resulting from federally sponsored evaluations. The most often mentioned example was the giving of priority to welfare recipients. Another example was elimination of duplicate accounting and checking activities after a federal audit; or the resetting of

cost objectives and objectives for number of 26 closures. Very few agencies (7) gave examples of changes resulting from another state evaluation, indicating that there is little communication between states concerning evaluation findings.

The following questions were asked about the significance of evaluation results:

Do you believe your evaluation activities have led to significant changes in program strategy, resource allocation among programs, or efficiency with which a given program strategy was implemented?

A number of the states were able to say that significant changes had occurred. The following table shows the responses breakdown for this question. Such a response attests to the value of program evaluation when it is conducted.

	<u>In program strategy</u>	<u>In resource allocation</u>	<u>In efficiency of implementation</u>
Several significant changes	17	12	18
1 or 2 significant changes	12	21	12
Several minor changes	7	3	7
1 or 2 minor changes	4	4	3
None	5	5	4
Don't know	4	5	5

Results of evaluations might be even greater if study findings were disseminated more widely. The survey indicated that evaluation findings generally take the form of briefings, memos, and "in-house" reports. Only two states routinely distributed studies to rehabilitation agencies in other states. Evaluations findings were most often shared among state central office personnel,

district and local administrators, supervisors and counselors in the state. Federal regional S.R.S. offices were also sent the results fairly frequently, but from our earlier information that few agencies had been affected by other state evaluations, it is doubtful that the Regional Offices are in turn disseminating these findings. Region VIII agencies did seem to rank higher than others in disseminating data to R.S.A. and the regional offices. However, in order to determine if regional offices are in turn disseminating findings which they receive, additional information is needed.

A final series of questions was presented with the intention of determining in a summary way the amount and type of evaluation being done in the states. The responses to these questions seemed very inconsistent with the rest of the questionnaire, however, indicating that the questions were not well understood. Three levels of evaluation were described, as below, and the respondents asked for various programs and groups, whether each type of evaluation was done routinely or occasionally:

- Level A-Routine:           audit, budgetary-control, monitoring staff performance
- Level B-Program:        measures of effect of a program on target population; evaluation of a program with respect to outcome and cost; analysis of causes of and barriers to outcome; ability of a program to be carried out successfully.
- Level C-Comprehensive: comparison of alternative strategies and programs with respect to costs, side-effects of programs, how programs achieve objectives

As expected, level A evaluation was reported as performed most often, followed by B and, finally, the more sophisticated and time-consuming level C. However, the number of positive responses for

level C exceeded what the answers to the rest of the survey seemed to indicate. A plausible explanation of this inconsistency is that informal evaluative thinking is done along the lines specified in the definition of level C. It seems clear, however, that only very infrequently is a formal evaluation design set up and conducted to answer such questions.

Overall, the results of the survey indicate that a significant number of state agencies are aware of the need for evaluation and at least to some degree have begun to set up divisions to work in this area. Some of the respondents' comments and letters indicated that the receipt of this and other similar surveys has served to increase state agency awareness of the need for evaluation. A number of evaluative measures and statistics are currently generated, but apparently are not used to greatest advantage. One problem is that most evaluation personnel do not have formal training in evaluation design, program analysis, and methodology, or in the complicated problems involved in the evaluation process. Of course, this is a field in which specific training has not been readily available. The backgrounds of evaluation staff are primarily in the substantive rehabilitation field, rather than in fields which might provide methodological "handles" for evaluation design and thinking. Evaluation, at this point, has largely focused on measuring the success of the rehabilitation process in terms of numbers, and has not been oriented as much to getting behind these statistics to determine what is responsible for program or client success or failure, what alternatives are available, and which alternative achieves objectives in the most effective and efficient way.

## APPENDIX I

## PROBLEMS AND PROCESS OF QUESTIONNAIRE DESIGN

This paper deals with the process and problems involved in designing a particular survey questionnaire. The survey discussed was mailed to 79 state agencies dealing in the vocational rehabilitation (VR) of the physically disabled and blind. Called a "Survey of State Program Evaluation Practices," the questionnaire probes the administration, methods and impact of state activity directed toward evaluating programs of service to the disabled. Do these state agencies have mechanisms for evaluating their programs? How sophisticated are their methods? Have evaluation results been instrumental in effectuating program changes and improvements? How can a project such as the Project for Research on Benefit-Cost Analysis and Resource Allocation for Rehabilitation Services tailor its activities to meet the needs of states? These were the sorts of questions to which we sought answers.

Purposes.

A questionnaire is significantly shaped by the purposes of the survey. Such a statement may seem to belabor the obvious: the purpose of the survey has been rather succinctly stated above, and it seems obvious that questions will be included in the questionnaire to elicit the required information. However, as we will see, a more explicit delineation of all the aims of the survey is needed if the questionnaire is to be as useful as it might.

At this point, our discussion will be facilitated by examining the particular purposes of our survey and seeing their effects on the content and style of our questionnaire. The survey was intended not only:

- 1) to determine, and more importantly, document the status of evaluation activity in state agencies,

but also

- 2) to establish the value of program evaluation,
- 3) to create an awareness of efforts (e.g. by certain states and by our project) to improve evaluation techniques
- 4) to establish a network, however informal, for coordination and exchange of information in the area of evaluation.

If the purpose of the questionnaire were limited to determining the status of evaluation activity, the questionnaire could have been shorter and simpler. Since program evaluation is a relatively new area of involvement for most of the relevant agencies, we could be fairly sure that only a limited amount of activity was going on and that methods were fairly simple. Thus, we could perhaps have eliminated questions asking if sophisticated marginal analysis techniques, data sources, environmental controls, etc. were used. Researchers might well challenge our presumption of the conclusions of our survey at the outset of survey design. However, since the survey had as its purpose the documentation of the status of evaluation activity, our presumptions needed to be treated as hypotheses and tested.

A second goal was to establish the value and necessity of evaluation, as well as provide indications of the particular attributes of methodologically good evaluations. This second purpose suggests that our questionnaire include reference to as many facets of good

program evaluation as possible. Of course, constraints -- such as length of the questionnaire and others discussed below -- limit the extent to which the questionnaire can pursue that second goal. But the content of the questionnaire is significantly different because of this second purpose.

This example makes clear the role of the various purposes of the survey play in the design of the questionnaire. The importance of the purposes will be seen again when the iterative process used to design the questionnaire is discussed.

### Content

The questionnaire, in light of our purposes, had to constitute a treatise on evaluation in question and answer form. Therefore, the primary source for the questionnaire's content was the body of literature on program evaluation. Using both articles dealing with evaluation in general and more specific memo-type "literature" dealing with evaluation of vocational rehabilitation programs, we considered evaluation in its administrative, organizational, methodological and procedural aspects. We considered the purposes of evaluation and the impact and effectiveness of evaluation. We developed typologies of evaluation, differentiated according to purpose, method, type of program evaluated, and level of sophistication or comprehensiveness. Some of our major sources were Suchman's "A Model for Research and Evaluation on Rehabilitation,"<sup>5</sup> Glennan's "Evaluating Manpower

---

<sup>5</sup>Howard A. Suchman, "A Model for Research and Evaluation on Rehabilitation," in Marvin Sussman, Sociology and Rehabilitation.



Programs,"<sup>6</sup> the Urban Institute's Federal Evaluation Policy,<sup>7</sup> and in-house memos from RSA and SRS.

Our method was to develop an initial memorandum-style list of the attributes and requirements of good program evaluation, as discussed in the literature. From this starting point, we took the questionnaire through a series of iterations in which, based on our own discussions and consultations with persons trained in survey design, we constantly revised and improved the questionnaire in terms of content, style and clarity.

The question of the appropriate respondent to the survey was our first problem. Among the 50 states, there is a diversity of organizational placements of the vocational rehabilitation agency. Some are autonomous agencies, whereas others are subordinate departments within some other department, such as Education. Some states have separate agencies dealing with the blind, whereas others combine service to the blind with that for all the disabled. The placement of the evaluation function, then, varies with each of these arrangements and may even vary among states where the agencies are similarly

---

<sup>6</sup>Thomas K. Glennan, Jr., Evaluating Federal Manpower Programs: Notes and Observations, the Rand Corporation, Santa Monica, California, 1969.

<sup>7</sup>Joseph S. Wholey (et. al.), Federal Evaluation Policy, the Urban Institute, Washington, D.C., 1970.

placed. Perhaps the Director of VR has full-time evaluation staff under him. Perhaps he has some personnel who do some minimal evaluation. Perhaps the "parent" agency for vocational rehabilitation (e.g. Department of Education) performs the evaluation function, in part or in toto. To which of these groups do we address our questionnaire? Naturally, we are interested in evaluation as it exists at each of these levels. But we wanted to ensure that our survey reached that person (or persons) who had primary responsibility for evaluation. This problem could have been partially solved for us by RSA sponsorship of the survey: the fact that RSA takes responsibility for the mailing of our survey, and directs it to the states through its regional organization, takes advantage of the regional offices' knowledge of the appropriate respondent. Unfortunately, although the questionnaire was designed at RSA request for RSA use, administrative clearance problems arose. RSA ultimately decided that we should independently send out the questionnaire, under the aegis of the Institute on Rehabilitation services, which was sponsoring a task force on program evaluation. We used the first question in the survey to both give us information as to the various evaluation agencies and to make clear to the respondent our interest in that office closest to the VR agency with evaluation as its primary responsibility.

RSA interest in and support of the survey eased or eliminated other potential problems. Our questionnaire was certain to be only one among many that the agency receives constantly. Such items are not often given the highest of priorities, and may be delegated to rather low-level staff persons. With RSA support, we had less worry about length of the questionnaire (questionnaires which appear

too cumbersome may find their way quickly to the wastebasket). We still had to consider that an overly lengthy questionnaire might yield diminishing returns to later questions, but we could be more assured of a high overall rate of response. A final benefit of RSA interest was our ability to "check" on certain answers -- for instance, the question as to whether state evaluation agencies communicate evaluation findings to RSA.

#### Iterative process

During this process of numerous revisions, our main concern was making the proper choices in light of the different trade-offs involved: including a given question might make the survey more complete, but would also make it longer and more time consuming for the respondent; wording a particular question in one way might make it more exact, but wording it slightly differently would increase the probability of the respondent understanding it easily; choosing response sets required some trade-offs between completeness, feasibility, and complexity.

Changes in content. As indicated above, length of the questionnaire was a major concern here. Certain questions were dropped because we knew they had been asked in other surveys. Naturally, we may have wanted to "replicate" the other studies (especially since some of these surveys only went to a sample of agencies), but we had to consider the attitude of the potential respondent who could react to our questionnaire as a useless, redundant waste of time. Some questions were eliminated because the information could be better obtained by phone follow-up. This was especially true of detailed

information. For instance, one question concerns the development of measures of non-monetary benefits to clients (e.g. emotional/psychological adjustment). Our questionnaire simply asks whether such measures are being or have been developed. Naturally, when they are being developed we will want as much information as possible on what measures are being used, etc. But since we suspect that very few agencies will be working in this area, we do not include such "detail" questions -- they would be irrelevant to most respondents. Follow-up by phone can be conducted for those few agencies which warrant it.

Style. Recall that the purposes of our questionnaire included not only gathering information for our use but providing as much information as possible to respondents about good program evaluation practices. We wanted our questions to be informative. Thus, the question "Have replicative studies ever been conducted?" became "Have replicative studies ever been conducted, i.e., to evaluate a particular program in two different time periods or places or to check the results of an earlier evaluation?" It clarifies such a term as "replicative study" for respondents who might not be familiar with it, and also gives an indication of the value of replicative studies -- why one might consider doing them.

We were also concerned with the tone of our questionnaire. Realizing that program evaluation is a new area of activity for many agencies, we did not want to intimidate respondents with a barrage of questions about things they never did or could never hope to do because of staff, time, and money limitations. Thus, we emphasized in our introduction to the questionnaire our realization

that this was, indeed a new area. We provided "No," "Don't Know," "Rarely" responses to every question. (We even considered placing the "No" response ahead of the "Yes"). As an example of a change in tone, we changed the response choices in our question about whether samples were used from

_____ Yes		_____ Yes
_____ No	to	_____ Rarely
_____ Don't Know		_____ Never, we want full 100% representation
		_____ Don't Know

Clarity. Paramount in our minds here was the fact that respondents to the questionnaire might range from DVR chiefs, to heads of evaluation staff, to evaluation staff members, or even to staff who have no direct relation to program evaluation. Similarly, we had to deal with the fact that "evaluation" has a second -- entirely different -- meaning in the VR field. A prospective VR client is "evaluated" in terms of his disability and his potential. We had to avoid confusion between this type of evaluation and program evaluation. Furthermore, we wanted to distinguish between formal program evaluation and the informal, more routine "monitoring" activity that any supervisor would do. This was the function of the introductory statement. We also included a section in the questionnaire where different levels of evaluation were described.

Clarity was a concern with each individual question as well -- especially because of the range of respondents we could expect. Thus as with the example given in the section on style above, we worked at making each question as self-explanatory as possible -- defining sophisticated terms, illustrating, giving examples, anticipating possible responses.

Response sets. Not all questionnaires provide response sets -- checklists, multiple choices, etc. -- for each question. Questions can be left open-ended. We chose to provide the response choices for various reasons: Such a questionnaire is easier to fill out and less time-consuming for the respondent; providing responses also makes the questionnaire easier for us to read and code -- there will be uniformity of answers; and providing response choices increases the informational value of the survey. Consider the following example:

"Have you developed, or are you attempting to develop, some measure of any non-monetary benefits so that these benefits may be included in the evaluation of programs or comparison of alternative programs?"

	<u>have developed</u>	<u>attempting to develop</u>	<u>have not worked in this area</u>
unpaid family work (e.g., housekeeping, child care)	_____	_____	_____
improvement in func- tional ability not job-related (e.g., self-care, recreation)	_____	_____	_____
emotional/psycho- logical adjustment	_____	_____	_____
other: please specify	_____	_____	_____

\_\_\_\_\_

The inclusion of the response choices makes the question much more informative by giving the reader an idea of the types of non-monetary benefits which might be important.

The design of response sets does involve problems: the range of response choices must be appropriate and complete, yet flexible. The problem of appropriateness arises especially with regard to numerical responses. If, in inquiring about man-years

spent in evaluation, we provide a choice of "1-5 years, 6-10 years, 10-15 years," and no state spends more than five years, the responses will be meaningless. We must do research as to likely numbers of years so that the correct range of responses will be offered. If we are going to ask about the background of evaluators, and we provide a checklist, we want to include all the likely backgrounds (completeness) -- rehab counselling, education, public administration, mathematics, etc. -- but we want to allow for the unusual or unanticipated response (flexibility). Thus, we provide an "other: please specify" response for many questions.

A final problem with response sets is frequency, or what I call the "sometimes" problem. One can ask, "Are control groups used in evaluating programs?" If one provides a response choice of "\_\_\_ Yes \_\_\_ No \_\_\_ Don't Know," where does the agency fit which has used a control group once, maybe twice in the past few years? We solved this problem in many cases by providing a more complete response set, such as "\_\_\_ Yes, routinely \_\_\_ Yes, occasionally \_\_\_ Rarely \_\_\_ Never \_\_\_ Don't Know." In certain other questions, where the range of responses was already length or complicated (a matrix might be offered), an additional request to indicate frequency would not only be overly complex for the respondent but would also complicate coding. Here, the trade-off between completeness and complexity had to be weighted for each individual case. In some cases, we indicated that a "yes" answer should only be given when a procedure was used routinely. In others, we had to depend on additional comments from the respondent to indicate frequency.

The pre-test. A survey pre-test (in such a case as this mail survey) usually involves mailing the questionnaire to selected respondents typical of or similar to the group for which the survey is designed. Return of the completed questionnaires is awaited, and answers are assessed for relevance -- did the respondents seem to interpret the questions as intended? Because our questionnaire was rather technical and we anticipated a need for substantial revision on the basis of a pre-test, we felt a need for direct consultation with the "pre-test" respondent. Naturally, this involves greater costs. Thus, our "pre-test" consisted of direct consultation with evaluation staff at California DVR, who independently filled-out the questionnaire. We had to trade quantity in a pre-test and diversity (sending it to a number of states) for what we felt was most fruitful -- receiving comments directly from respondents. Because California is one of the most advanced states with regard to program evaluation, we could interpret its experience as a sort of "upper limit." Thus, while we pre-tested in only one state, we know where this state fit in the spectrum and could allow for this.

These are some of the issues and problems we contended with in designing our survey. Perhaps the discussion of the forces which shaped our questionnaire and the factors we considered in revising it will be generalizable to other survey designs, and in conjunction with the questionnaire itself, will be useful to those interested in evaluation as well.



APPENDIX II  
THE QUESTIONNAIRE

The following questionnaire was initially mailed to all state Vocational Rehabilitation and Blind agencies. After eight weeks, at the request of RSA which desired a 100% response if possible, a second copy of the questionnaire was mailed to nonresponding agencies. After several more weeks, again at RSA request, phone calls were made by project staff to all nonresponding agencies requesting that they complete the questionnaire. In several cases, the Federal regional offices were called first to identify the correct contact within the state agencies. A third copy of the questionnaire was mailed to those agency staff in charge of evaluation who stated by phone that they had never received the questionnaire. After a few more weeks, RSA staff made calls to several state agencies which still had not responded, and also urged completion of the questionnaire at one or more conferences with Regional office staff and state agency directors.

During this period of repeated contacts, which prolonged the completion of the survey from March to September, 1972, a series of preliminary tallies of questionnaire responses were made available to RSA staff. The initial paper reporting the results of the survey was made available in April, 1972, following the analysis of the responses from the first mailing of the questionnaire. This draft paper was circulated at the meeting in Minneapolis of the Institute on Rehabilitation Services, which sponsored a study group on program evaluation for rehabilitation services programs.

Copies of all questionnaires will be made available to any interested state agency or research group, at Xeroxing and mailing cost. The estimated cost as of February, 1973, for obtaining a complete set is \$30.00. Punch-cards are also available. Copies of the code sheets and card format are, of course, available for those who desire copies of the punch cards.

## SURVEY OF STATE PROGRAM EVALUATION PRACTICES

There is a growing interest among agencies of all types, at all levels, in formal program evaluation. In the field of vocational rehabilitation, where the need for services exceeds the resources available to public and private agencies serving the disabled, we must be able to establish priorities among alternative programs in order to achieve maximum output from limited resources. Further, evaluation of rehab programs can provide the type of evidence -- of return for the rehab dollar -- which will demonstrate that increased investment in rehabilitation services is warranted.

This is a new area of concern for state rehabilitation agencies. Few states have as yet embarked on any sort of full-scale program of evaluation. The purpose of this survey is to obtain a clear idea of the current status of formal program evaluation in state vocational rehabilitation agencies. We wish to determine how frequently or infrequently formal program evaluation is conducted in state agencies; what type of evaluation is conducted; for which programs; and to what end. Our aim, as an RSA funded project in Cost-Benefit Analysis and Evaluation Design, is to provide materials and designs that will respond to agency needs regarding program evaluation. An understanding of where formal evaluation stands, and a knowledge of state agency priorities for program evaluation, is essential to this task.

You will note that our real interest is in formal program evaluation, as opposed to the type of monitoring activity which any good supervisor or administrator performs. While certain questions do refer to this more routine form of evaluation, the survey is mainly concerned with the more comprehensive program evaluation.

None of the questions should require time-consuming search for answers. Where such an effort might appear necessary, your "best guess" will be appropriate.

Please list below the name of the chief evaluation officer or individual immediately responsible for evaluation activities, whom we might inform of the results of this survey and of other evaluation studies being undertaken by this and other university research projects.

Name \_\_\_\_\_

Title \_\_\_\_\_

Agency \_\_\_\_\_

State \_\_\_\_\_

Phone \_\_\_\_\_

I. The following questions concern administrative aspects of evaluation.

1. a. Is there a component within the state VR agency or bureau whose primary responsibility is VR program evaluation?

\_\_\_\_\_ Yes \_\_\_\_\_ No \_\_\_\_\_ Don't Know

Note: If "Yes", give title of this office: \_\_\_\_\_  
REFER TO THIS OFFICE IN ANSWERING THE REMAINDER OF THE QUESTIONNAIRE.

- b. Within the state government, what other offices or departments are responsible for evaluation of VR programs?

\_\_\_\_\_ Human Relations Dept.      \_\_\_\_\_ GSA  
\_\_\_\_\_ Dept. of Education      \_\_\_\_\_ Legislative Analyst  
\_\_\_\_\_ Dept. of Finance      \_\_\_\_\_ Other, please specify:  
\_\_\_\_\_

NOTE: If you have answered "no" or "Don't Know" to question 1.a., which office in 1.b. has primary responsibility for evaluation of VR programs? \_\_\_\_\_  
REFER TO THIS OFFICE IN ANSWERING THE REMAINDER OF THE QUESTIONNAIRE.

2. a. How many man-years are spent in planned evaluation each year at the state level? Include for example, program analysts, research analysts, special project evaluators. Do not include persons indirectly involved with evaluation, such as clerical help, data processors, etc.

\_\_\_ 0 \_\_\_ less than 1 \_\_\_ 1-3 \_\_\_ 4-6 \_\_\_ 7-9 \_\_\_ 10+ \_\_\_ Don't Know

b. How many man-years are spent in planned evaluation each year at the regional, district, and local levels? Again, do not include clerical help, etc.

\_\_\_ 0 \_\_\_ less than 1 \_\_\_ less than 2 \_\_\_ 2-3 \_\_\_ 4-5 \_\_\_ 6+ \_\_\_ Don't Know

c. Do any of these persons have formal training in evaluation, either through their academic specialization or job-related training?

\_\_\_ Yes \_\_\_ No \_\_\_ Don't Know

If "yes", how many people obtained this training through:

Special workshops or training programs in evaluation \_\_\_\_\_  
 Informal on-the-job training \_\_\_\_\_  
 B.A. \_\_\_\_\_ in what field(s) \_\_\_\_\_  
 M.A. \_\_\_\_\_ in what field(s) \_\_\_\_\_  
 Ph.D. \_\_\_\_\_ in what field(s) \_\_\_\_\_

d. What are the general (i.e., not specifically related to evaluation) backgrounds of personnel involved in evaluation? Respond by indicating the number of personnel with each of the following backgrounds.

___ Rehab Counseling	___ Economics
___ Education	___ Math
___ Sociology	___ Psychology
___ M.D.	___ Other, please specify:
___ Public Administration	_____

3. a. Do district and local offices ever conduct formal program evaluation, either at your agency's request or on their own?

\_\_\_ Yes, for particular projects at state request  
 \_\_\_ Yes, they initiate formal evaluations on their own  
 \_\_\_ Both of the above  
 \_\_\_ Rarely  
 \_\_\_ Never  
 \_\_\_ Don't Know



5. a. We would like to identify programs for which we could provide demonstration evaluation designs that would be of value to state agencies. Have priorities been established for evaluation among programs managed out of your agency?

Yes  No  Don't Know

b. If yes, which program(s) have priority? (If a memo is available regarding such priorities, you may attach a copy. In lieu of this, you may list the priorities here or, if you prefer, send a letter indicating your priorities or suggestions of programs for which demonstration evaluation designs would be useful.)

c. On what basis are priority programs for evaluation chosen? Would a program be chosen for evaluation because of a particular interest in the client disability type served by the program, because the purpose of the program is of special interest, because the scale of the program or its cost makes it important relative to other programs, because of problems with its execution, or are all programs evaluated in some sequential order? Check as many as apply. If you check more than two criteria, circle the one most frequently used.

- scale of program (# of clients, cost,...)
- purpose of program is of special interest
- disability type(s) served
- degree of problem with program execution (e.g. program in a crisis situation)
- all programs are evaluated in a sequential order
- other: please specify \_\_\_\_\_
- Don't Know



II. In the next group of questions, we are interested in the "methodology" of evaluation: what data and statistics are used; what is measured; how often evaluations are conducted; types of studies done.

1. What sources of data do you use in obtaining information needed to evaluate programs? Check all appropriate answers. If you check more than two, circle the two most important data sources.

- R-300 reporting form, federally required
- R-300 reporting form, additional state requirement
- other routinely collected program data
- data collected in special state-agency studies
- data acquired in contracted studies
- data collected by other parties (e.g., university or federally contracted studies) and made available
- data voluntarily reported (e.g., workshop reports)
- other; please specify \_\_\_\_\_

2. a. Which of the following evaluative measures and statistics are routinely generated by state information systems, for internal planning and management?

	routinely	occasionally
1. Clients/counselor or man-year	_____	_____
2. Rehabs/counselor or man-year	_____	_____
3. Cost/rehab	_____	_____
4. % applicants accepted	_____	_____
5. Successful closures/total closures	_____	_____
6. Length of time from referral to plan	_____	_____
7. Length of time from plan to closure	_____	_____
8. Cost-benefit ratios	_____	_____
9. Pay back period	_____	_____
10. Pay back rate of return	_____	_____
11. Other: please specify _____	_____	_____

b. Are any of the above measures broken down for each individual disability type?

Yes, routinely      Please Specify: \_\_\_\_\_  
 Yes, occasionally      "      "      \_\_\_\_\_  
 Rarely      "      "      \_\_\_\_\_  
 Never  
 Don't Know

c. Are any of the above measures generated for particular client disability types only?

Yes, routinely      Please Specify: \_\_\_\_\_  
 Yes, occasionally      "      "      \_\_\_\_\_  
 Rarely      "      "      \_\_\_\_\_  
 Never  
 Don't Know

d. If cost-benefit ratios are generated, what discount rate is used?

---

3. Have you developed, or are you attempting to develop, some measure of any non-monetary benefits so that those benefits may be included in the evaluation of programs or comparison of alternative programs?

	have developed	attempting to develop	have not worked in this area
unpaid family work (e.g., housekeeping, childcare)	_____	_____	_____
improvement in functional ability not job-related (e.g. self-care, recreation)	_____	_____	_____
emotional/psychological adjustment	_____	_____	_____
other: please specify	_____	_____	_____

4. a. Have follow-up evaluations -- of clients after case closure -- been conducted?

Yes, for most statuses

Yes, for a few statuses

Rarely

Never

Never

Don't Know

Please specify statuses: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

b. How long after closure were these studies usually conducted?  
You may check as many as appropriate.

	less than 3 mos.	3-6 mos.	7 mos. - 1 year	13 mos. - 2 years	More than 2 years later
successful closures (26)					
ineligible clients (08)					
voluntary drop out clients (08, 28, 30)					
clients closed who received plan services but were not rehabilitated (28, 30)					
other: please specify					

5. a. Have your evaluations ever used a sample of relevant clients, rather than all those involved in a particular program?

Yes, routinely  
 Yes, occasionally  
 Rarely  
 Never, we want 100% representation  
 Don't Know

- b. If yes, which of the following sample types have been used? Check all which are appropriate.

random  
 stratified  
 not chosen by statistically designed procedure  
 Don't Know

6. Have control groups ever been used in evaluating a program?

Yes, routinely  
 Yes, occasionally  
 Rarely  
 Never  
 Don't know

7. Have replicative studies ever been conducted, i.e., to evaluate a particular program in two different time periods or places or to check the results of an earlier evaluation?

Yes, routinely  
 Yes, occasionally  
 Rarely  
 Never  
 Don't know

8. In evaluating programs, has there been explicit consideration of any of the "situational" factors which might affect a program's outcome, but are beyond program control, such as conditions in the community, and economy?

	Yes, routinely	Yes, occasionally	Rarely	Never	Don't Know
conditions in national economy					
conditions in local economy					
community attitudes toward disabled					
other private and public agencies serving disabled in the communiyt					
other: please specify					

III. The following questions concern the impact of evaluation on programs and policy.

1. Is your agency involved in the formal evaluation of proposed programs -- that is, evaluation of programs not yet started -- in addition to the evaluation of ongoing and existing programs? You may check more than one.

Yes, we do formal evaluations of programs proposed within our own agency.

Yes, we are involved in formal evaluations of federally-proposed programs prior to their initiation.

No

Don't Know

2. Expansion of a given program will often involve serving more severely disabled, or less easily rehabilitated, persons -- thus making incremental costs and effects different from average costs and effects. Have any evaluations projected the effect of a program expansion?

Yes, considered incrementals as well as averages

Yes, but did not explicitly consider incremental differences

No

Don't Know

3. One of the problems with evaluation has been that courses of action which appear to be the logical result of evaluation findings are not taken. This problem of "no results" from evaluation may be due to unanticipated negative findings, or it may be due to failure to plan ahead for the possible changes which might be warranted after the evaluation. Prior to an evaluation, has there been an explicit attempt to identify the alternative courses of action to be taken for each possible outcome of the evaluation?

Yes, routinely

Yes, occasionally

Rarely

No, but there has been informal planning

Never

Don't Know

4. Can you give an example(s) of a program expansion, cancellation, internal change or other change due to an evaluation? Specify briefly.

---

5. Can you give such an example which resulted from a federally-sponsored evaluation? Specify briefly.

---

6. Can you give such an example which resulted from information from another state program evaluation? Specify briefly.

---

7. We are interested in how widely your evaluation studies and results have been distributed. Do you disseminate findings and proposed implementations of evaluations...

	routinely	occasionally	on request
to RSA in Washington	---	---	---
to VR Regional offices	---	---	---
to non-evaluation personnel at the state level	---	---	---
to district and local administrators	---	---	---
to supervisors and counselors	---	---	---
to clients	---	---	---
to VR agencies in other states	---	---	---
to non-VR agencies in your state	---	---	---
to private or quasi-public agencies	---	---	---
to professional organizations	---	---	---
to university groups involved in this field	---	---	---

8. Do you believe your evaluation activities have led to significant changes in program strategy, resource allocation among programs, or efficiency with which a given program strategy was implemented?

	in program strategy	in resource allocation	in efficiency of implementation
several significant changes			
one or two significant "			
several minor "			
one or two minor "			
None			
Don't Know			

IV. These final questions request a sort of "summary" of the types and frequency of evaluation conducted by your office. In answering, please refer to the following description of evaluation levels.

- Level A - Routine:                   audit, budgetary-control, monitoring staff performance
- Level B - Program:                   measures of effect of a program on target population; evaluation of a program with respect to outcome and cost; analysis of causes of and barriers to outcome; ability of a program to be carried out successfully
- Level C - Comprehensive:           comparison of alternative strategies and programs with respect to costs, side-effects of programs, how programs achieve objectives

1. Please indicate whether the following programs are evaluated routinely or occasionally at any of the above-described levels.

	Level A		Level B		Level C		not managed out of this office
	routinely	occasionally	routinely	occasionally	routinely	occasionally	
Basic Services (section 2)	___	___	___	___	___	___	___
Innovation (section 3)	___	___	___	___	___	___	___
Research and Demonstration (section 4)	___	___	___	___	___	___	___
Training of Personnel	___	___	___	___	___	___	___
Program Expansion	___	___	___	___	___	___	___
Workshop & Facility Programs (sections 12&13)	___	___	___	___	___	___	___

2. For those programs (in question IV.1) which are routinely evaluated, are there any which are evaluated, not as a total program, but only for particular groups, projects, or offices?

only for particular disability groups of interest?           \_\_\_ Yes   \_\_\_ No   \_\_\_ Don't Know

If yes, specify program and level A, B, or C. \_\_\_\_\_



only for particular local projects?  Yes  No  Don't Know  
 If yes, specify program and level A, B, or C. \_\_\_\_\_

only at particular district or local offices?  Yes  No  Don't Know  
 If yes, specify program and level A, B, or C. \_\_\_\_\_

3. Please indicate whether programs for the following disability groups (regardless of funding source) are evaluated routinely or occasionally at any of the above described levels.

	Level A		Level B		Level C		not managed out of this office
	routinely	occasionally	routinely	occasionally	routinely	occasionally	
Public Offenders	___	___	___	___	___	___	___
Drug Addicts	___	___	___	___	___	___	___
Alcoholics	___	___	___	___	___	___	___
Mentally Retarded	___	___	___	___	___	___	___
Blind	___	___	___	___	___	___	___
Public Assistance Recip.	___	___	___	___	___	___	___
SSDI Recip.	___	___	___	___	___	___	___

4. Note: Omit this question if no programs are routinely evaluated. For those which are routinely evaluated (in question IV.3), are any evaluated:

only for particular local projects?  Yes  No  Don't Know  
 If yes, specify program and level A, B, or C. \_\_\_\_\_

only at particular district or local offices?  Yes  No  Don't Know  
 If yes, specify program and level A, B, or C. \_\_\_\_\_

5. When reports of evaluation studies are made are they in the form of briefings to staff members, informal memos, written reports for "in-house" use, or published reports? You may check as many as appropriate.

briefings  written reports for "in-house" use  
 informal memos  published reports  
 some combination or sequence of the above: please specify:  
 \_\_\_\_\_

6. Please attach copies, if available, of two or more program evaluation studies. If you have conducted evaluations using samples or control groups, it would be helpful to include examples. If you have conducted cost-benefit studies, an example would be very useful. In general, we would like to obtain examples of 1) a study which is typical of evaluation activity in your state and 2) a study which you consider outstanding in quality and impact. Please mark documents to indicate whether they are examples of sample studies, cost-benefit studies, use of control groups, or of a typical or outstanding study.
7. Please list any other agencies which you are aware are involved in evaluation of state VR programs; for example, departments or agencies to which the VR agency is responsible, budget bureaus, other state executive agencies, federal agencies.
- 

Thank you for your time. Your answers will be helpful to us in directing our efforts toward areas of greatest need and interest to state agencies.

Please return to:

Professor Frederick C. Collignon  
Institute of Urban and Regional Development  
University of California, Berkeley  
Berkeley, California 94720