DOCUMENT RESUME

ED 212 958 CG 015 752

AUTHOR Reagles, Kenneth W., Ed.; Crystal, Ralph M., Ed.
TITLE Program Evaluation for Rehabilitation: A Book of Readings from "Program Evaluation for Rehabilitation Agency Personnel," Syracuse, New York, May 4-6, 1978. Michigan Studies in Rehabilitation, Series I, Monograph IV.
INSTITUTION Michigan Univ., Ann Arbor. Rehabilitation Research Institute.
SPONS AGENCY National Inst. of Handicapped Research (ED), Washington, D.C.
PUB DATE 81
GRANT NIHR-15P-59021/5
NOTE 190p.; Some tables are of marginal legibility. For related documents, see ED 169 402 and ED 182 514.

EDRS PRICE MF01/PC08 Plus Postage.
DESCRIPTORS Anthologies; *Evaluation Criteria; *Evaluation Methods; Program Evaluation; *Rehabilitation Programs; Self Evaluation (Groups); *State Agencies; State Programs; *Training Objectives; Vocational Education; *Vocational Rehabilitation; Workshops

ABSTRACT

This monograph contains presentations from a training program that was designed to link state vocational rehabilitation agency personnel with the current technology of program evaluation in rehabilitation settings. Part One discusses a model program evaluation unit in state rehabilitation agencies, state agency program evaluation units from the perspectives of federal and state agencies, and university resources in program evaluation research and practices. The papers in Part Two focus on program evaluation standards and data resources for rehabilitation. Part Three on program evaluation methods and measures contains several papers about various evaluation issues such as case review schedules, benefit-cost analyses, and management information systems. Lists of contributors and trainee-participants are also included. (NRB)

Michigan Studies in Rehabilitation

Series I, Monograph IV

General Editors:

Don K. Harrison
Juliet V. Miller

PROGRAM EVALUATION FOR REHABILITATION:

A BOOK OF READINGS

From "Program Evaluation for Rehabilitation Agency Personnel," Syracuse, New York, May 4-6, 1978.

Edited by

Kenneth W. Reagles

and

Ralph M. Crystal

The University of Michigan

Rehabilitation Research Institute

Ann Arbor, Michigan 48109

1981

Published by

The University of Michigan
Rehabilitation Research Institute
1323 School of Education Building
Ann Arbor, Michigan 48109

The Michigan Studies in Rehabilitation are supported by Research Grant #15P-59021/5 from the National Institute of Handicapped Research, U.S. Department of Education.

Interpretations and opinions expressed in this publication do not necessarily represent the interpretations or opinions of the Department of Education or the National Institute for Handicapped Research.

TABLE OF CONTENTS

Chapter

FOREWORD

The apparent coincidence between the increased attention given to the severely disabled by the state-federal vocational rehabilitation program and the increased efforts of the evaluation of the program is not, of course, a coincidence at all. The Congressional Oversight Hearings of the early 1970's, chaired by Representative Brademas of Indiana, provided a national forum which witnessed a loud outcry of dissatisfaction from consumer and advocate groups representing the severely disabled. Congress was so moved by the testimony collected at these hearings that its intentions were reflected in the Rehabilitation Act of 1973. A major policy shift was reflected in the legislation - the severely disabled were to be given priority for services.

Not only did Congress mandate priority services for the severely disabled, it backed up its intent with another mandate. Congress mandated that the Secretary of the Department of Health, Education and Welfare report annually on the progress made towards the priority goal of serving the severely disabled. Reporting implies some sort of data collection. The impact of the requirement to report on progress annually was to dramatically accelerate the generation of program evaluation capacity within the rehabilitation system, especially within individual state rehabilitation programs.

While the state-federal program had served the severely disabled prior to the Act and conducted some evaluations of its efforts, the increased activity towards meeting the ends intended by Congress provoked controversy, confusion, and concern. The controversy stemmed from the concurrent demands made upon the system to respond to the totality of needs of eligible severely and non-severely disabled persons. The confusion revolved around the definition of who was or who was not "severely disabled." Concern was expressed for the future: What were the ultimate implications of the abrupt change of program policy? Thus, the increased attention upon the severely disabled and the greater efforts towards the evaluation of the state-federal rehabilitation program are tied intimately together.

In the spring of 1977, Syracuse University's Rehabilitation Counselor Education Director, Dr. Kenneth W. Reagles, submitted an application to the Rehabilitation Services Administration for funding of a short-term training program entitled, "Program Evaluation for Rehabilitation Agency Personnel." The proposal was reviewed favorably and the training was conducted on May 4-6, 1978, in Syracuse, New York. Approximately 120 individuals representing state vocational rehabilitation agencies, both general and blind, attended the training. The trainees were predominately program evaluation specialists, researchers, and some administrators. The purpose and content of the training were intended to link such persons with current technology of program evaluation in rehabilitation settings. An Advisory Committee was formed and assisted Professor Reagles with the planning and content of the training. The members were:

Rodney Pelton
Director, RSA Division of Evaluation
U.S. Department of Education/OSERS/RSA
Washington, D.C.

Robert Struthers
Coordinator of Program Evaluation
Michigan Department of Education
Bureau of Rehabilitation
Lansing, Michigan

Ralph Crystal
Research Director
Rehabilitation Research Institute
The University of Michigan
Ann Arbor, Michigan

Stanley Sadofsky
Program Evaluation Specialist
Regional Office
Rehabilitation Services Administration
New York, New York

Reed Greenwood
Director of Research
Arkansas Research and Training Center
University of Arkansas
Fayetteville, Arkansas

The objectives of the training were:

1. To impart knowledge and understanding of the historical antecedents to current program evaluation efforts in the state-federal vocational rehabilitation program

2. To increase knowledge and understanding of the origin and development of the General Standards for Evaluation, including:

   a. knowledge of the original Standards

   b. knowledge of the limitations of the original Standards

   c. knowledge of the review and modifications of the original Standards

3. To examine the evaluation Standards within the context of the total program evaluation effort and the methods of collecting program evaluation data so that the individual states may be able to respond to the reporting mandates

4. To understand the concept of "performance level" associated with the evaluation Standards, their origin, development, limitations, and potential application

5. To engender an appreciation of the various audiences of program evaluation data and results concerned with vocational rehabilitation agencies

6. To increase an appreciation of the need for management reports with which program evaluation data and results may be fed back to the various audiences of vocational rehabilitation information

7. To further the development of knowledge and skill in methods and procedures of utilization of program evaluation data and results for the purpose of program development

8. To impart knowledge and skill in methods and procedures for utilizing program evaluation data and results for the purpose of program planning

9. To prepare and publish an account of the training content for dissemination to vocational rehabilitation agency personnel, as well as to persons interested in program evaluation in rehabilitation settings

A compilation of the presentations made by persons participating in the short-term training in Syracuse is contained herein. Although considerable time has passed since the presentations, the issues addressed by the program evaluation authorities are still relevant. Moreover, the presentations may be viewed from a historical perspective to gain an appreciation for the developments in rehabilitation program evaluation. In addition to presentations made at the short-term training conference, additional readings have been included which are related to the topic of the training.

Kenneth W. Reagles

Ralph M. Crystal

PREFACE

The University of Michigan Rehabilitation Research Institute (UM-RRI), co-sponsor of this publication, was established in 1976 with funding from the National Institute of Handicapped Research in response to the mandate of the Rehabilitation Act of 1973 that rehabilitation programs and projects be evaluated. The UM-RRI's efforts are directed toward research and related activities to assist states in evaluating management practices and service delivery systems.

The UM-RRI has been working on several long and short range objectives in rehabilitation program evaluation to:

1. Develop alternative conceptual models that may be used as a framework for comprehensive program evaluation in the state-federal rehabilitation program

2. Conduct research on existing program evaluation instruments to determine their feasibility for current use and to determine their need for additional development and validation

3. Identify, design, develop, test, validate, demonstrate, and disseminate program evaluation instruments, techniques, and methodologies that are consistent with conceptual models for comprehensive program evaluation in rehabilitation

4. Develop criteria for designing, developing, testing, and validating new and existing program evaluation instruments, techniques, and methodologies that consider measurement of impact, effectiveness, effort, efficiency, and output

Production of this monograph is consistent with the objectives outlined of the UM-RRI.

Ann Arbor
August, 1981

Don K. Harrison

CONTRIBUTORS

LANA C. BRENES is a Training Specialist with the San Diego State University Rehabilitation Center.

MARGARET BROWN is Project Coordinator of the Rehabilitation Indicators Project at the New York University Medical Center.

CHARLES P. COLE is Director of the Abt Associates project to revise the Rehabilitation Services Administration Management Information System.

RALPH M. CRYSTAL is Assistant Professor in the Rehabilitation Counselor Education program and Research Director of the Rehabilitation Research Institute at The University of Michigan.

LARRY GLASS is a researcher with the Rehab Group, Inc.

DON K. HARRISON is Associate Professor and Director of the Rehabilitation Counselor Education program and Director of the Rehabilitation Research Institute at The University of Michigan.

BRUCE MALOOF is Director of Rehabilitation Studies for Abt Associates and is involved with the project to revise the Rehabilitation Services Administration Management Information System.

FRED McFARLANE is Director of the Rehabilitation Counseling Program at San Diego State University.

JOHN H. O'NEILL is with United Cerebral Palsy of New York State.

RODNEY PELTON is Director of the Evaluation section within the Rehabilitation Services Administration Policy and Planning branch.

STANLEY PORTNY is with Portny and Associates in Alexandria, Virginia.

GARY PRAZAK is with Walker and Associates in Minneapolis, Minnesota.

KENNETH W. REAGLES is Professor of Special Education and Rehabilitation at Syracuse University and Director of the short-term training program from which this monograph was produced.

SUSAN SHEA is Vice President in Charge of Operations for Berkeley Planning Associates

ROBERT D. STRUTHERS is Coordinator of Program Evaluation for the Michigan Bureau of Rehabilitation

ROBERT M. THRALL is Professor in Jones Graduate School of Business Administration, Department of Mathematical Sciences, Rice University in Houston, Texas.

RALPH R. TURNER is a Project Manager with Abt Associates and is affiliated with the project to revise the Rehabilitation Services Administration Management Information System.

SARA WAGNER is Project Director for JWK International in Annendale, Virginia.

PART ONE

A PROGRAM EVALUATION UNIT IN STATE REHABILITATION AGENCIES

INTRODUCTION

A Program Evaluation Unit in State Rehabilitation Agencies

The theme of the conference revolved around the concept of an ideal or model state agency program evaluation unit. For too long a rather ambiguous conception has existed of what a program evaluation unit ought to be, what it ought to include, and what capabilities it ought to have. To provide a perspective, it is necessary to define the dimensions of a model or ideal state agency program evaluation unit and then examine that ideal in light of what actually exists. Necessary variation in program evaluation units exist with regard to placement of the unit within the administrative structure of the agency, staffing patterns, internal and external information needs, and the mandate to the unit provided by the rehabilitation agency.

Providing information to management on the functioning of the rehabilitation program is a primary function of the program evaluation unit. Certain reporting minimums are necessary to respond to needs for evaluation data at the federal and state levels. Sufficient capacity should exist to provide a response to the unique information needs of the various audiences within each state rehabilitation agency. Thus, a commonality or generic component appears to exist in the evaluation unit which is considered to be common to all state agencies. Evaluation units should also be capable of responding to the unique evaluation needs of the respective states. Further, the need exists for the evaluation unit to become more pro-active, i.e., to plan evaluation studies. Often in the past, evaluation studies have been conducted in reaction to federal or state reporting requirements. It is important that a program evaluation unit have the capacity to translate the expressed agency evaluation needs into specific studies.

At the conference three perspectives were presented on the parameters of a model program evaluation unit. First, from the Rehabilitation Services Administration (RSA) in Washington, Rodney Pelton discussed a Central Office perspective. Second, Robert Struthers of the Michigan Bureau of Rehabilitation presented a view of program evaluation from the vantage point of a state rehabilitation agency. Third, Ralph Crystal, from The University of Michigan Rehabilitation Research Institute (UM-RRI) gave an overview of a programmatic research thrust in program evaluation in a university setting which may assist both RSA and state agencies in program evaluation capacity development and utilization.

CHAPTER I

DIMENSIONS OF A STATE AGENCY PROGRAM EVALUATION UNIT: A FEDERAL PERSPECTIVE

Rodney Pelton

Recent Developments

Before discussing model state programs, an overview will be provided of some of the new developments in rehabilitation evaluation. The Rehabilitation Act of 1973 is being revised, and will introduce, for the first time, a broad program of projects on independent living. Emphasis will be placed on utilization of information from the rehabilitation program.

New developments are underway. For example, a data link is under development with the Social Security Administration on job and wage information. New evaluation Standards are being developed. The developmental phase of the new Standards will be completed in FY 1978; the Standards will be pre-tested in FY 1979, and put into operation in FY 1980. (This timetable has since been revised.) Also, a new medical and vocational facilities information system is being developed. The design for the information system will be completed in FY 1978; it will be pre-tested in FY 1979, and put into operation nationally in FY 1980-1981. A revision of the RSA information system will start later this year and proceed over a period of appproximately 18 months. It will include information inputs from all RSA projects. In addition, new information will be available on the system sometime in the time period of FY's 1980-1983.

Measures of functional limitations and rehabilitation indicators will change the approach from defining disabilities conceptually to measuring them. Measures will become available of the degree to which restoration and training of a severe disability is accomplished. A great deal is happening in cost-benefit analysis. A nationwide Delphi has been completed to identify rehablitation benefits. A potential exists in cost-benefit analysis which particularly focuses on physical restoration.

Among other recent developments is a study in the State of Arkansas on weighted case closures conducted by the Arkansas Rehabilitation Research and Training Center. Another exciting activity is the development of a taxonomy of rehabilitation, a project at Texas Tech, which began with a grant from the National Science Foundation. The main taxonomic dimensions will contain classifications of disability, rehabilitation intervention, and function change, which will assist in planning; establishing objectives; identifying goals; categorizing services, costs, outcomes, and client populations.

Obstacles to State Agency Evaluation

RSA has encountered some problems in establishing evaluation systems in state agencies. First, program planning and development does not generally exist at the level needed for evaluation. Evaluation demands measurable objectives which often go beyond the existing planning technology or the resources available for planning in many agencies. The planning which exists at a measurable and objective level is often fragmented. Few comprehensive plans are interrelated to identified problems. Very often, managers are not explicit about their objectives, and evaluators fail to insist on the guidance needed to define evaluation criteria. Moreover, covert uses of evaluation have been found to exist especially when evaluation, as a rational enterprise, is undertaken for non-rational reasons. Use of evaluation for non-rational reasons may be to delay a decision or to justify a controversial decision. Too frequently attempts are made to shape evaluation designs or findings.

Evaluations are sometimes used to publicize already successful programs. Evaluation may be used only to fulfill grant requirements, rather than as a management opportunity. When evaluation units are established, too often they are placed with other units with most resources going to units other than evaluation. New evaluation systems are often viewed as an unnecessary burden. Some people consider evaluation systems as collecting and recording useless information; others say that evaluation systems dehumanize the process of helping people. Still others say that evaluation systems are an encroachment on the professional's domain. Evaluation systems are particularly vulnerable to budget squeezes. Many feel that diverting funds from direct services to evaluation is unjustified even though evaluations help with resource allocation.

In a study of VR program evaluation units (Ridge, 1973), the following was found: (a) units at that time were under development, but limited in size; (b) the staff had limited evaluation training; (c) staff was minimally involved in the evaluation process; (d) evaluation studies tended to be crisis oriented and focused on specific programs, explaining problems after the fact; (e) case review was used to evaluate case flow, but less frequently to evaluate follow-up of cases; and (f) the R-300 was the primary data source.

A Model Evaluation System

An overview will be presented of a model evaluation system. An evaluation system should be based on a conceptual framework of the information needed for decision making in administration, data processing, planning, budgeting and fiscal controlling, program developing, and maintaining consumer interaction and interagency linkages. Evaluation affects decisions at the strategic or policy making levels, and it affects decisions on the

operational and tactical levels. The evaluator interacts with agency personnel at the conceptual level and the empirical level.

Five steps are part of the evaluation process. First, determine the purposes of evaluation, i.e., whether it is for program justification, planning, organizational development, or organizational change. Second, determine the levels of evaluation, whether it is of the total state agency, program management, or service delivery. Third, determine the type of evaluation, whether it's input evaluation, process evaluation, or output evaluation. Fourth, determine the evaluation criteria; for example, the extent of the program, number of clients served or measures of effort, effectiveness, impact, or efficiency. Fifth, describe methodology and measures, and develop additional methodologies and measures when necessary.

A broader prospective of evaluation as a process may emerge after describing the program. Before initiating a formal evaluation, it may be necessary to provide a detailed description of the program to be evaluated. Understanding the program is important and insight may develop after the program has been described. Additional judgments about the program are created through the process of evaluation. Formulating recommendations for change and selling the recommendations to decision makers are critical factors. Implementing recommendations for change is the final step in the evaluation process.

The Management Process

The management process will be described and related to evaluation. The responsibility for program evaluation lies with the manager of the program. Where should the organizational placement of evaluation be in the management system? It should be on a level consonant with its mission. To determine its mission, questions need to be asked: Is evaluation directed at policy questions? Is evaluation directed at specific program variations?

The evaluation unit should maintain a greater degree of autonomy than other support functions. It has a specialized and sensitive function; it identifies with the control functions of management; it is frequently subjected to political unilization; it frequently has to respond to questions of credibility. Evaluation units should be placed at the highest organizational level where they are accountable and responsible for assessing the direct delivery of services. It needs to be underscored that evaluation is a top management function. As a top management responsibility, evaluation may avoid disadvantages of internal evaluation units where evaluation exists under the actual program unit which is being evaluated.

Another concern is presented by placing the evaluation function outside

the organization providing the services. Under such a structure, knowledge is missing of the enormous complexities of the real world program. Evaluation may be limited to concerns of only economy and effort, since little substantive knowledge exists of the program. An argument made for placement of the evaluation function outside the organization pertains to credibility. However, placement of evaluation within the highest level in the agency also has credibility.

Models of Input, Process, and Outcome

Program evaluation can be considered using a systems theory input, process, and outcome model. Consider the processor, or the process, to be the organization itself. The organization can have feedback on its performance or outcomes, so that it may correct the amount and kind of input and process activities that occur.

Taken separately, input evaluation is divided into two parts: (a) the evaluation of effort, and (b) needs assessment. Effort evaluation is concerned with types of clients, services, staff time, and monetary expenditures. In short, it relates to the extent and type of program effort undertaken. Effort evaluation is usually routinized into statistical systems and budget justifications. Evaluation of effort assesses utilization of resources in program priorities and changes in the demand for services. Because an evaluation of effort is essentially just accountability, it needs to be used in combination with other evaluation procedures to have credibility in evaluation. Evaluators should be sure, in effort evaluation, that the use of statistical reporting systems, budget narratives, and other sources of effort assessment are not duplicated.

The second part of input evaluation is needs assessment. To be useful for evaluation, needs must be assessed at more than one point in time to determine if needs have changed. Needs assessment methods may utilize key informants, such as community agency leaders or community forums. RSA has used the participatory planning process, one of which was on evaluation in 1977. Clients receiving rehabilitation services can be examined. The evaluation of clients in the program can be examined in terms of incidents and prevalence of disability. Social indicators are indirect correlates of need. One example of a social indicator is the amount of delinquency as an indicator of need for mental health services. The survey approach to needs analysis is the most valid and reliable and also the most costly.

In turning attention away from input to process evaluation, it should be remembered that the rehabilitation agency is the processor. Some issues are critical in process evaluation. First, program monitoring provides impressionistic data about how projects are going. Second is client tracking, which includes acceptance, transfer, movement and exit

of clients from the system. The actual can be compared with the ideal. Client tracking can improve the efficiency and the responsiveness of the agency to client needs. Third is cost accounting, or just accounting, which pertains to the cost of delivering different types of services. However, the use of accounting by itself has a danger of making the cheapest the best, thereby confusing efficiency with effectiveness.

Another aspect of process evaluation is compliance with requirements: using legal requirements rather than standards which are professionally developed. Usually legal requirements are concerned with safety and protection of program participants, rather than with standards of professional practice.

The adequacy of program standards constitutes another issue of process evaluation. Standards can apply to outcomes, but sometimes they apply to program processors. Standards define the professional estimates or statistical calculations of what the program should be doing at a minimum. Standards have the advantage in that they are repetitively collected data. An administrator can look at what's happening in the agency longitudinally.

Finally, process evaluation is concerned with goal directness. Reference is not to achievement of goals or outcomes, but reference is made to the degree to which the program is actually sticking with stated goals, objectives, and policies.

Outcome evaluation addresses the question of whether clients have changed as a result of receiving services. Outcome evaluation is often answered by follow-up surveys, frequently the most satisfactory form of outcome evaluation. Measurements can occur at different points in time, and control groups and even sampling procedures used when possible. Follow-up surveys can closely approximate the experimental design of a before and after control group configuration. A limitation of surveys is that they can't establish direct cause-effect relationships and linkages with the effects at follow-up, and services previously offered. The existence of an outcome effect does not automatically prove that the effect is due to the services provided and was the cause of the outcomes. Many uncontrolled intervening variables exist that may have caused the outcome.

To attempt to handle uncontrolled intervening variables, evaluators have used client satisfaction as a moderator variable. If the outcome following rehabilitation services is positive and the client is satisfied with the service, the assumption may be made that the service contributed to the outcome. The use of client satisfaction as a moderator variable is for the purpose of controlling the uncertainty of the relationship of service activity to client outcomes. Client satisfaction measures should be used in combination with other measures.

Another outcome measure is client goal attainment, as specified in

the IWRP. The use of the IWRP limits outcome evaluation to goal achievement related to specific problems which the client presents, rather than to more general outcome measures such as "adequate functioning" and Status 26, closed rehabilitated closures. Efforts are made to pinpoint a client's success rather than placing the client in a general class of good outcomes. Other service outcomes pertaining to training, education, and production activities are harder to measure because they are indirect. Under such conditions, it's difficult to measure success, except when goals are explicitly stated.

Types of Evaluation

Cost-benefit analysis is a promising approach to evaluating outcomes. The rehabilitation program may be ahead of other human service programs because of the research currently underway in cost benefit analysis. Cost-benefit combines a concern for business management with human services accountability. In other words, the cost of the program is related to the benefit to people. Cost analysis includes the monitoring of client movement and the identification of where and how costs are incurred in that process. The benefit analysis includes defining and isolating units of benefit. However, isolating units of benefit is a major conceptual and methodological task.

Impact evaluation, another type of evaluation, assesses the impact the program has on the target population--a difficult problem. Programs are intended to alleviate widespread problems. Some of the measures used in impact evaluation are social indicators. Social indicators, which indicate need, present a problem if the indicator changes at a later time; there is no assurance that the program really caused the change and the reduction of need.

Survey evaluation is another way of assessing the impact of a program where measurement is made of the level of existing need at various times. Cost-benefit analysis, impact evaluation and survey evaluation are types of evaluation which a program administrator can use to manage a program. Evaluation is a management function. However, the manager not only has responsibility for evaluation, but for other management functions which may complement evaluation.

Important management functions include planning and implementation. Implementation is critical to the location and allocation of resources, coordination with other programs, selling the program to users, sponsors, and coping with personnel morale. The implementation function involves evaluation least, although sometimes the evaluator compares program implementation with process standards; the issue of client tracking is sometimes related to implementation. Since the rehabilitation program pursues multiple goals with limited resources, the manager is usually concerned about how to maximize the effectiveness of the program.

Controlling the Program

Evaluation can assist in answering many questions to increase program control. For example, where should resources be directed? This question relates to need assessment. Which of several alternative ways should be supported to reach similar goals? This question pertains to process and outcome evaluation. How should resources, particularly human resources, be organized? This question relates to program continuation, program effectiveness, and program efficiency.

Another consideration is obtaining a balance in resource allocation between service delivery and administration. How much money should go into administration of the program and how much money should go into delivery of services? Administration includes the money that is invested in program evaluation activities.

Linking Evaluation and Planning

The relationship between evaluation and planning (note the similarity between planning and evaluation) has been called the link between knowledge and action. In other words knowledge, without planning, can't be put into action. Or one can go into action, but one may not use knowledge effectively unless one plans. Planners are concerned with identifying problems, stating goals and objectives, costing alternative programs, identifying key program variables, unit costs, cost trend analyses, correlation of variables, program linkages, problems in data processing, supervision and control over performance, evaluation and monitoring of costs, identification of program evaluation gaps as well as setting the focus of future evaluation.

The focus of program evaluations should come out of planning. Some planning tasks are to develop long and short-range plans, to link with other plans (for example state plans and federal plans), to make recommendations for resource allocation to the administrator, and to make recommendations for action and development of legislation. Note that some planning and research techniques involve evaluation. For example, management by objectives -- which is really based on a monitoring system -- may have evaluation overtones. Evaluators are also interested in most of the topics of interest to planners. All planning methods have in common the requirement that outcomes be specified for an ongoing program within a specified time frame. Evaluators should be involved in the planning process to assure that the structure and process of planning is such that evaluation is possible.

In specifying objectives, the objectives must be realistic and measureable. Some of the steps that evaluators take in defining program objectives (administrators should insist that evaluators be associated with planners in the development of program objectives) are to define

multiple criteria by which programs are to be judged, to assure that objectives coming from the legislature, policy makers and administrators are addressed, and to develop objectives with language which incorporates desired program results. The favorable setting for evaluation is one in which legislation specifies a clear mandate for program evaluation. Often objectives for program operations have to be extracted by reading between the lines. This can be a difficult task, and constitutes one reason why evaluation and planning should be linked.

Evaluation and Program Development

A few observations will be made regarding the relationship of evaluation to program development. Evaluation, in relation to a program, should assist in: (a) identifying the potential target population; (b) defining the needs of the potential target population; (c) specifying client outcomes based on that need; (d) designing program activities to achieve objectives; (e) monitoring program activities and measuring the client achievement of outcomes; (f) making judgments or evaluations about the programs; and (g) making program changes or creating new programs. Just as evaluation is linked with planning and management, it is also linked with program development, monitoring, and outcome evaluation.

Evaluation and Data Processing

In the relationship between evaluation and data processing, evaluation has the responsibility to increase the reliability and validity of routinely collected data. To develop data which will indicate program effectiveness and permit comparisons of cost, process, output, and follow-up, one must also establish a system which permits the convergence and linkage of program and budgetary information for the purpose of cost-benefit considerations. To promote the use of standardized measures, it is essential that measures are used nationally or if used in other states, the measures should be standardized within the state.

Many evaluation designs are faulty because the data base is too limited. As reliable and valid measures are developed, it is the evaluation specialist's responsibility to introduce those measures into the data system so that these valid and reliable measures can increase the possibilities for evaluation. It's also necessary to consider the multiple outcomes of the rehabilitation process. Variables should be included in the data base which would allow measurement of psychological and social functioning, physical functioning, economic independence, independent living, vocational functioning, and various intermediate outcomes resulting from restoration, counseling, and training. A need exists to recognize and include in the data base those multiple variables impacting on rehabilitation outcomes, including variables in the broader socio-economic context, such as the number of competitive placements related to the economic structure of the state. A computerized data base is very important in evaluation. The data base should not exist without evaluator knowledge

of the multiple possibilities existing for evaluation within the data base. A conceptual framework should link all the data in the evaluation system to agency objectives and the potential solution of problems.

Conclusion

In summary, state evaluation model units will attempt to address some of the issues identified. The model evaluation units should relate directly to the target population, goal structure, and interaction with state agencies. Model units should clearly analyze and form linkages between agency components and the environment in which the agency operates. Relationships include those between state evaluation model units and universities for the development of new techniques and technologies, particularly in conceptual and methodologically difficult areas such as program impact. These linkages should be strengthened with the federal evaluation program. The state agency should read federal evaluation requests for proposals (RFP's) to keep abreast of developments in the federal program. An evaluation plan should exist which is either developed or updated at the beginning of each year and which might range over a two to three-year period. The emphasis of the evaluation plan should be on feasible studies of major programs. Development of state agency commitment and personnel involvement is important in data-base decision making, or decision making as far as possible using information obtained by conceptually guided evaluation studies.

A variety of information should be available to make decisions about rehabilitation programs. Evaluation should involve administrators, supervisors, and counselors in the design of evaluation studies. Whenever possible, evaluation studies ought to be linked with other studies such as those which may be concluded by financial management units in auditing. Evaluation units should establish formal systems to disseminate and implement evaluation results. Model evaluation units should involve consumers in evaluation, both in the planning of evaluation and in the follow-up to determine client satisfaction. The evaluation unit should use outside consultants and expertise for instrument development, RFP preparation, and training. Lastly, evaluation units should provide for adequate long and short-term training.

A considerable amount can be learned by doing, but learning about evaluation occurs by taking advantage of other opportunities. For example, the training materials at Research and Training Centers, Rehabilitation Research Institutes, and Research Utilization Laboratories should be available to everyone. Adequate funds should exist to purchase relevant professional evaluation and research literature, and adequate provisions should be made for training in statistics, profile analysis, rehabilitation indicators, outcome measure selection, study design, sampling, development of implementation plans, construction of measures, information systems and computers, and writing RFP's. An

evaluation unit has a lot of responsibility but the type of interaction required of an evaluation unit is widespread when it is developed and linked with the rest of the system.

CHAPTER II

DIMENSIONS OF A STATE AGENCY PROGRAM EVALUATION UNIT: A STATE AGENCY PERSPECTIVE

Robert D. Struthers, Ph.D.

Introduction

In preparing this paper, it occurred to me that I was becoming one of the old-timers in agency program evaluation, and that there has been a lot of progress in this field in the last few years. Meetings for program evaluation people from different states are still not an everyday occurrence, but in the early 1970's they were even more rare than they are today.

When program evaluation meetings did occur, there was always an air of mystery. Evaluators would turn to each other and inquire, rather furtively, "What do you do in program evaluation in your agency?" And the answers would range over a variety of subjects. Some would say, "Well, we're planning to do a follow-up study;" others would describe their casework review process, and usually a few evaluators would attend the meeting because of their interest in evaluating clients in rehabilitation facilities. The meetings consisted primarily of discussing the issue, "What is program evaluation anyway?"

Now when evaluators meet, they converse at great length about standards, effectiveness, efficiency, models, outcomes, and indicators. Some even think that all evaluators mean the same things when they use these words!

Evaluators are obviously not trying to say that they don't have a long way to go, but they believe they have seen considerable progress, and attribute much of that progress to the activities of colleagues in the federal office, and especially to the promulgation of the Federal Program Evaluation Standards. Many evaluators take turns at pointing out the deficiencies of the Standards, but in my opinion, the Standards perform an absolutely essential function of illustrating the scope and content of a program evaluation system. Furthermore, the Standards provide a common target for discussion and thought, providing at least some structure on which to develop further state program evaluation activites.

The subject is "A Model Program Evaluation Unit From the State Perspective." This subject will be approached rather indirectly by discussing first the beliefs about what a program evaluation (PE) unit should do, and then how a unit might be structured to accomplish its mission. In the process, some examples will be given of the problems faced.

Functions of a State Program Evaluation Unit

In very simple terms, there are four capabilities that a state PE unit needs to have:

1. The unit needs to know what to do. It needs to have a strategy which determines what needs to be studied or reported.

2. The unit needs the technical capacity to be able to design and manage the projects or systems dictated by its strategy and to interpret their findings.

3. The unit needs to have the support capacity (person power) to carry out those studies, or other functions which are dictated by its systematic approach.

4. The unit needs to be able to convey the information it obtains within the organization and promote changes in operations.

Now let's discuss each of these subjects in more detail.

The Framework

The unit needs to know how to do program evaluation in a systematic manner. It needs a scheme or framework or model by which to determine what subjects are to be addressed and what information is to be generated as a result of program evaluation efforts. An appropriate framework is an issue which has been dear to my heart every since this program evaluator sat in a meeting and realized that any group of people within the agency could generate more questions in 30 minutes than could be answered in the next three years. Since that time, the most pressing question has been, "What is the most important thing that needs doing in program evaluation, and what kind of a scheme can be used which will assure that efforts are placed in the most fruitful areas?" This question can be considered paramount in planning, organizing, and staffing program evaluation units at the federal and state levels.

It seems that many in the field believe that this first step of deciding what needs to be done, and how and when to do it, is still a major stumbling block. It is because of this need that the publishing of the federal Standards, despite their shortcomings, was such an important task. It provided a start. However, it is realized that the Standards in and of themselves do not comprise a total evaluation system.

For example, the Standards very appropriately include a requirement that a determination be made about the retention of rehabilitation benefits. When the requirement to conduct follow-up studies was put into effect, a belief seemed to exist that conducting a follow-up study would automatically provide information for counselors, their supervisors, agency directors, RSA, and the Congress, which would result in program improvement at all levels. The recognition now exists that there is a

significant difference between generating a general description of the degree to which employment is retained, and placing in the hands of our field people conclusions about how they can serve their clients better with regard to retention of benefits. The same type of problem exists with much of the other data that is gathered. It is very different to gather data to make a statement about the types of phenomena which are occurring within a state agency than it is to gather data in such a way that will assist in the planning of agency programs, the supervising of counselors, or in facilitating counselor activities with clients.

Federal and State Agencies. Before going further, a few observations will be made about the relationship between federal and state rehabilitation agencies. As one approach, imagine a system in which the federal government plays a very minor role and the states develop their own versions of program evaluation units and systems.

A second approach is for the federal government to set up the desired system and expect the states to comply. We have seen the latter occur to some extent.

The development of strong state PE units without adequate federal coordination would result in a wide variety of methods being used with lack of quality control or comparability in findings. Such an approach promises little to improve the planning or management of the total VR system on a national basis. On the other hand, over-emphasis on the planning of VR evaluation systems at the federal level may result in simple reporting systems with little prospect that they will accurately reflect or have the capability to improve field operations within the states.

The approach which is being assumed requires that there be considerable technical strength and capacity for conducting activities at both the federal and state levels. The states probably do not have either the inclination or the ability to build units which are technically self-sufficient, and which will make major innovations in evaluation strategy. The task, therefore, is simply too large to expect such a result. Also, it is clearly inefficient to have people in 82 rehabilitation agencies (general, blind, and combined) all trying to solve the same problems, each with very limited resources. The reality seems to be that once a practice is adopted at the national level, work done in an individual state will be superseded. The federal level has a primary responsibility for utilizing the results of evaluation efforts to justify the funding of the total vocational rehabilitation (VR) system and attempt to see that it improves on a national basis. It also has the capacity to provide the needed impetus to make new technical developments and to coordinate knowledge among the states.

It is assumed that much of the initiative for developing strategies in program evaluation is going to remain at the federal level, and that considerable standardization of methods and reports is necessary. Nonetheless, state agencies are where evaluation activities have to be converted from theory into practice, and it is going to be necessary for states to have a very thorough knowledge and understanding of evaluation

strategies even though they may not be the primary developers of those strategies.

The State PE Unit. The primary roles of the state PE unit will be to see that (a) tested methods are used in its studies, (b) the results of trying new approaches are reported back to be considered in the planning of new federal initiatives, and (c) program evaluation findings are used in improving rehabilitation services. When referring to an evaluation system, the concern is about a single compatible system for both the federal and state levels.

What should a system be like? The purpose of program evaluation is to get information to those individuals who are making decisions and taking action in the VR program so that they can take the most effective and efficient actions. The unique quality of information program evaluation provides is that it relates to the achievement of the agency's objectives, and specifically to the effects our services have on the clients being served. This is a more restrictive charge for program evaluation activities than some would give, but it seems that it is the heart of what is being attempted.

Furthermore, there is a best way to accomplish these things, and the best way has not yet been found. A workable evaluation system can be developed which reports much more than is currently known about the effectiveness of the VR program and which is so understandable that all state agencies can use it. Once the right questions are asked, some indicators of program performance will be seen as far more revealing than others, and some schedules for reporting will be seen as far more productive than others. The system could be described in a paperback book that can be put on shelves. It could be called, "How to do Program Evaluation in a State Rehabilitation Agency for Fun and Profit, Including the Items to be Reported and When and How to Report Them." By using similar systems from state to state, we can build a body of knowledge which will ultimately improve rehabilitation, and along the way will dictate to a great extent the form and activities of our program evaluation units.

Working Toward an Improved PE Model. Before leaving this subject, let's linger a moment to suggest three subjects which need to be addressed in working toward an improved PE model: decision making within the rehabilitation program, describing client change, and technical capacity.

1. Decision making within the rehabilitation program. First, a very intensive study is needed of who makes the decisions within the rehabilitation program, what decisions they make, and what information they need to make those decisions. Proper recognition is needed of what decisions are made by the Congress of the United States, by the federal rehabilitation agency, by the state director, by field supervisors, and by counselors, and exactly what type of reports can be placed in their hands which will assist them in making primary decisions. While all of these agents may refer to certain basic program descriptors, it is

clear that the decisions are different at different levels. The Congress may be greatly concerned about how many people need the services of the agency overall, but that may be a secondary concern for persons who are delivering services to clients.

The issue of who can use our information is at the heart of planning for evaluation activities, along with the need to do a better jol of ascertaining the users of the information before studies are initiated. Evaluators in state agencies have a primary responsibility to understand clearly how organizations operate, and to convey to the people who are doing research on evaluation methods or attempting to devise measuring instruments, just exactly who can use what information and under what conditions.

2. Describing client change. The second issue is the need for an additional indicator to describe client change in addition to the criterion of rehabilitated versus non-rehabilitated. This ancient issue will not be gone over again other than to say that agencies still do not have exactly what they need to serve as an indicator of client change for the purpose of describing program effectiveness. This subject will also respond to careful analysis of what we need the measur for. Under the Rehabilitation Act of 1973, a "handicapped individual" is to be served if he or she can reasonably be expected to benefit from services in terms of employability. We are probably not going to get very far in describing the effectiveness of our services until we have an indicator which says something like: The general population has an employability level of 38, persons who describe themselves as disabled in surveys have a level of 24, persons who enter our employment enter with an average level of 20, we set goals at a level of 34, they leave with a level of 30, and are found at follow-up to be functioning at a level of 28.

It is difficult to believe that such a measure needs to be any more complicated than the Oklahoma Service Outcome Measure. That is not to say it would necessarily permit the weighting of case closures or assist in the supervision of counselors or assist counselors in developing plans for clients. The point is that we need more discriminating indicators of client change to better describe what is being accomplished and to serve as dependent variables to identify the correlates of client change, and to build up a body of knowledge on effective means of service delivery.

Analysis should be done on three programs: (a) one designed to lead to increased employability, (b) one to lead to improved functioning as a homemaker, and (c) one to lead to increased functioning in independent living. The goal for each client should be chosen on an individual basis. Once an

individual is in a program, we should be able to describe his or her status at the beginning and at the end of services relative to the chosen goal.

3. Technical capacity. In addition to a systematic and standardized approach to evaluation, the PE unit must have the technical ability to design projects and systems and to obtain, analyze, and interpret data. Technically, there are three kinds of functions which should be available in a program evaluation unit:

   a. The ability to provide reports required by the state-federal program evaluation system and the ability to contribute to improved methods and strategies for the system

   b. The ability to conduct studies of particular interest, over and above items which may be dictated by a standard program evaluation system

   c. The ability to provide technical consultation within the agency on a variety of subjects and to elicit such information as can be obtained from "informal" data

   The first two functions are rather clear, but the third may not be so clear. Experience has taught that the program evaluation or research person becomes the reigning technical expert of the agency. This person is presented with all the quantitative problems of the agency. How many rehabilitations should we expect counselors to produce? How large should a caseload be? How many people need services in this or that county?

The success or failure of the evaluation unit may rest with the ability to respond to these types of questions. A considerable part of evaluation is, and will continue to be, the ability to make guesses with partial data, which is somewhat better than having no data.

Obviously, if the program evaluation projects and systems have been carefully planned to provide those indicators which are of greatest utility, the findings may not require much interpretation. However, in the case of evaluations which are not done frequently, and even more commonly in the case where efforts are made to draw conclusions from data which were not reported in the most usable form, the PE unit needs considerable technical competence.

The author is very quick to state a feeling of inadequacy in attempting to apply appropriate statistical methods to data which are generated in the agency and to assure that the data are gathered in the most usable fashion from the beginning of the project.

A similar situation exists with regard to computer capability.

Generally, adequate capability exists within our agency (Michigan general) to handle the vast quantities of production data which are generated on a regular basis. However, when previously unfamiliar data are generated, our agency has difficulty getting timely programming, appropriate statistical analysis, and the printing out of the information in usable formats.

The obvious answer to these problems is to have strong technical people on the staff. The type of person we need should have:

1. Agency experience, preferable at the counselor and supervisory levels

2. A good background in research and evaluation methods, including the use of experimental and quasi-experimental designs

3. Knowledge of statistical methods, including simple parametric and non-parametric methods, plus multivariate methods

4. Competence in writing questionnaires and designing and conducting evaluation studies

5. Complete knowledge of sampling methods, including methods for survey sampling and internal auditing

6. Knowledge of the methods of economic analysis, including knowledge of labor market statistics, methods to compare the costs of products, projecting the costs of new projects, and utilizing cost-effectiveness and cost-benefit formulas

7. Knowledge of computer capabilities and the ability to utilize computers as an aid in projects

8. Knowledge of methods of management science, including linear optimization, goal programming, and simulation

9. A knowledge of modern management thinking, including methods of management by objectives and methods of training personnel and gaining utilization of research findings

10. The ability to work easily with persons at the counseling level and to function smoothly with the director of the agency, representatives of the Bureau of the Budget, legislators, and other state executives

11. The person should have extensive experience, but be relatively young, and be employable under the affirmative action plan

12. It is extremely helpful if the person can stop speeding bullets, leap over tall buildings, and swim raging rivers

The point made seems to be clear. More difficult technical problems are faced than state units are likely to be able to solve without technical support.

Support Personnel and Electronic Data Processing Capability

After the road map has been organized to guide the production of needed information and the technical ability to conduct studies put in place, an important factor is the capacity to gather necessary data. Reference is not about technical capability, but simply the number of people and the necessary reporting systems to gather data and get it available for analysis. Some information which is needed for managing the agency may be required on a monthly or quarterly basis. Other information may be meaningful only on an annual basis. Other types of information are needed only occasionally.

More people are needed if the desired information is hard to obtain. It is not feasible to do a survey of the total state population to determine the need for services every year. Several problems exist at this point with follow-up studies. With the difficulty in getting accurate information due to incomplete responses, the relatively small differences found from year to year, and the cost of getting the information, serious doubts exist whether it is worthwhile doing a yearly follow-up study.

Likewise, the required special studies of the needs of the severely handicapped, the need for rehabilitation services, and the need for services in general probably should be done on a periodic basis, perhaps once every three years. Scheduling projects on a rotating basis allows the agency to keep a relatively small staff busy on a regular basis. The total amount of person power needed may be greater in a larger agency or more populous state than in a smaller one.

On the subject of studies of needs, it appears that it may be necessary to provide some supplementary funds for the conduct of studies of the need for services. One impression is that most of the work in estimating needs currently done by state agencies is not really usable for meaningful planning activities. For example, is unlikely that we know how to describe client needs in terms of program options.

If the state has an ongoing reporting system and data can be gathered by simply recording a couple of new items, then there may be no need for additional program evaluation staff. However, there may be instances where it is far better to use program evaluation staff for a short period of time and do a single study on a sampling basis rather than to involve many more hours in the recording of data on every single case which goes through the system. One assumption is that everything possible would be done to minimize the time taken from service delivery to provide data for program evaluation.

Just as the need may exist for specialized technical ability to utilize computer capability, the need may exist for extra computer capacity to process data from program evaluation studies. An assumption is that most agencies have some computer capability to assist in reporting data, but often it is difficult to make additions to regular reporting procedures or to utilize computer capabilities for special projects.

Utilization of Findings

The fourth function of the program evaluation unit is in some respects the most difficult of all. That is the function of getting information utilized within the agency and influencing program change. There is a considerable amount of admiration, and also sympathy for our federal agency people on this subject. They have tried almost everything to get research utilization within the state agencies, and we feel they would attest that it is not easy.

In this area most of the progress which has been made in program evaluation can be attributed to federal regulations. The requirement that program evaluators should report to the director of the agency and the existence of reports on the Federal Program Evaluation Standards has forced the states to at least recognize the existence of program evaluation.

Nonetheless, it is the author's opinion that program evaluation is generally seen as an add-on activity to other more important functions. The program evaluation function is not seen as a high status role or position within the agency. In contrast, it may be seen as the bearer of bad tidings and as a threat to persons who feel they may be subject to evaluation.

No simple rules are known to change this status. For the most part, it depends primarily on the attitudes and activities of the persons doing program evaluation, and the value of the information the unit is capable of providing.

Optimism can continue in the belief that when we have a sufficient understanding of what evaluation has to offer and the agency can present it in a convincing way, evaluators will gain support for program evaluation activities. Some agencies are making progress in this regard.

One concern in this regard is that evaluators will find that some useful functions can be performed for the agency by the evaluation unit in summarizing quantitative information. But the evaluation unit may never develop the capability to the point where evaluators are, in truth, able to reflect to managers the consequences of their decisions in terms of what is happening to clients.

Implications for the Organization of a Program Evaluation Unit

Several implications will be provided of the remarks which have been made for the organization and operation of a state program evaluation unit.

Location of a Unit. First, the federal guidelines for staffing and organizing a program evaluation unit are very appropriate. It will be remembered that the guidelines emphasize that the program evaluation unit should have close access to the director of the agency and should include or work closely with the program planning unit. Certainly that approach is supported.

One of the frequent problems with evaluation is that persons not familiar with the potential and limitations of quantitative methods may fail to grasp the importance of certain findings or, conversely, may over-interpret very shaky findings. In addition, when problem statements go through several layers of management they sometimes look very little like the original problem. As program evaluation progresses in development, a much better job of "training" our administrators should be achieved in the use of quantitative information for decision making, as well as helping them learn how to use a technical support system.

At the same time, it is important to be familiar with the informal factors which go into decision making in our agencies and to be able to become more responsive to our agencies' needs. It is likely that most evaluators do not have as much access to decision making in their agency as would be desired in a model system.

Organization of the Unit. As to the organization of the unit, the author has spoken with several people in state agencies on this subject, and admits having difficulty translating the various job descriptions into common terms. Some evaluators have relatively elaborate organizational structures which may offer advantages which are not addressed here.

As indicated before, the tendency exists to define the functions of the program evaluation unit rather narrowly. Not included are the operation of the ongoing data reporting system of the agency or planning which is done to meet the requirements of the Federal Program and Financial Plan or support budget requests. It is assumed that these functions are present in every agency, with or without a program evaluation function. The program evaluation unit should provide information to support their activities, but is an addition to them. Functions which may be called research in many agencies are not included. As will be noted later, there is a tendency to classify many research projects as evaluation related.

There is not a great deal to say about the organizational relationship between the program evaluation unit and the ongoing data reporting system, the fiscal unit, the unit for casework review, or the unit for planning and budget preparation. In most states these functions are grouped in some type of a management support unit. The working relationship is the crucial element, and at various times the program evaluation unit needs to work with all segments of the agency, and the methods of operation are more important than the organizational chart.

It is obvious that the program evaluation unit needs to maintain the best possible relationship with other units within the agency. As little experimentation as possible is recommended which interferes with the delivery of services. New projects or potential items should be field tested as thoroughly as possible before requiring general participation. In general, the program evaluation unit should seek to reduce the reporting requirements of service deliverers rather than increase them and, if possible, should get along with sample information unless it is absolutely necessary to require reporting of each item on each case.

The program evaluation unit is not seen as being the manager of changes within the agency. The product of the program evaluation unit is information. Obviously, the higher the information can be presented within the agency, the more likely it is that action will result, but it is not a managing unit. The type of information which should be generated by the program evaluation unit is primarily that of new systems or special projects. If information is worth gathering on a systematic basis, then the program evaluation unit should aid in the initial interpretation and utilization of the information. However, once it is reported on an ongoing basis, then presumably managers will be given the information to use and the training to use it. Parenthetically, the program evaluation function cannot reach its ultimate goal of changing services if the agency does not have a strong field management system. The program evaluation unit probably needs to do training in how to use quantitative data in managing, but it cannot compensate for a management system which is inherently weak.

Staffing. When hiring people to staff this unit, start with a person who has familiarity with the agency and some ability to deal with quantitative data and some interest in it. In any agency of almost any size this one person will find more assignments than he or she is capable of performing. This person will need to spend considerable time in corresponding, communicating, and meeting minimum requirements to comply with federal reporting requirements.

The second person hired would be the most capable technical person available with the ability to utilize analytical methods and, hopefully, to manage the development of computer capability to support the technical projects. This person would have training and/or experience in counseling, psychology, or the social sciences. The problems are sufficiently complex to challenge persons with doctoral degrees, if we can afford them, but this would probably be a person with an M.A.

The next acquisition would be support people to work with our technical personnel in data gathering and analysis until it appears that the technical person is working at maximum efficiency. The next person sought would have a background in an area such as economics or management science who would be able to bolster our capacity to make cost estimates, determine cost effectiveness and project costs of future program alternatives.

Frankly, our imagination does not run beyond this point. Probably one should go no further, expanding rather slowly to see how much capacity is generated by the addition of each person. Obviously, larger states would require somewhat more data gathering and analysis capacity than smaller states. If reporting systems are developed for new programs it may be necessary to assign some personnel permanently to the utilization and interpretation of findings.

Concluding Comments and Recommendations

This has been a rather sketchy description of program evaluation and the possible functioning of a state program evaluation unit. It has been noted that activities at the federal level have provided considerable leadership to date, and will probably have to continue to provide much of the initiative in the future. At this point, it is more comfortable to talk about common state-federal problems than about state units themselves.

To express a couple of opinions, and make a few recommendations, first, the rather biased view is offered that all of the research and development projects being sponsored in RSA should be support for service delivery activities, and especially agencies in the state-federal VR program.

Second, the perspective of program evaluation should carry a great weight in the spending of research and development, and innovation and expansion funds. Without looking at the titles of the projects that are being done in the name of research, development, innovation, and expansion, the judgment can be made that most of them should be called program evaluation projects.

In almost every instance the intent of evaluation projects is to determine that if services are provided to a given target population in a given locality, or of a given type, it may be expected that a certain number of persons can be rehabilitated with a certain amount of rehabilitation gain at a given cost. These are evaluation questions and should be considered as such instead of having a feeble requirement that each project include an evaluation component. It should be expected that the product coming out of such projects will be a statement which can be used in the planning of future programs, and until it is known how to do projects so that such statements are generated, evaluation reports will continue to gather dust on shelves.

Now for some quick recommendations:

1. While preferring not to turn immediately to this subject, it is true that a greater capacity is needed for doing program evaluation in the states. Most states need a minimum of two individuals to do anything more than simply respond to reporting requirements and very short-term agency questions.

2. It appears that the federal staff also needs more resources in order to maximize the gains which should be made in coordinating the efforts of contractors, grantees, and the states. Success in program evaluation will depend on the ability to identify the most critical subjects to be addressed and to assure that usable products will result. This requires very intimate knowledge of the problems and close supervision of projects.

3. State agency program evaluation people need to be involved in more of the planning of federal strategies and projects. The participatory planning sessions for research and for program evaluation begun in the last two years have been an important step forward. However, the process needs to be made more frequent and more intense. Also, it has become commonplace for state agency people to be on advisory committees for projects being done by contractors and by grantees. This involvement should be continued and increased. It can be a learning process for all concerned as well as improving the products of the projects.

4. More training meetings are needed so that information can be shared among program evaluators, and to get a sense of commonality of purpose and content in what we are doing.

5. More purely technical training of program evaluators is needed. Experience has shown it is difficult to obtain training directly relevant to rehabilitation evaluation projects.

6. More grant funds should be provided directly to state agencies to conduct projects which promise to add to program evaluation knowledge and which require more resources than the state could normally generate. Joint grants to the states and competent technical people from universities and technical centers would be the most desirable.

7. Consideration should be given to a type of technical support which could be provided by roving technical consultants from our research institutes and research and training centers who would advise state program evaluation personnel about the design of new projects, pass on new ideas from other states, provide analyses of data by computer, and aid in interpreting that information. People engaged in such a function would soon be more knowledgeable about what is going on in our field than any of us are now, and would provide the type of know-how which would result in significant progress.

There are many facets of this issue which cannot be addressed. Hopefully, one will detect a great deal of enthusiasm and hope for program evaluation in rehabilitation. Potentially, program evaluators are in a key position to have great impact on the rehabilitation movement. At the same time, one must confess to being not at all sure that it is possible by quantitative program analysis to demonstrably improve rehabilitation services, especially in such a way that they can provide guidance for counselors with clients and for their immediate supervisors.

Nonetheless, as said at the beginning, we can point to significant progress. Program evaluators have some exciting challenges and are hopeful that in this rather obscure enterprise in one of the smaller governmental endeavors the day will come when it can be said that program evaluators have performed a service to our citizens in seeing that their money was well spent and, even more, that the services of our society to the handicapped people we seek to benefit have improved.

CHAPTER III

PROGRAM EVALUATION RESEARCH AND PRACTICE: UNIVERSITY RESOURCES

Ralph M. Crystal, Ph.D.

The Rehabilitation Research Institute at The University of Michigan (UM-RRI) is directed to program evaluation and to assisting states in conducting evaluations of their management practices and service delivery systems. The research objectives are focused on the following issues:

1. To develop alternative conceptual models for program evaluation processes
2. To develop and test instruments, techniques, and methodologies in vocational rehabilitation program evaluation

Core Area of Research

The University of Michigan requested funding from the Rehabilitation Services Administration (RSA) for a five-year long-term project to establish a rehabilitation research institute beginning October 1, 1976. The mission of the UM-RRI is to implement programmatic research in rehabilitation program evaluation and to identify, develop, demonstrate, and encourage the utilization of evaluation techniques and strategies for the measurement of rehabilitation program effectiveness as a basis for program improvement.

In order to accomplish the objectives of the Institute, a two-level approach to rehabilitation program evaluation has been implemented. This approach relates to (a) research development and (b) research utilization, both of which are goal and end-product oriented.

Research development pertains to projects and activities which are primarily developmental in nature and which may require a research effort over an intermediate to long-term period. Research utilization refers to projects and activities which may be short-term or intermediate in time, but the efforts are directed toward disseminating, "packaging," and modifying materials in such a way that users in the field can apply a technique or approach, and be provided with training or technical assistance. Although Institute staff members have been assigned specific responsibilities for these areas, it should be noted that these two areas are viewed as interdependent and interrelated.

It is the philosophy of the UM-RRI to continually survey the rehabilitation program evaluation field as a needs assessment approach to the

identification of areas requiring programmatic research. The ultimate objectives of this procedure are to (a) develop a comprehensive conceptual approach to rehabilitation program evaluation, (b) assess and analyze problems and concerns in rehabilitation evaluation, (c) compile, test, and validate new and existing program evaluation instruments, (d) identify completed and ongoing rehabilitation evaluation studies, and (e) provide technical assistance and conduct utilization conferences for agency personnel.

Through the research (product) development-research utilization network, the UM-RRI maintains close and continuing contact with rehabilitation program administrators, evaluators, and planners. These linkages provide further information for needed research development and research utilization. It is the view of the UM-RRI that program information must be utilized to be meaningful.

Since the Institute began in October 1976, a major emphasis has been placed on surveying the rehabilitation program evaluation field to determine the current state-of-the-art. This has resulted in a number of monographs completed or to be completed by December 1981. Among these documents are the following:

1. Crystal, R. M. A survey of the current status and program evaluation needs in the state-federal rehabilitation program. Michigan Studies in Rehabilitation, The University of Michigan Rehabilitation Research Institute, Ann Arbor, 1978, 1 (2).

2. Crystal, R. M. and Lee, C. C. Program evaluation terms: A glossary. Michigan Studies in Rehabilitation. The University of Michigan Rehabilitation Research Institute, Ann Arbor, 1979, Utilization Series: 3.

3. Harrison, D. K., Lee, C. C., and Zawada, M. A. Competency evaluation in rehabilitation (CEIR): Rehabilitation counseling competencies. Michigan Studies in Rehabilitation. The University of Michigan Rehabilitation Research Institute, Ann Arbor, 1978, 1 (3)

4. Harrison, D. K. and Miller, J. V. Rehabilitation program evaluation: Problems, objectives and projects. Report of an RSA sponsored participatory planning conference, Ann Arbor, 1977.

5. Miller, J. V., Lee, C. C., Wargel, J., and Won, H. Program evaluation approaches for state rehabilitation agencies: Current status and future directions. Michigan Studies in Rehabilitation. The University of Michigan Rehabilitation Research Institute, Ann Arbor, 1977, 1 (1).

6. Miller, J. V., and Wargel, J. F. Handbook of program evaluation studies. Michigan Studies in Rehabilitation. The University of Michigan Rehabilitation Research Institute, Ann Arbor, 1978, Utilization Series: 1.

7. Miller, J. V. and Wargel, J. F. Identifying and recruiting priority clients: A guide to planning and evaluation. Michigan Studies in Rehabilitation. The University of Michigan Rehabilitation Research Institute, Ann Arbor, 1979, Utilization Series: 2.

8. Miller, J. V. and Wargel, J. F. State agency facility planning: A guide to planning and implementation. Michigan Studies in Rehabilitation. The University of Michigan Rehabilitation Research Institute, Ann Arbor, 1979, Utilization Series: 4.

Relationship Between the UM-RRI and a State Evaluation Unit

Simply stated, the basic purpose of program evaluation is to provide information to management and other constituents of the rehabilitation program about the functioning of the program. Evaluation addresses issues related to the impact of the program on the client, the agency, the counselor, and the community. On the basis of these data, decisions can be made for program planning and development, innovation and expansion, organization development, the implementation of innovative programs, and program and policy analysis. Program evaluation data are also used for monitoring the program and justification of the program to sponsors and other external funding sources.

The purpose and use of evaluation information depends, in part, on who is using the information. Different constituents of the rehabilitation program (administrators, program evaluators, program planners, counselors, clients, program sponsors) have different information needs and uses for evaluation data. However, a common objective of evaluation is the effective utilization of manpower, service, and monetary resources in helping clients achieve their rehabilitation objectives.

The UM-RRI can be viewed as a resource to model evaluation units. The Institute can help these units develop and utilize a conceptual and information base to increase a state's capacity to utilize rehabilitation programs using evaluation data. Three areas of input are presented:

1. To provide technical assistance and support for the utilization of existing Rehabilitation Services Administration (RSA) evaluation capacity

2. To help states increase their evaluation capacity through the development of a comprehensive evaluation strategy. This would provide continuous monitoring of program activities as well as a system for measuring change in client functioning as a result of the provision of rehabilitation services

3. To help states develop and build capacity to effectively identify key program questions and develop evaluation strategies to answer those issues

First Area of Input

In recent years, RSA has increased rehabilitation evaluation capacity through various activities. Included among these are the continued development of the vocational rehabilitation program evaluation Standards, the linkage of data records between the Rehabilitation Services Administration and the Social Security Administration (SSA) the development of a model facilities information system, the project to revise the current RSA management information system, and the model state evaluation units.

The UM-RRI has the capacity for providing evaluation units with technical assistance to utilize the information generated by RSA evaluation projects. This support can facilitate a state's ability to understand and evaluate its rehabilitation program. The Institute can help states determine (a) what information is being provided by these data sources, (b) to what program components the information is related (including the limitations of the data), and (c) how the data can most effectively be implemented as a means of capacity building.

As an example, an evaluation methodology is ng developed by the UM-RRI to help states identify explanatory factor om the data elements of the federal evaluation Standards. These explanatory elements are being viewed as antecedent and consequent variables. If a state is out of compliance on a particular Standard data element, it will be possible to indicate reasons for that situation and factors which might be affected in the future.

Thus, a resource such as the UM-RRI can be viewed as having the ability to help states translate evaluation information into a context which is understandable, meaningful, and relevant for constituents of the program. This relates to the next step in the evaluation process - utilization of information.

The planning, implementation, and evaluation of change is a critical element of the program evaluation process. By not being intimately involved in the day-to-day operation of the rehabilitation program, university resources can provide a perspective by which it is possible to help states assess the quality, use, and impact of evaluation data. This type of consultative support can be provided through technical assistance projects and research utilization conferences. The key issue for input from university resources is how to most effectively utilize the evaluation information being generated through current and new sources.

Second Area of Input

There is a need to help states build their evaluation capacity in terms of the identification of change in client functioning as a result of the provision of rehabilitation services. Based on the research con-

ducted to date by the UM-RRI, a need to develop and implement a comprehensive rehabilitation evaluation system was identified. Existing and new information (as discussed under "First Area of Input") can be incorporated into this evaluation approach.

The evaluation design is seen as broad enough to encompass present and potential evaluation questions and meet future changes in program directions. With independent living legislation in rehabilitation, it is increasingly important to indicate the competencies needed by clients to function effectively in vocational and non-vocational domains. The following developmental activities, proposed by Miller, Lee, Wargel, and Won (1977) and expanded by Harrison (1978), provide a Competency-Based Client-Outcome Evaluation strategy as a systematic response to the program planning and evaluation efforts needed in the rehabilitation field:

1. Document the Program Development/Program Evaluation Process - A clear description is needed of the parallel and dependent efforts between rehabilitation program development and program evaluation. This statement should include a description of the steps involved in the planning and delivery of rehabilitation programs; e.g., needs assessment, goal setting, program design, program delivery, outcome evaluation, program judgments, and program changes. It should also include a description of the key points throughout this planning process where program evaluation information is needed to improve decision making.

2. Develop a Taxonomy of Rehabilitation Outcomes - A taxonomy of rehabilitation outcomes is needed. This taxonomy should describe (a) broad domains of rehabilitation outcomes (e.g., psycho-social functioning, physical functioning, economic independence, and vocational functioning and potential); (b) specific terminal (end of rehabilitation process) outcomes related to each of the broad domains; and (c) intermediary (during program) outcomes which could be used to monitor client progress throughout the rehabilitation process. This taxonomy would help define the nature of the rehabilitation process. At present, evaluation is focusing primarily on terminal outcomes. This means that little pre-assessment and during-program data are available. Such early monitoring is essential for the adjustment of services and for identifying real changes resulting from the rehabilitation process.

3. Develop a Catalog of Measures Related to Outcomes - Once a taxonomy of outcomes is developed, existing measures should be linked to specific outcomes and new measures should be developed where necessary. This, of course, is a massive effort requiring the cooperation of state agencies and developmental groups. Use of the taxonomy of outcomes could, however, guide efforts in this area. It would respond to the existing needs to specify uses of existing measures and to guide the development of new measures.

4. Relate Outcomes to Rehabilitation Services - As the taxonomy of outcomes was developed, it could be related to existing areas of rehabilitation services; e.g., outreach, assessment, counseling, restoration, or training. This would help provide guidelines for determining what outcomes might be expected to occur during particular stages in the rehabilitation process. At present, it is not clear what criteria are used to determine changes in client status. Perhaps it is possible to base changes in status on the achievement of specific intermediary outcomes. This would help counselors become more systematic in the delivery of services.

5. Description of Variables in the Vocational Rehabilitation System - There is a need to identify and measure the multiple variables impacting on rehabilitation outcomes. Using a systems model, it would be possible to describe the typical state rehabilitation agency and its socio-economic content, and specify the range of variables which might impact on rehabilitation outcomes. Once the variables have been identified, a list of indicators should be developed which could be utilized in evaluation efforts.

6. Train Program Evaluators to Use the Evaluation System - The developmental activities described thus far are components of an evaluation system. They include: (a) a program development/ program evaluation process guide; (b) a taxonomy of rehabilitation outcomes; (c) a catalog of measures related to outcomes; (d) a guide to relate rehabilitation outcomes to rehabilitation services; and (e) a catalog of indicators for variables impacting on rehabilitation outcomes. These products would provide a system which could be used within state agencies to design ongoing evaluation efforts which support the program development process. This system would be flexible and would provide the process and the content for evaluation design. Effective use of the evaluation system would depend on training state agency program evaluators to use the system to design evaluation procedures appropriate to the needs of their own agency (Miller, et al., pp. 87-90).

The developmental activities for this effort can be viewed as complementary and interrelated with current RSA evaluation projects and other activities. It is very likely that much of the data for the evaluation system will be derived from the revised Standards for evaluation, the facilities' reporting system, and the revised RSA management information system. The proposed evaluation system would provide states and RSA with additional evaluation capacity which is sensitive to client functions without deviating from current RSA evaluation projects and priorities. The proposed evaluation system is parallel and interdependent with other RSA evaluation activities.

Third Area of Input

A critical evaluation concern is to determine key evaluation questions and issues. The ability to articulate rehabilitation program problems, needs, and concerns and translate these into questions which can be developed into evaluation studies is an important aspect in the identification of program evaluation issues. Needs assessment techniques and RSA's use of participatory planning conferences are procedures which have been used in the past to identify evaluation issues.

State agency person need to have skills which enable them to identify program concerns and translate these into evaluation studies. Knowing the questions to ask is the first step in the evaluation process. If the "wrong" questions are posed, the results of the evaluation will not be satisfactory and, therefore, will be ignored. Effective utilization of evaluation information is dependent on the articulation of evaluation questions and the development of studies which answer those questions.

By determining critical evaluation questions, it will be possible for state agencies to make recommendations to RSA regarding the instruments and the data which need to be collected to effectively answer present and future evaluation questions. With such a perspective, it will be possible to build into the revised RSA management information systems items which will provide data for answering rehabilitation evaluation issues.

In evaluation, as in other types of research, it is important to first have the question and then decide what data are needed to answer the question, rather than having the data and deciding what questions can be answered or attempting to answer questions with data that do not completely fit the questions. Working together, university resources and state agencies can help determine what information is needed to answer the rehabilitation questions (including the federal evaluation Standards) and make recommendations to RSA that such reporting, where feasible, be made a part of the new information system.

Other Areas of Input

Another area in which university resources and state evaluation units can work cooperatively is in the design, development, validation, and critique of new measures. There is a need to develop measures which accurately reflect what is occurring in the vocational rehabilitation program. New measures and instruments, such as follow-up questionnaires, case reviews, indicators of client status and functions, and other instruments can be developed and validated through field testing.

A final area of cooperation is in field testing and demonstration of new evaluation approaches. State agency evaluation units can serve as demonstration sites for innovative evaluation efforts in such areas

as outreach, restoration and training, and job placement. With technical assistance from university resources, the information gained will greatly increase the evaluation capacity of state agencies and RSA.

New dimensions in evaluation can be explored in the model units without detracting from required evaluation activities. Included in this can be studies and projects which examine the most appropriate methods to provide information to management and personnel needed to function effectively as an evaluation unit. New methodology and statistical procedures can be developed which will help determine the effectiveness of the rehabilitation program.

Summary

It was the purpose of this paper to indicate several areas in which a university-based rehabilitation research institute can serve as a resource to state agency evaluation units. An intent of the paper was to stimulate thinking in terms of possible cooperative efforts between state agencies and university facilities. Without doubt, all possibilities have not been exhausted.

References

Harrison, D. K., Lee, C. C., and Zawada, M. A. Competency evaluation in rehabilitation (CEIR): Rehabilitation counseling competencies. Michigan Studies in Rehabilitation, Series 1, Monograph III, 1978.

Miller, J. V., Lee, C. C., Wargel, J., and Won, H. Program evaluation approaches for state rehabilitation agencies: Current status and future directions. Michigan Studies in Rehabilitation, Series 1, Monograph 1, 1977.

PART TWO

PROGRAM EVALUATION STANDARDS AND DATA RESOURCES FOR REHABILITATION

INTRODUCTION

At the federal level a comprehensive evaluation system for the state-federal program is being developed by RSA. The initial response to the Rehabilitation Act of 1973, mandating state agencies to conduct evaluation studies, led to the implementation of what became known as the nine General Standards. Although still in effect, the limitations of these initial Standards were recognized. Supported by RSA's Division of Evaluation, a number of projects have been undertaken to further develop and refine RSA's evaluation capacity.

Under the direction of Ms. Susan Shea, Berkeley Planning Associates has developed revised evaluation Standards for the state-federal rehabilitation program. Program Standards were developed for the formula grant program and project Standards were developed for the eighteen legislated project authorities. The new Standards were designed using feedback from the current Standards, but more importantly to the development process, Berkeley Planning Associates reviewed the rehabilitation process and its outcomes more comprehensively to develop a conceptual base for the Standards.

Mr. Gary Prezak of Walker and Associates has been involved in the development of the Model Federal-State Facilities Reporting System for Medical and Vocational Facilities. The nationwide adoption of the system will constitute a pioneering effort at a systematic collection of data relating to rehabilitation facility operations in such previously unexplored areas as cost-efficiency of a facility, client distribution, and client outcomes under various programs. With the implementation of the facility reporting system, RSA will be able to better control how money is spent on rehabilitation facilities, which accounts for approximately one-third of case service expenditures.

Currently in private practice, Mr. Stanley Portny was instrumental in the development and implementation of the current evaluation Standards. He speaks on the topic of utilization of evaluation data with much background knowledge and experience. A concern of RSA has always been how information can be employed by program managers at the state, regional, and national levels after it becomes available. Mr. Portny addresses the topic of matching evaluation questions with evaluation data.

CHAPTER IV

STANDARDS FOR THE EVALUATION OF REHABILITATION PROGRAMS

Susan Shea

Introduction

Standards have been and will continue to be a major focus of attention at any conference addressing program evaluation needs in vocational rehabilitation. Unfortunately, one of the only favorable remarks to be made about the existing Standards has been that they formed a common target against which all program evaluators could unite. My current role seems to be to provide a new, improved target in the form of new, improved Standards. One of the main problems is that the present Standards are seen as a perfunctory and time-consuming reporting exercise. They have not really been a part of good program evaluation and have not really had any impact on the program.

This is not to say that there weren't attempts to accomplish some sort of integration of the Standards with program evaluation and with improving the program. Work on performance level scores and other techniques for integrating the Standards with other data in order to learn something from them has been attempted. But, in general, we have failed to go far enough in this direction and the Standards have remained a once-a-year data-copying exercise for submittal to the Rehabilitation Services Administration (RSA).

The failure to integrate the Standards data with other information concerned Berkeley Planning Associates (BPA) greatly as work was initiated on revising the Standards. So BPA interpreted its mission to be not just a revision of the Standards, but also the development of a whole system for using them and making them the focus of program evaluation activities in vocational rehabilitation (VR).

This paper will focus on the developments of BPA in terms of revised Standards, data elements, and potential approaches to performance levels. It should also acquaint the reader with the various issues confronted in refining the Standards and the need to ensure that this time around the mistakes of the past are not repeated.

Development of the Proposed Standards

Much criticism had been leveled at the original nine Standards from the beginning and much work was done addressing how they might be improved. Many program evaluators participated in what became the "New Orleans Report," which documented the results of the regional studies

of each of the nine Standards.

When work began on revising the current nine Standards, BPA became familiar with all available material, but it was also critical to obtain first-hand information. BPA took a step that had not been taken in developing the first set of Standards. Direct visits were made to several states to determine the "state of the art." All of the states could not be visited but as many as possible were, covering different regions, states with sophisticated data systems, and states without sophisticated data systems to find out first hand what state agency personnel thought of the Standards. A choice was made to speak to evaluators, administrative people (the director or his or her representative), and field service people in each of the agencies.

Somewhat surprisingly, very little response was obtained from the people in administrative or field service roles. Most of them knew of something called Standards. The main concern of administrative and field service personnel, based on limited familiarity, was that somehow the Standards were going to be used as sanctions. The sanction point of view was often the limit of their acquaintance with the Standards. From the evaluators, however, quite a lot was learned. Sharing the "laundry list" of problems and desires for an ideal Standards system provides a good context for understanding the steps taken in developing a refined set of Standards.

In general, the comments reflected a sense that the nine Standards, in concept, went in the right direction. But the operationalization and execution left much to be desired.

The Data Elements

The focus of the greatest criticism was that the data elements do not measure the intent of the Standards. A myriad of examples were offered. For instance, everyone agreed that undue delay should not exist in serving clients (Standard 3), and that measuring time in process is not a good indicator of undue delay. Who is to say that 22 months is the maximum acceptable time for serving every client, regardless of disability, case severity, or services needed? Some clients might legitimately be in a case service status for well over 22 months; others might be experiencing tremendous delay if in a case service status for only eight months.

Performance Levels

No satisfaction was expressed with the central tendency performance level approach, based on the mean with plus or minus one standard deviation level established as the cut-off point for acceptable performance. The major criticism was one that showed real concern with the policy

implications of the Standards; i.e., that this approach leads to mediocrity. Why should currently less than average performance be considered up to standard? On the other hand, when in a given area all states may be doing well, the use of this central tendency approach would then automatically declare several states out of acceptable limits. This would be of particular concern should Standards eventually be used for sanctions.

A concern which was expressed in many ways but particularly in terms of the performance levels was that the current Standards and performance levels took no account of differences between states. With the exception of separate treatment of blind agencies and general agencies, no consideration was given to differences in client mix. Similarly, states with different kinds of economies, different amounts of funding, or different kinds of service resources were all treated the same on almost every Standard.

Emphasis on Quantity

A third problem with the current Standards was the perception that they emphasized quantity to the exclusion of concern for quality. Many administrative and field service staff in the states wished to see some attention to service quality, rather than solely attending to production.

Inappropriate Topics for Standards

The preponderance of observers felt that most of the nine Standards were conceptually on target. Even if the data elements were unsatisfactory, at least the Standards themselves addressed the correct issues. Specific Standards were critized on this basis, particularly the Standard on post-employment services. A question existed in people's minds as to the rationale for singling out post-employment service from among all VR services. Similar doubt was expressed as to the validity of setting a standard for reasons of case closure.

An Exercise in Duplicative Reporting

A fifth criticism, which was that the Standards represented nothing more than a reporting exercise, obligating states to report in a different and incompatible format the same series of numbers required for various other reports. Not only was duplication a problem, but definitions were just a little bit different, the populations addressed were just a little bit different, and the results appeared to be a time-consuming and wasted effort.

Data Collection Burden

Another criticism was that the Standards imposed a burden of data collection, primarily due to the follow-up survey. Follow-up was con-

sidered time-consuming and expensive for the agency. The follow-up survey was the one aspect of the Standards with which field service staff were familiar, and it was not a particularly positive experience.

Field service staff felt the tracking of former clients' addresses required excessive time, with very low yield, because of the difficulties of out-of-date information and resultant low response rates. Further criticism was that the data being collected on client satisfaction produced the same results year after year with little utility to the agency.

Lack of Feedback

State agency staff consistently complained of having received little or no feedback from RSA on the Standards. While it is admittedly a complicated process to obtain, digest, and report back on data from 50 states (32 agencies in all) these agencies had exerted considerable effort to report on their status, apparently to no avail. They had turned in their reports, at times with some fears of being out of compliance, yet they heard nothing in response.

Standard Objectives

The foregoing were the problems heard. What did people want in the way of Standards? They wanted Standards which would really direct the program to its goals and a system which would measure whether the program was meeting its ultimate goal of serving the disabled client. They wanted issues of quality as well as quantity to be addressed. They wanted meaningful data but they also wanted a minimum burden to the states in collecting new data. They wanted to cover the important elements of the program but preserve the flexibility of state agencies. This particular desire reflected an awareness that it is important to be able to measure whether VR is achieving its goals, but a wariness against any attempt to standardize all of the processes and inputs to VR. States wish to have varying service approaches or administrative structures, as long as they can accomplish the required end result and meet the ultimate goals of the VR program; there was a reluctance to impose standardization. BPA knows of no "one best way," no single technology or state of the art that "works" for VR. Much of the work reported in the "New Orleans Report" and other research made this same point. Standards which would require agencies to do things a certain way, not previously established as the most effective, might actually be detrimental.

Standards Within the State System

States wanted a system which did not just measure how they were doing but could tell them why. That seems like a big order but it is something BPA has spent a lot of effort trying to achieve. If one state achieves many successfully rehabilitated (competitive employed clients) and another

state has very few, or if one state has high cost effectiveness and another state does not, the reasons why need to be known. If within a state one district does well and another district does not do well, why is that? The point, well taken, was that having Standards will be of no value if the status reports are submitted and then ignored. The desire was for a system which would enable states to learn from one another and enable the program to improve operations and ultimately program achievements.

States wanted a system which recognized individual differences in terms of numbers of severe clients, disability mixes, client socioeconomic backgrounds, resources, and funding situations. All of these characteristics require differential consideration of performance assessment.

Finally, states wanted to have useful feedback from RSA central and regional offices. States did not wish to continue sending in reports and hearing nothing in return.

This is the scene into which BPA entered when work started toward refinement of the Standards. There were many directions which could have been taken: maintaining the existing nine Standards, trying to make minimal changes in the data elements; moderately revising the nine Standards, but avoiding radical change; starting from scratch and thinking through, from the very basic concepts, what the Standards should be. BPA chose the latter course. This is not to say that BPA was committed to throwing out everything about the current nine Standards, but a decision was made to not be tied to them. The criticisms and concerns about the Standards were sufficient so that it seemed to make sense to start from the ground up.

Building a Standards System

The first consideration was to identify aspects of VR which should be addressed by the Standards. BPA found immediately that the items and issues people felt needed to be covered were different in nature and scope, but that these differences were not being recognized. Some candidates for proposed Standards related to goals or outcomes of vocational rehabilitation: clients should be successfully rehabilitated and become employable; clients should be placed into competitive employment whenever possible; clients should retain the benefits of services and maintain employment and reasonable salaries over time.

Other candidates for Standards concerned processes and procedures for achieving these outcomes: procedures for providing services and for protecting client interest, the Individualized Written Rehabilitation Program (IWRP), the time clients spend in the service process, the size of counselor caseloads and approaches to caseload management. Finally, some candidates for Standards included inputs and factors which might be

taken into account in understanding program performance in achieving such goals as client mix (age, amount of education, disability), agency budget, organizational structure, and staffing patterns.

The current nine Standards, and a lot of the work toward revision of the Standards treated all of these things similarly. Standards were set or proposed for what an agency's client mix should be, how much should be spent on administration, what caseload size should be, and a Standard for client retention of benefits. BPA, working with RSA and the Council of State Administrators of Vocational Rehabilitation (CSAVR) Advisory Committee, came to the conclusion that it was inappropriate to treat these disparate concerns similarly. They all have a role in evaluation and some have a role in the Standards, but they were not all elements for which we should be trying to achieve consistency across states. So BPA developed a system which would distinguish each and put each in its appropriate role for program evaluation.

Classifications for Performance, Procedure, and Support

The system contains three dimensions: (a) Performance Standards, (b) Procedural Standards, and (c) Supportive Evaluation Elements.

Performance Standards measure the achievement of a desired outcome or mission of VR. For example, that there should be competitive employment closures to the maximum extent possible. It is recognized that legitimate differences exist among states in the proportion of clients who are found to be eligible, so a Standard should not be set which says that a state shall find 72%, or any standard proportion of their applicants eligible.

Procedural Standards address important elements of the process, but ones for which a standardized expected performance level across states is inappropriate. Data elements for Procedural Standards obviously would be of a different nature. Basically, they would involve use of what most states already have in place: some form of case review procedure on a sample basis.

The last portion of the system is Supportive Evaluation Elements. These elements of the VR process for which data are collected and analyzed, not necessarily on a routine basis, are used to help explain differences in performance on the Standards. They include cost data. One state may do better in one area than another because they have more money or spend their money in a different way, perhaps spending more on case service than on administration, or vice versa. Client mix is another element. One state may do better than another because of the type of clients it takes in. A third example is the type of services offered. One state may have identified a very effective service mix for

certain kinds lients a d do better than others for that reason.

These are examples of information that might help explain why some states do better than others, or why some states did better this year than last year. Appropriate analyses would be done as necessary. The analyses of interest would change from year to year depending on how states performed and the particular issues arising. While the distinctions in concepts may be a little bit abstract, they are necessary in order to achieve a workable system.

The Standards

The proposed system includes eight Performance Standards, four Procedural Standards, and a number of Supportive Evaluation Elements. These are as follows.

Performance Standards

1. VR shall serve the maximum proportion of the potentially eligible target population, subject to the level of federal program funding and priorities among clients.

   This Standard addresses the dimension of program coverage: How much of the target population is being served? Are there unmet needs for service?

2. The VR program shall use resources in a cost-effective manner and show a positive return to society of investment in vocational rehabilitation of disabled clients.

   This Standard addresses the relative effectiveness of agencies in achieving goals and generating benefits in terms of their levels of resource investment.

3. VR shall maximize the number and proportion of clients accepted for services who are successfully rehabilitated, subject to the meeting of other standards.

   Traditionally, success in VR has been measured by the number of "26 closures" obtained. State agencies need to know how many individuals are successfully served and should be encouraged to rehabilitate as ma persons in need as possible.

   Literature on the "26 closure" measure points to the need for more refined outcome measures. As a measure of outcome, counting "26's" ignores qualitative types of gains and does not differentiate clients on variables of difficulty or the level of effort and efficiency with which clients are rehabilitated. Assessments of programs and counselors, based solely on production measures such as this, have met with much

valid criticism from counselors and administrators alike. Relying simply on the number of 26's is only a partial approach for Performance Standards on client outcome and needs to be supplemented by other proposed outcome Standards.

4. Rehabilitated clients shall evidence increased economic independence.

This Standard addresses questions which are central to the meaning of gainful employment. Do job earnings allow a client to be free from other forms of financial support, such as welfare? Do individual rehabilitants achieve higher earnings after having received VR services than they would have without them?

For competitively closed cases, a gain in wages should represent an increase in economic independence. For those in homemaker, unpaid family worker, or sheltered workshop closures, the concept of reduced dependency represents a qualitative dimension parallel to the wage gain measure for competitive outcomes. Do such closures show an increase in economic independence by having improved ability to perform personal care without assistance, to perform household duties, or to achieve a reduction in dependency on public funds?

5. There shall be maximum placement of rehabilitated clients into competitive employment. Non-competitive closures shall be in accordance with the IWRP goal and shall represent an improvement in gainful activity for the client.

Competitive employment may not be the appropriate placement for all clients, and this may be particularly true given new emphasis on serving severely disabled clients. VR regulations require that any placement of a successfully closed client be into a "gainful activity" and that it be "consistent with his/her capacities and abilities," whether in competitive, sheltered, or non-competitive employment.

There is much speculation in the field about abuse of "homemaker" and "unpaid family worker" placements, and about using these categories to assure a "success" rather than because the placement is appropriate. While maximizing the proportion of competitive closures is important to the VR mission, it will not ensure that sometimes legitimate non-competitive placements are suitable for the client. The second part of this Standard addresses the legitimacy of non-competitive closures.

6. Vocational gains shall be attributable to VR services.

The intent of this proposed Standard is to provide a measure of VR's effectiveness in its attempt to rehabilitate clients; that is, to measure the extent of a causal relationship between the VR services provided and the outcomes on vocational gains achieved. Studies have

indicated that both VR clients and non-clients will show, over the same time period, increased earnings, increased levels of skills, and other vocational gains. This Standard seeks to measure how much of a client's vocational gains can be reasonably attributed to the provision of VR services.

7. Rehabilitated clients shall retain the benefits of VR services.

Retention of benefits gained through VR services is most important to the rehabilitated client. It can also be used as a measure of overall program effectiveness. A high drop-out rate or high recidivism rate following successful closure implies program failure and points to incongruence of program goals vis-a-vis individual client goals.

8. Clients shall be satisfied with the VR program, and rehabilitated clients shall appraise VR services as useful in achieving and maintaining their vocational objectives.

This Standard responds to consumer advocacy interests. There are problems, however, with using such satisfaction measures as an evaluation measure of program effectiveness. Nearly all clients express satisfaction, even those clients not successfully rehabilitated. Satisfaction measures need to be buttressed by client assessments of the utility of VR services in terms of actually having contributed to their getting a job and functioning in it. Utility assessment offers a valuable entree for probing areas needing program improvements and for ensuring consumer involvement in improving the responsiveness of VR services to client needs.

Procedural Standards

1. Information collected on clients by the R-300 and all data reporting systems used by RSA shall be valid, reliable, accurate, and complete.

The VR service delivery system needs an objective data base from which to measure performance. A valid and reliable R-300 system including appropriate data items would be particularly beneficial to the state relative to the Standards, since states collect the data, report on the R-300 to RSA, and have RSA do the actual analysis for the Standards report. Yet inconsistencies and errors in reporting currently exist among and within VR program data systems. Confusion and misunderstanding of definitions exist and need to be minimized. This procedural Standard would ensure that state agencies maintain acceptable levels of validity and reliability in reporting of R-300 and other data.

Indicators for acceptable tolerances of validity, etc., should be

set for this Standard. However, responsibility will rest with the states to carry out the procedures necessary to detect error rates in reporting. There is legitimate variability among states in data systems used to derive the R-300 and other reported data items.

2. Eligibility decisions shall be based on accurate and sufficient diagnostic information, and VR shall continually review and evaluate eligibility decisions to ensure that decisions are being made in accordance with laws and regulations.

The determination of an applicant's qualifications for eligibility is a critical point in the VR process for both the client and the agency. This Standard seeks to protect client interests by requiring state agencies to install procedures for monitoring eligibility decisions in a sample of cases and ensuring that the decisions are appropriate.

3. VR shall ensure that eligibility decisions and client movement through the VR process occur in a timely manner appropriate to the needs and capabilities of the clients.

This Standard seeks to avoid delays in the VR process which are likely to impede or hinder successful rehabilitation of the client. Rather than set a Performance Standard using time-in-status to define "undue delay," this Procedural Standard requires that each state have a monitoring or flagging mechanism for cases remaining in statuses over a given length of time, and a follow-up procedure to evaluate the appropriateness of any case delay. Many state VR agencies already have variations of such a system in place.

4. VR shall provide an Individualized Written Rehabilitation Program for each applicable client, and VR and the client shall be accountable to each other for complying with this agreement.

There are several aspects of the Individualized Written Rehabilitation Program which would be addressed in this Procedural Standard: (a) compliance with the requirement that an IWRP be fully developed for clients accepted for services or extended evaluation; (b) assurance of the protection of client rights and client awareness of the remedies available for mitigating dissatisfaction; (c) joint client/counselor development of the job goal and the service plan; (d) mutual client/counselor responsibility for follow-through on the agreement and annual review of its progress and appropriateness; and (e) the appropriate handling of plan revisions.

The regulations concerning the IWRP indicate that compliance with this Procedural Standard will ensure compliance with the legislative intent

of the IWRP. Since research has supported the premises underpinning the IWRP by showing that the process and the possession of the IWRP affect client outcomes positively, adherence to the IWRP requirements become a powerful norm for quality case management in VR, as well as a protection of client interests and rights.

Supportive Evaluation Elements

Several elements important to the VR service delivery process and organizational structure have not been covered in the proposed Performance and Procedural Standards. Earlier discussion of the role of Standards has indicated that there will be items which, while they have potential importance as explanatory factors with respect to differences in performance, are not appropriate for establishing Standards directly. Elements to be explored as explanatory factors include the following:

1. Outreach and referral practices
2. Staff mix
3. Staff training
4. Administrative and organizational structure
5. Administrative costs, fiscal and budget elements
6. Caseload size
7. Caseload balance
8. Caseload management: continuity, frequency of contact, supervision, monitoring service delivery
9. Job placement activity
10. Service mix, including post-employment services
11. Reliability of service delivery
12. Service costs
13. Use of similar benefits

As discussed earlier, the supportive evaluation structure for the Standards system provides the ability to answer the all important question of "why." What are the factors which lead to successful performance on the Standards in some agencies and unsatisfactory performance in others?

Alternatively, why does a given state agency do well on some Standards, but not on others? Without this evaluation structure, the Standards exist in a vacuum and bear no relationship to improving services and state performance. With supportive evaluation, RSA and the states can examine differences across agencies in, for example, organizational structure, costs, caseload sizes, staff mix, or case management practices, to determine which of these contribute to differences in performance on the Standards. Agencies are then in a position to identify what changes should be made in current structures and practices to improve performance on the Standards and, because the Standards are framed in terms of the ultimate goal of VR - quality outcomes for clients - better serve the disabled population.

Supportive evaluation and analysis will often be most beneficial when they occur not within a single state but across states so that important differences can be identified. Program Administrative Reviews, contracted studies, and analysis of R-300 and IWRP data are all potential vehicles for this type of analysis. The particular methods appropriate to the various supportive evaluation elements will be considered as the new Standards are field tested.

It is important to understand that the 13 items listed do not form an exhaustive list. At different times, questions about program performance will vary, in terms of the particular Standard of concern, or hypotheses about problems in performance. New elements, other than those already listed here, may become important to look at in terms of supportive evaluation. So these items are of a different nature than the 12 Standards. They are a list, but only an illustrative one, of the kinds of factors which might be important to look at in explaining how states perform on the Standards.

The Role of Standards in Evaluation

Before looking at specific data elements, it is important to understand how the Standards should be used and how they should fit into a state evaluation unit. The role of Standards in evaluation is a topic which is central to ensuring that the Standards do not just become a cumbersome reporting exercise to be gotten out of the way once a year.

While the following picture may not be an accurate one of what goes on in every state, too often a state turns in its data on the current nine Standards and then in an entirely related series of activities, carries out other program evaluation activities which address issues of concern. For example, some states are very interested in how many Status 26 closures are produced and how one district compares to another. Some are thinking about the quality of closures and developing quality review systems. Others are concerned with administrative

structure and identifying the best way to administer the program. These interests are not seen as bearing any relation to the Standards; they are simply two separate activities.

BPA would like to see something substantially different materialize. Our basic concept of how the Standards fit into an evaluation unit is as follows. Most reporting on the data elements should come directly from the R-300 or perhaps from the Social Security data link, and thus would be RSA's responsibility. Annually, RSA will report on state and national program performance on the data elements.

This will be supplemented by an annual analysis of performance on Standards which identifies differences in how the states are doing and looks at the reasons for differences and identifies problems and strengths and what knowledge and procedures can be transferred from one state to another. This must be done centrally because no one state can compare all the states to each other and see why some are doing better than others. It requires "looking at the whole picture."

The Role of the State Program Evaluation Unit

What is the role of the state program evaluation unit? First, the states will be responsible for the procedural Standards and those data elements of performance Standards requiring client surveys or case review. BPA's preferred recommendation is that a follow-up study not be a requirement; rather, Social Security data should be used to look at client retention of economic benefits, and questions about satisfaction and utility should occur at closure for a national sample of clients. This would save the resources that go into expensive follow-up studies which attempt to track clients and yield a very small and perhaps biased response anyway.

Second, BPA sees the program evaluation unit designing its activities to form "supportive evaluation" focused on the concerns addressed in the Standards and using their evaluation resources to try to explain why they are having problems and to try to figure out what can be done to improve matters. The essential point is that the Standards should not represent an isolated effort one week out of the year for an evaluation unit, totally separate from its main evaluation activities. Evaluation activities should be built around each other.

Data Elements: Measures for the Standards

Data elements are the crux of the Standards refinement. The Standards will stand or fall based on the quality of the measures. The data element is that which receives attention; people may not be able to recite the Standard word for word, but they will be able to recite the data element

and identify every problem in collecting it and reporting it and, if appropriate, indicate exactly how it fails to accurately measure the Standard. Thus, an approach has been taken that if we can't come up with good data elements for a Standard, the Standard shouldn't be there. Poor data elements not only lack congruence with the concept supposed to be measured, but they can actually take the program in directions that are not helpful to it. The 26 closure is the prime example of such problems. To some extent, it has taken the program in the wrong direction of valuing "numbers" more than quality rehabilitation.

Thus, Standards built on the best of intentions, but built around data elements measuring something else, with problems of consistency and comparability across the states, are unacceptable. The main criteria for data elements are that they accurately reflect the concept of the Standard and be validly measurable. One other criterion that was followed is that of simplicity. Where a very complicated data element could have been chosen, the more simple and straightforward measure available was selected. Table 1 presents a summary of the proposed data elements for the Performance Standards.

Performance Levels

The problems with the current central tendency approach to performance levels have been discussed earlier. What are the ways to set performance levels? That is, given a Standard and its data element, how are expected levels of performance established for states? Performance levels could be set in the way they have been set before: determine the average state agency performance and set performance minimum or maximum levels at plus or minus one standard deviation from the mean.

Performance levels could be set by policy. Policy makers can make explicit the level of performance expected. Many state agencies have suggested this approach. Agencies want to know what RSA expects from them and they believe performance levels should be goal oriented, not oriented to what happens to be the level of past performance. Some problems exist with setting performance levels by policy: Who does it? Is it a committee? Who is on the committee? How do members of the committee agree on what is the acceptable performance level?

Another possible way of looking at performance levels is to establish upper or lower cut-off points based on percentiles. The bottom ten percent of the agencies, for example, on any given measure can be considered to be outside accepted performance expectations. This is subject to the same criticisms as the current central tendency approach. It does have the merit of ease of interpretation and doesn't require agreement on what the ideal policy expectation is.

An approach which has not been talked about before but which BPA

believes has merit is something called "progress norms." Instead of setting a performance level expectation for any given year, "progress norms" is an approach which concerns year to year progress and expects states, until they achieve a certain high level of achievement, to be improving at a given rate from year to year. The performance level would require a two percent increase or a ten percent increase, for example, in the number of clients competitively employed. The performance level is established "over time," rather than setting an absolute level for any given time. This approach clearly directs the program toward better and better performance.

The last possibility is that there be no performance levels at all. Perhaps arbitrary performance levels should not be set, but an examination should be made of the state performance and supportive evaluation used to analyze why some states do better than other states, and a determination made whether a given state could do better. To the extent that legitimate and substantial reasons exist for state differences, perhaps no standardized performance level is necessary or appropriate.

Table 1

Summary Table of Data Elements Recommended for Further Consideration as Measures of the Eight Performance Standards

Standard	Data Elements	
#1 Coverage	(i)	% national caseload served / % national expenditures
	(ii)	# served in given year / 100,000 population
	(iii)	# served in given year / RSA-generated eligible population based on current methodologies (i.e., Ridge-Worrall formula for General Agencies and JWK-Westat formula for Blind Agencies)
#2 Benefit-Cost and Cost Effectiveness	(i)	Total agency expenditures per competitive employment closure
	(ii)	Total agency expenditure per rehabilitation
	(iii)	Benefit-cost ratio (RSA model)

Table 1 (Continued)

Standard	Data Elements	
	(iv)	Discounted net present value (RSA model)
#3 Successful Rehabilitations	(i)	# of 26 closures / # of 26 + 28 + 30 closures
#4 Increased Economic Independence	(i)	% 26 closures with wage at/above federal minimum wage
	(ii)	Mean wage of competitive closures / Mean wage of employees in state
	(iii)	Mean wages year following closure / Mean wages year preceding referral or Mean wages at closure / Mean wages at referral
	(iv)	Comparison of % receiving public assistance at closure and referral
	(v)	% change in amount of public assistance from referral to closure
#5 Competitive Employment	(i)	% 26 closures competitively employed
	(ii)	% 26 closures who have: (a) Competitive goal and competitive outcome, or non-competitive goal and non-competitive outcome (b) Non-competitive goal but competitive outcome (c) Competitive goal but non-competitive outcome
	(iii)	% non-competitive closures judged to show improvement in gainful activity

Table 1 (Continued)

Standard		Data Elements
#6 Causality		Recommendations Pending
#7 Retention of Benefits	(i)	% of 26 closures with jobs at closure, who retained jobs at follow-up (1 year)
	(ii)	% of 26 closures with jobs at closure, who retained or increased wages at follow-up (1 year)
	(iii)	% of 26 closures remaining off or going off public assistance between closure and follow-up
	(iv)	% of non-competitive closures assessed at follow-up to have maintained or improved closure skills
	(v)	% of sheltered workshop and BEP 26 closures with competitive employment positions at follow-up (1 year)
#8 Client Satisfaction and Utility Appraisal	(i)	% of closed clients satisfied with overall VR experience
	(ii)	% of closed clients satisfied with specific aspects of VR
	(iii)	% of clients judging services received to have been useful in obtaining their job/homemaker situation
	(iv)	% of clients judging services received to be useful in current performance on the job/homemaker situation

CHAPTER V

FACILITIES INFORMATION SYSTEM: A PROGRAM EVALUATION RESOURCE

Gary Prazak

Walker and Associates is currently under contract to the Rehabilitation Services Administration to develop a model Facilities Information System. The purpose of the project is to develop a model information system for use in measuring, evaluating, and managing rehabilitation facility performance. Rehabilitation facilities constitute a significant part of the rehabilitation services network. The Greenleigh Study found that there are approximately 3,000 workshops providing services in the U.S. Further, it is estimated that there are an additional 200 to 300 vocational facilities that are not workshop connected, and 200 to 300 medically-oriented facilities. Nationally, about 16% of federal case services money spent by state VR agencies was through rehabilitation facilities.

Project Goals

Eight major goals for the system have been identified. It is intended that the facilities information system will assist in achieving these goals at the national, state, and local levels. The goals are:

1. Volume: Assure that adequate numbers of types of facilities are available
2. Quality: Assure that quality facility resources are available
3. Distribution: Assure an equitable distribution of facilities
4. Use of Facilities: Assure that facilities are maximally utilized
5. Clients Served: Assure that facilities are serving those individuals they are intended to serve
6. Quality of Use: Maximize the quality of facility utilization (timeliness and appropriateness to use)
7. Benefits: Maximize the benefits obtained from facility program services
8. Efficiency: Maximize the efficiency of facility programs

Components of the System

The Facilities Information System consists of four major components. They are:

1. An inventory of facility resources available for obtaining and presenting descriptive information regarding the types and amount of resources available in rehabilitation facilities

2. The Information/Evaluation System for the collection and presentation of information on facility programs regarding who is served, benefits achieved, and costs

3. Accounting System Guidelines for rehabilitation facilities which will be used primarily for the facilities in determining program costs for agreements between state rehabilitation agencies and the facilities

4. Guidelines for Agreements between state agencies and facilities, designed for use by state agencies in arranging for the purchase of services and the review of facility performance

The following is a more detailed discussion of the system components.

Inventory of Facility Resources

The current plans for this inventory are that it would be administered annually to collect descriptive information on all facilities used by state agencies. Descriptive information would be summarized by the state agency and would be used at the state and facility levels for a variety of purposes. The statewide summary information would also be forwarded to the central RSA office to be aggregated with similar inventory information on facility resources from other states. For the first time descriptive information regarding all the facilities used by vocational rehabilitation in the nation would be assembled.

Some of the categories of information being considered for inclusion in this descriptive inventory are as follows:

1. The type of facility, organizational setting, and principal programs

2. The types of accreditations the facility has achieved

3. Number of staff by major types and amounts of revenue for the total facility

4. Major target population served

5. Specific programs provided to clients of the state VR agency

Information/Evaluation System

The system will be used to generate quarterly reports of information from facilities which would also go to state agencies for statewide aggregation. Annual summaries would be prepared by the state agencies and sent to the RSA central office for national aggregation and preparation of national reports. The focus of this information would be on who is being served by facilities, benefits achieved, and the efficiency with which these benefits are achieved. The plan is to summarize this information at the program level within the facilities. "Programs" means such things as work adjustment, skill training, vocational evaluation, in-patient physical restoration, and the like. More emphasis will be given to information on benefits achieved rather than on the process itself.

The Information System is being designed with a central core of information that would be common from all facilities and from all states. Individual states may also collect additional, perhaps more detailed information from facilities for their unique purposes. It is anticipated that facilities themselves will be interested in the information going beyond this common core.

The areas of benefit information currently being considered for inclusion in this system are, first, the achievement of various types of rehabilitation such as competitive employment, sheltered employment, etc., as well as the levels of earnings of persons achieving such employment. Also considered are benefits which are "proxy" to rehabilitation status, such as "in placement for employment" or "in training" or "in an employment goal program." The third type of benefit information collected would be defined by the individual facility and would be specific to the particular program type as an indicator of clients who are not vocationally rehabilitated, but nonetheless have achieved some other benefits considered significant by the facility.

The next category of information, beyond benefits information, would be information regarding who is being served by the facilities. The number of severely disabled persons will likely be included, as well as a summary of major disability types and other demographic variables.

Just as the federal evaluation Standards call for cost information related to benefits, the facilities information system will include efficiency information such as the amount of state-federal VR costs for clients served by facilities and the relationship of these costs to benefits. The length of program participation will likely also be included in this area of efficiency information.

An additional area of information that state agencies are expected to be interested in would be that which would assist them in monitoring their use of facility programs, such as referral rates, waiting lists, program capacity, etc.

Accounting System Guidelines

Additional components to the project on (a) facility accounting, and (b) agreements between state agencies and facilities are the responsibility of a project subcontractor. These components were included in the project largely in response to a GAO Audit Report which was critical of the relationship between the state agencies and the facilities, and to strengthen this relationship.

The accounting system guidelines for rehabilitation facilities is being designed to help facilities determine the real cost of providing services to VR clients. It should also assist facilities in preventing some of the problems that have been identified by various audits, such as inadequate documentation of costs, duplicate payments, etc. It will not be a uniform accounting system requiring all facilities to keep track of costs in exactly the same fashion; it should, however, assist facilities in designing a system unique to their particular needs.

The guidelines for agreements between state agencies and rehabilitation facilities will include suggestions for improving the arrangement for purchasing services, presenting various alternative arrangements for the purchase of services, and describing the elements that should be part of any contract or purchase of services agreement. Guidelines will also be included regarding the monitoring of facility performance by state agencies in terms of who should be responsible for various aspects of monitoring, specific procedures that can be used, and the frequency with which monitoring should be done.

Design Issues

Several major system design issues have been identified at this point in the development of the Facilities Information System. Several design issues will be presented. The first issue is the need to strike an appropriate balance between including enough information in the system so that it is truly useful at the facility, state agency, and federal levels, and at the same time keeping the system simple and concise enough to be accepted. It is recognized that many, perhaps most facilities get substantial income from sources other than the state vocational rehabilitation agency. The varied income base limits the leverage of state agencies and also means that most facilities have to respond to other reporting requirements as well. An inclination is to design this system so that it includes the simplest, most basic information that can be demonstrated to have value for direct management, planning, and resource allocation at the facility level, for state VR agencies, and for federal RSA.

The second design issue involves developing a universal typology of programs for the categorizing of facility activities with VR clients. The typology will include those programs which are clearly vocationally focused, such as skill training, vocational evaluation, and work adjustment. It will also need to include programs provided by medically oriented

facilities, such as in-patient and out-patient restoration, medical evaluation, and comprehensive rehabilitation programs. It will also include special adjustment programs for the blind and those special grant programs within facilities that have similar objectives across facilities, such as Training Service grants and Projects With Industry.

A third design issue that has been identified involves the responsibility for aggregation of data. An earlier RSA effort at gathering facility information, the Reciprocal Rehabilitation Reporting System, was designed for individual facility data to flow to the RSA central office. The data were then to be summarized there and flow back to state agencies and facilities. This system failed to work for a variety of reasons, principally because the system was voluntary and few facilities chose to participate, and because of the insufficient resources for processing individual facility information in a central location.

Current plans call for the aggregation of data at the facility level to be done by the facilities themselves. Data aggregation provides the advantage of faster feedback and enables facilities to readily integrate this information with their other informational needs. For example, facilities which are collecting and aggregating the information expected to be included in this system will be well on the way to having many of the necessary elements of an internal program evaluation system such as called for in the Commission on Accreditation of Rehabilitation Facilities Standards, Section 9.

The state agency would be responsible for aggregating information across all facilities in the state and submitting the statewide aggregations annually to RSA regional and central offices. Most of the state agencies contacted have indicated they are not interested in or able to process individual client data for individual facilities. However, the system will allow for the option of processing for those states that do choose to process data for individual facilities.

The need is recognized for the integration of the Facilities Information System with several other reporting efforts currently being planned or revised, such as the general Standards for evaluation and the state facility plan. It will be important before national installation of the Facilities Information System that all of these efforts be coordinated so that the same categories of results are used, similar definitions apply to commonly used terms, and duplication of effort is avoided.

Finally, some comments will be made regarding the relationship between state agency evaluation units and the Facilities Information System. It appears that most state program evaluation units do not now have any significant role regarding state agency use of facilities or in evaluating facility performance. The Facilities Information System will be generating considerable information for use by various people within state VR agencies and it would seem appropriate that the program evaluation unit be involved in the interpretation and use of this information. The information should be of considerable value, not only to facility specialists, but to planners

and program staff, including counselors. It may be an appropriate role for the program evaluation unit to assist these people in using the information regarding facilities.

Potential Uses of the System

Some of the potential uses that could be made of this information at the state level include:

1. Comparing facilities with similar programs regarding who they are serving, costs and benefits achieved

2. Comparing facility performance within the state to performance in other states

3. Enabling individual counselors to better know what to expect from specific facility programs they may be considering using

4. The information should lend itself to special analyses regarding differential results by different types of facilities such as comparing the performance of state operated to privately operated facilities

5. It will also be useful in determining the impact on costs or results of changes made within the facility. This might include changes brought about by specific grant programs, changes in staff, or specific program changes

The field tryout of the system should provide useful feedback for making necessary changes in the model systems and fine tuning it to improve its utility. The field tryouts should also be useful as a demonstration of the value of the system as a tool for facilities, state agencies, and for RSA toward the ultimate better management and improved performance of rehabilitation facilities.

CHAPTER VI

THE USES OF PROGRAM EVALUATION DATA WITHIN A STATE REHABILITATION AGENCY

Stanley E. Portny

Introduction

The 1968 Amendments to the Federal Vocational Rehabilitation Act for the first time specifically required state rehabilitation agencies to evaluate the services they were providing. For the first time, federal funds for evaluation of the vocational rehabilitation (VR) program were made available. During the following ten years a tight economy has limited the availability of federal program funds, and a growing emphasis has developed on accountability in federally sponsored programs. As a result, there has been a considerable increase in the amount and types of evaluation activities conducted in and of the VR program.

Evaluations of the VR program are conducted by a wide variety of offices and organizations, ranging from the U.S. Congress through the various levels of the Department of Health, Education, and Welfare (DHEW), state legislatures, and state rehabilitation agencies. Since these evaluations all address various aspects of program operations and/or service provision, the assistance of the state VR agency is often requested to facilitate the collection of data on or about clients and the provision of VR services. Due to the diversity of requests and the individuals initiating them, the generic similarities in the issues explored and the information required to explore them are sometimes overlooked. It is also often not realized that the ultimate purpose of all evaluations conducted is to generate information which will enable more competent decisions to be made regarding program mission, goals, structure, policy, and operations procedures.

This paper explores the nature of evaluation, specific characteristics of evaluation in VR, the purposes of evaluation activities, and ways in which evaluation, at all levels, can ultimately lead to program change and improvement. In addition, suggestions are offered for increasing the effectiveness of evaluation and the potential utility of the results. Possible uses of data generated in the conduct of these evaluations for purposes other than those originally intended are also discussed. Comments are offered on how to complement and augment the capability of the state VR agency evaluation staff through the use of specialized consultants.

Definition and Role of Evaluation

The term "evaluation" has come to be widely used in the literature of management and administration; however, the particular definition of the word frequently reflects the context in which it is employed. A broad and comprehensive definition is proposed by Suchman (1967, pp. 31-32) as being "the determination...of the results...attained by some

activity...designed to accomplish some valued goal or objective.:." The process of evaluation consists of three separate steps:

1. Selection of performance indicators and criteria with respect to which performance will be assessed

2. Monitoring or observation of the activity

3. Assessment of the observations with respect to the criteria

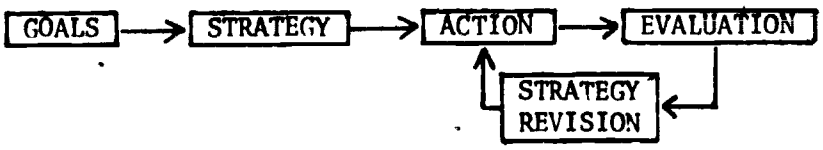
To facilitate subsequent use of evaluation results, it is helpful to include a fourth step in the process:

4. Formulation of recommendations based upon the results of the assessment

Evaluation is an essential element in any Management by Objectives (MBO) system designed to help a program or organization achieve its mission and goals. The specific elements of an MBO system include (a) development of goals and objectives; (b) formulation of a strategy for achieving those goals and objectives; (c) implementation of the strategy; (d) evaluation of program operations; and (e) modification of the strategy, if required. MBO is a continuous and cyclical process, as illustrated in Figure 1.

Figure 1.

A Management by Objectives (MBO) System



The particular scope and focus of an evaluation will be guided by the program goals of immediate concern, the program components being examined, and the types of action plans being contemplated. An evaluation may be summative or formative. A summative evaluation is designed to answer the question, "Does this program work or not?" Summative evaluations result in comprehensive statements about the net worth of a program and alternative actions based upon such evaluations may include either continuation or termination of the program.

A formative evaluation addresses the question, "How may this program be improved?" Formative evaluations do not seek to make judgments about whether a program should continue to exist; rather, they are designed to determine the quality of program performance, possible areas for improvement, and ways in which to achieve this improvement. In recent years, the trend in VR activities has been toward formative, rather than summative evaluations.

Evaluations may examine the VR system in terms of input, process, or outcome. System inputs include the level of resources utilized and the number of characteristics of clients entering the program. Process refers to the activities conducted in the provision of services. Outcome refers to changes in clients' vocational/economic status or other non-vocational chracteristics which occurred as a result of program participation.

Evaluation may focus on the effectiveness, impact, or efficiency of a program or activity. Effectiveness refers to the extent to which program objectives related to desired changes in individual clients are attained as a result of the program. Impact refers to changes which occur in a specific target group or in society in general, as a result of the program. Efficiency is the economy with which available resources are used to achieve program objectives.

It is possible for a program to be very effective but to have a relatively small impact on the entire target population. Such a program may achieve remarkable results with those people it serves, but may serve such a small fraction of those eligible that the overall problems of the target population are only minimally improved.

Likewise, it may be that a program is very effective in serving individual clients but is highly inefficient. In this instance, it is possible that the program could be expending far more resources per client than in fact is necessary to achieve a certain level of effectiveness. However, since efficiency is a relative assessment of the utilization of resources to achieve program objectives, an ineffective program (one which does not achieve its objectives with respect to individual clients) can never be classified as efficient. It should be noted though that an ineffective or inefficient program can have a significant impact upon the target population and society.

Evaluation in the VR Program

The VR program is a state-federal partnership. The program was established by federal legislation, and both federal and state funds are utilized to support program operations. Responsibilities of the federal government include the administration of federal program funds, the development of program policies, regulations, and guidelines, as well as the funding and coordination of evaluation, research and demonstration, technical assistance and training support for the general benefit of the program.

State rehabilitation agencies provide and coordinate the provision of services to clients, administer state and federal program funds, and develop administrative policies and procedures for program operations. A variety of government offices and agencies have some degree of authority with respect to aspects of the VR program. As illustrated in Figure 2, four distinct federal government bodies, including the Congress, offices with general program and fiscal responsibilities in the Department of Health, Education, and Welfare, and the central and regional offices of the Rehabilitation Services Administration (RSA) have some measure of

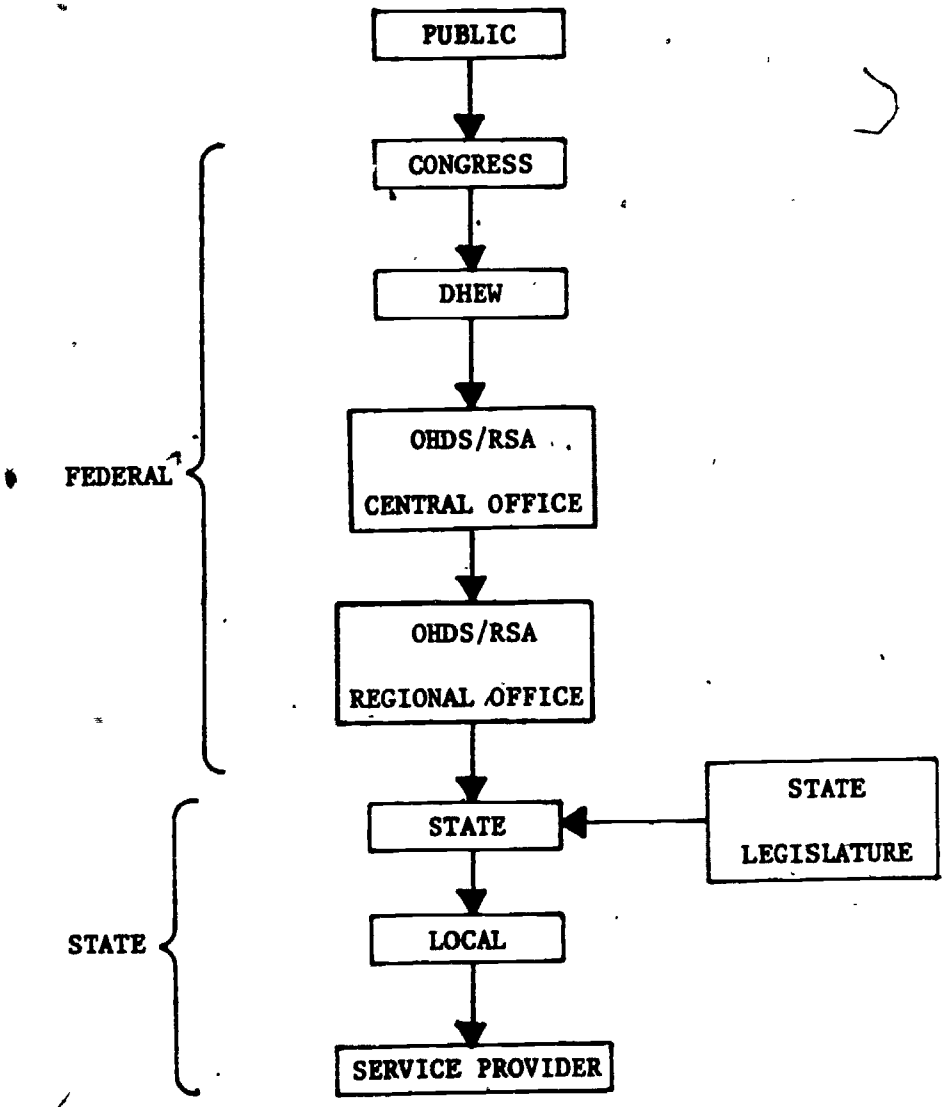


Figure 2. Levels of Authority in the VR Program

responsibility for VR program operations. The state VR agency is in turn responsive to both the federal government and the state legislature. In general, higher levels of authority are concerned with program issues of broader interest and applicability. The highest level of authority, the U.S. Congress, is ultimately responsible to the public.

Each individual responsible government office has its own program responsibilities and, consequently, its own goals and objectives to achieve. To ensure that goals and objectives are being achieved, each must conduct its own evaluation of the particular program elements of interest.

Various government offices use different methods to conduct their evaluations. The U.S. Congress may use the General Accounting Office (GAO) Audit Agency to investigate programmatic and fiscal operations, or it may require, through legislation, that specific evaluation activities be conducted either by offices within the federal government or within the state VR agency. DHEW and RSA can have evaluations conducted internally by agency staff, they can require that state VR agencies conduct particular evaluations, or contract for evaluation studies to be conducted by private organizations. The state VR agency can conduct its own evaluation internally or contract with private organizations.

The ultimate purpose of evaluation at all levels is to assist each responsible agency to achieve its goals and objectives with respect to the VR program. However, government agencies and offices have available different methods for implementing action plans to affect program performance. The U.S. Congress can modify and amend federal program legislation. The federal agencies can develop and modify program policy and guidelines and the state VR agency can modify operational procedures.

In general, particular government agencies have the authority to enforce implementation of their plans and requirements, within the scope of their particular operations. That is, directives from the U.S. Congress will apply to individual programs within all states, while directives from a state agency will affect operations within the agency. In general, action or implementation plans at any level will affect operations at the same and lower levels. Since service provision occurs at the state agency level, all action or implementation plans will ultimately affect state agency operations in some way.

Evaluation in the State VR Agency

Evaluation activities are conducted in the state VR agency for one of two purposes: (a) to provide information and respond to the directions of audiences outside the agency; and (b) to facilitate effective internal management of agency operations. Often, the general issues addressed by both types of evaluations may be similar, but the level of detail, as well as the perspective of the information presented may differ. Though it is frequently perceived that evaluations for internal use will affect agency operations while those for external audiences just result in the production of reports, it is important to remember that the information

contained in those reports can form the basis of program actions taken at other levels of government which ultimately will affect state VR agency operations.

Evaluations for External Audiences

Evaluations in the state agency may be conducted in response to the demands of the U.S. Congress, DHEW, RSA, or the state legislature. The specific issues addressed may relate to (a) compliance, (b) program justification, (c) program accountability, or (d) program improvement.

Compliance

Either an evaluation of the program may be conducted to determine the extent of compliance with specified requirements, or the very conduct of the evaluation activity itself may constitute compliance with the requirements. An example of the first type of compliance-related evaluation is the preparation and submission of reports on case cost, case flow, and other program operations addressed in the federal VR Standards. The information presented in these reports is compared with required performance levels to determine whether the program is satisfying the requirements.

Examples of the second type include the conduct of follow-up studies of closed clients as required in the VR Standards, or the annual review of ineligibility determination as required by program legislation. In these cases, the compliance requirement is to conduct the studies, rather than to obtain specific results or to take certain actions based upon the results. It is anticipated that the agency will use the information to correct any problems identified. Results of these evaluations may actually form the basis of amendments to program legislation or to program policy at the federal level.

Program Justification

Evaluation for program justification may include (a) assessment of the program's net impact on, or benefit to the target population in order to justify its continued existence, or (b) analysis of program cost and operations in order to justify and support specific budget requests. Actions in response to these evaluations might take the form of increases (or decreases) in federal or state appropriations, legislative amendments modifying the program purpose or structure, and so forth.

Program Accountability

Analyses related to program accountability seek to insure that program funds are being expended in accordance with prescribed guidelines and plans, and that basic program goals and objectives are being achieved. Actions in response to such evaluations might include modifications of program legislation or program policy and procedures.

Program Improvement

Finally, evaluations may be conducted in response to the directions of external audiences, in order to identify problem areas and to develop recommendations for improving program operations, including service delivery. Examples of such activities include the federally coordinated and conducted Program Administration Review (PAR), as well as certain federally sponsored evaluation studies performed by private organizations. These projects are designed to examine state VR agency program operations in order to identify ways in which program performance could be improved. Actions based upon the results of such studies are implemented through plans developed jointly by federal and state program offices.

Evaluations for external audiences may either be performed by outside organizations, performed jointly by an outside agency and state VR agency staff, or conducted entirely by the state VR agency. In each of these cases, the state VR agency may be required to perform several different types of activities in support of these evaluations. Required program data may have to be collected, statistical reports may have to be prepared and submitted, or agency records, caseload data, and other information may have to be made available to external evaluators.

Evaluations for Internal Use

Evaluations for internal use are primarily designed to facilitate the improvement of service provision or program operations. These evaluations may be systematic and ongoing or ad hoc. Systematic evaluations are planned activities which are repeated at scheduled time intervals. Ad hoc evaluations are designed to address a particular problem or concern of the moment. Both types of evaluations serve as essential elements in the agency system for program management.

Prime audiences for results of evaluations conducted for internal use include agency top management, middle management, first line supervisors, counselors, and other service providers. The following are candidates for inclusion in the audience for a particular evaluation activity:

1. The individual or office who requested the evaluation
2. Those people with authority to implement change in the program elements being evaluated
3. Those who are being evaluated
4. Those who will be involved in or affected by any changes implemented as a result of the evaluation

The results of evaluations can be used in two ways to help improve program performance. First, the information can be presented to and discussed with those staff members who were evaluated. People are often unaware that their performances are not measuring up to expectations.

Therefore, the first step, reporting to them the results of assessments of their performance, will enable them to note problems with their performance and take steps to correct these problems.

Second, it is important to report the results of evaluations to program managers, so that they may devise action plans and appropriate employee incentive schemes to encourage the correction of problems and the continuation of desired performance. Toward these ends, evaluation data may be used in program planning and goal setting, performance assessment of agency sub-units, caseload management, and employee performance appraisal.

Comparison of Evaluations Designed for External Audiences With Those Designed for Internal Use

Though program evaluations for internal and external audiences often address similar issues, certain differences in their design and conduct are of importance. First, evaluations for internal use are designed to address issues of current importance in the operations of the state agency. The results of such evaluations may form the basis for changes in operational procedures which will have an immediate impact on agency operations. Evaluations for external audiences, on the other hand, may address issues of concern from a broader program perspective, but may be of little immediate operational utility for the state agency. As a result, staff may devote less effort to insuring that the evaluation is rigorously and thoroughly conducted. All program evaluation information has the potential for influencing actions at some level of program authority which ultimately will affect state agency operations.

Second, evaluations for internal use are designed and conducted specifically to identify ways to improve program performance. However, the motivations behind preparing reports for external audiences is often to convince these audiences that operations are satisfactory as they are. In other words, the purpose of these evaluations is to convince an audience that no program changes or actions are required. Therefore, the results of evaluations for external audiences often fail to explore potential problem areas in program operations to the extent necessary or desirable to facilitate the correction of these problems.

Finally, the end result of evaluations for internal use should be changes in operations which will result in improved program performance. The end result of evaluations for external audiences is often perceived to be just the completion of a report. This apparent lack of direct connection between evaluation results and program action may cause people to feel that evaluations for external audiences are of little real value. It is important to recognize that these reports may be instrumental in leading to program changes, even though the changes may be formulated at levels outside of the state VR agency.

Sources of Evaluation Data

Data required for program evaluations may be obtained from sources external and internal to the state agency. External sources include

federal audit reports, RSA statistical reports, federally funded evaluation studies and reports, or studies from other state agencies. Internal data sources might include case records, management and caseload information systems, and state agency conducted studies.

Though broad utilization of existing data is to be encouraged, there are certain potential problems which must be carefully considered when using data for other than their original purpose. It is important to insure that the definitions of the existing data elements are identical to those of the data elements being sought. Terms such as "target population," "active caseload," and "severely handicapped" may have different definitions, depending upon the context in which they are used. Similarly, the time intervals for which data apply should be specified and compared for consistency.

If the existing data were developed based upon a sample, it is essential to verify that the sample is valid and appropriate for the current purpose. As an illustration, data collected for the purpose of determining trends in national program performance are often insufficiently detailed to enable the determination of trends in individual state agency performance.

It is important to ascertain whether there may have been a pre-existing bias which affected the values of the existing data. This is of particular importance, for example, when considering the use of data for program improvement purposes which were originally collected in response to compliance requirements. Sometimes it is possible to correct for the bias with statistical procedures, while in other instances the data may be of no use for the present purpose.

Finally, it is important to insure that the existing data are current enough for the present purpose. Since the preparation of reports which present the existing data is a lengthy process, the data presented may be too old to be of use for the current requirement. More recent data may exist within the state agency which are more suitable for the purpose.

To illustrate some of the points mentioned in the preceding section, two examples of evaluation activities conducted in the state VR agency will be discussed. The first, a weighted case closure system, is a method for using information on the quality and characteristics of case service provision to improve agency operations. The second, the federal VR evaluation Standards, entails the evaluation of operations and the preparation of descriptive reports in response to an externally generated requirement. Particular attention will be directed towards exploring the potential utility for internal management use of data collected in accordance with the VR evaluation Standards.

A Weighted Case Closure System

A caseload analysis system is a set of procedures for describing the nature and quality of services provided to clients, the effort involved

in the provision of those services, and characteristics of the client population served. Caseload analysis systems may be designed to provide overall assessment of cases after they have been closed or they may be designed to provide continuing evaluation of cases through their active statuses, up to, including, and after closure. In the former instance, the quality of services provided is often reflected by the type of closure achieved, characteristics of the closure, and characteristics of the client at closure. Systems of this type are referred to as weighted case closure systems. In the latter instance, the quality of services provided to cases which are still active may be reflected by the extent to which intermediate objectives (i.e,, milestones reached on the path toward closure) are reached, professional standards of service provision are met, or other characteristics of service provision or program operations are achieved.

A variety of indicators may be used to reflect the quality of services and level of effort associated with the closure of a particular case. The a priori probability of achieving a successful closure for a client with the demographic characteristics of the case under analysis, the case service costs expended, hourly wages and other vocational/economic descriptors of the client's employment situation at closure, and non-vocational client descriptors may all be used to assess the services provided to the subject client. In addition, the extent to which agency priorities for serving specific target groups or for meeting particular budgetary or service goals are successfully achieved may be included in the overall assessment of a particular case.

The purpose of these systems is to encourage the provision of services to all eligible clients in accordance with program priorities. The purpose is achieved by providing information to staff regarding their performance and by facilitating the development of staff incentive systems which reward personnel in accordance with their performance assessments as developed by these systems. For example, greater credit might be received for:

1. Rehabilitating a person with a lesser probability of success
2. Achieving a quality rehabilitation at less than average cost for such a rehabilitation
3. Rehabilitating clients in identified priority groups

Information developed in accordance with a weighted case closure system can be utilized to improve program operations in two ways. First, the information should be reported back to the agency staff so that they can see to what extent their performance is meeting agency goals and objectives. They, in turn, can use this information to improve the delivery of services to clients. Second, information on counselor performance should be related to agency management. Management may then reward those staff members who are performing well and encourage and assist those who are not performing. The ultimate success of this system depends upon the extent to which agency personnel can coordinate their efforts to provide services of acceptable quality to the appropriate

target population.

Numerous sources of data are useful in the initial development of the weighting systems governing the relative credit to be given for achieving particular types of rehabilitations. Possible sources include existing agency caseload and management information systems, RSA-300 report sub-missions, and federally funded evaluation studies which examine the provision of services to particular types of clients.

Information on closed cases which is required to conduct ongoing evaluations of agency operations and services provided could be collected from existing statistical reports or specially designed reporting forms. Care should be taken to minimize additional reporting requirements.

In addition to guiding the types of clients served by the agency and the quality of services they received, information generated by a weighted case closure system has other possible uses. It may serve as input for more comprehensive evaluation of the effectiveness and impact of agency program operations, as a basis for a counselor performance appraisal system, or as input into agency planning and the selection of agency goals and operational activities. Further, the detailed analysis of service provision activities may be useful for estimating the need for future resources, developing supportive information for budget requests, and explaining the extent to which current operations are realizing the general goals and objectives of the VR program.

The Federal VR Evaluation Standards

The federal VR evaluation Standards were developed in response to a specific Congressional mandate included in the Rehabilitation Act of 1973. The Standards were designed to establish criteria for the evaluation of program effectiveness, to increase program accountability, and to require state VR agencies to carry out more comprehensive evaluations of their programs. To accomplish these purposes, the Standards were designed to address four basic issues:

1. The degree to which the eligible population is reached and affected

2. The degree of change towards the goal of gainful employment experiences by clients as a result of VR services

3. The extent to which the program meets the mandated priority for serving the severely handicapped

4. The efficiency with which resources are used to accomplish program goals

The Standards are structured to focus on certain program characteristics, processes, and outcomes which, if not achieved or conducted in accordance with specified criteria, would suggest that the program is not providing services of acceptable quality.

The state agency's responsibilities with respect to the Standards include the collection of program data, the performance of certain studies (such as client satisfaction and follow-up studies), and the preparation and submission to RSA of a report on agency performance with respect to the Standards. RSA then analyzes the data submitted, compares the data with national norms and established performance levels, and submits a summary report back to each state agency describing the extent to which the agency is in conformance with the Standards. RSA also prepares reports detailing analyses of the data from regional and national perspectives.

In addition to satisfying a mandated requirement, much of the information contained in the Standards reports prepared by the state agency and RSA may be useful for internal agency applications. Certain case flow information, such as time in and between various statuses, may reveal potential bottlenecks or problems in the rehabilitation process. Outcome information, such as mean weekly earnings at closure with respect to mean weekly earnings for the week before referral or compliance of job status and earnings at follow-up with job status and earnings at closure, yields insights into program effectiveness. Cost data, such as mean case service costs for different types of services for different closure statuses, and target population data, including estimates of size and characteristics of cases accepted and served, may be useful for state agency program planning, resource allocation, and budget requests.

The data required for and prepared in response to the Standards are aggregated to describe overall characteristics and behavior of the entire agency, rather than of sub-units, individual staff, or individual cases. Therefore, the Standards data are of most use to agency top and middle management who make agency wide decisions.

For those aggregated data which are calculated from individual case records, the RSA report to the state agency includes not only the mean value of the data element but also its standard deviation. If unusual performance is noted in the behavior of one of these data elements, examination of the ratio of the standard deviation to the mean may help to pinpoint the cause of the problem. If the ratio is small, in all likelihood the problematic behavior will be observed throughout the agency. In this instance, the mean represents a strong central trend and the variations in this trend are relatively minor.

If the ratio is large, there is strong indication of significant variation in the data element among the cases examined. Further analysis would be directed towards determining whether the distribution of the data was normal, uniform or multimodal and, in many cases, whether there were certain characteristics common to cases grouped together, such as by the same district, the same disability, the same counselor, and so forth. Upon determining particular places within the agency where the problem was localized, additional investigations could be conducted to determine specific causes of the behavior.

Since much of the information for the Standards reports is derived from existing caseload information systems, the charge that motivation

to present a favorable picture of agency performance might have affected the detail, perspective or bias of the data is less likely. However, it is still important to consider certain potential problems which might be encountered when using these data for agency program operation assessment and improvement.

The ultimate purpose of the Standards reports submitted by state agencies is to convince RSA that programs operations are satisfactory - that no changes are needed. Therefore, the reports may tend not to explore or identify potential problem areas fully, unless specifically required to do so in the Standards instructions. This may limit the extent to which favorable performance as described by the Standards data should be considered by agency management as a firm indication that no problems exist.

The submission of information by state agencies and the preparation of the summary reports by RSA is a lengthy process. Therefore, the data contained in the reports may not be the most up-to-date data available to the agency at the current time. Though the data may be useful in preliminary analyses, it may be necessary to obtain updated information for particular applications.

Data in the existing Standards do not completely address all aspects of importance in the VR program. Therefore, Standards data should be used as indicators of program performance, rather than as complete and comprehensive descriptors. Agency staff should develop and conduct additional evaluations, as required, to explore all vital areas of agency operations.

Implications for the State Agency Evaluator

The preceding discussion has explored aspects of the varied demands for evaluation information placed upon the state VR agency and of the possible applications for which this information will or could be used. As a final note, there are certain points which should be highlighted in order to increase the utility of program evaluations and the likelihood that results of the evaluations will be utilized in the most effective ways possible.

The need often exists in the state VR agency to satisfy the requirements of external audiences before activities with internal applications can be undertaken. However, when planning the allocation of scarce agency staff and financial resources for evaluation:

1. It is important to remember that information provided to external audiences may eventually form the basis for program change introduced at some level outside the state VR agency

2. It is always advantageous to consider how use might be made of this information in evaluations undertaken for internal application

Several points should be remembered when attempting to plan for the successful utilization of any evaluation information:

1. Always identify those with authority to act upon the results of the evaluation. These people should be a prime audience for the evaluation results

2. Always determine deadlines by which the evaluation results are needed, in order to impact upon relevant planning and decision making. An evaluation report, no matter how excellent and insightful, is of no value if the issue to which it pertains has already been resolved

3. Studies have shown that managers are primarily oriented towards action and do not respond well to written periodic reports. Therefore, communication of evaluation results should be clear, concise, and topic oriented and, if possible, should be verbal as well as written

4. Finally, evaluation results should always include proposed recommendations for action and anticipated consequences associated with those actions, rather than just raw data and statistics

References

Cox, C. B. Managerial Style: Implications for the utilization of program evaluation information. Evaluation Quarterly, 1, 3, 1977.

Davis, H. R. and Salasin, S. E. The utilization of evaluation. In Struening, E. L. and Guttentag, M. (Eds.), Handbook of evaluation research. Beverly Hills: Sage, 1975.

Harris, J. The uses of performance measures in rehabilitation programs. Institute of Urban and Regional Development, Working Paper No. 215/RS022. Berkeley: University of California, 1973.

Health, Education, and Welfare (U.S. Department of), Human Development Office. Vocational rehabilitation programs and projects: Evaluation Standards. Federal Register, 1975, 40 (245), 58956-58961.

Portny, S. E., Kossoy, H. S., Goodman, R., and Smith, L. Analysis of FY 1976 data submitted on the vocational rehabilitation standards. Washington, DC: JWK International Corporation, 1978.

Rehabilitation Services Administration, U.S. Department of Health, Education, and Welfare. Program regulation guide RSA-PRG-76-32: Rehabilitation services manual transmittal number 45 (studies and evaluation.) Washington, DC: Author, 1976.

Suchman, E. A. Evaluation research. New York: Russell Sage Foundation, 1976.

Walker, R. A. The ninth panacea: Program evaluation. Evaluation, 1, 1, 1972.

PART THREE

PROGRAM EVALUATION METHODS AND MEASURES

INTRODUCTION

Beyond having a national set of program evaluation Standards, a need exists for methods and measures which address evaluation issues beyond the Standards. A number of such projects are underway at present. While not encompassing all such activities, the projects are representative of some of the innovative approaches currently under development.

Fred McFarlane and his associates at San Diego State University recognized the need for a uniform and comprehensive case review procedure for the rehabilitation program. Using rehabilitation legislation as a guide, McFarlane and his group have developed a standardized case review schedule based upon agency compliance with RSA rules and regulations. A modified version of the case review has been included as part of the new evaluation Standards developed by Berkeley Planning Associates.

Rehabilitation Indicators constitute a new and basic tool with much potential for application within a variety of rehabilitation settings. Margaret Brown and her associates at New York University developed Rehabilitation Indicators in response to the continuing and long-standing difficulty of documenting the nature and scope of effective efforts in rehabilitation programs. These measures constitute a basic tool which describes observable elements of the client's life that can change as a result of rehabilitation efforts. Central to the development of Rehabilitation Indicators is a model of client change which defines sets of client variables that may change during the rehabilitation program, and relationships between the variables.

Robert Thrall and Larry Glass present work they have been conducting in the area of applications of cost-benefit analysis to rehabilitation. This approach has traditionally received mixed reactions in the rehabilitation community. Thrall and Glass attempt to deal with the concerns of rehabilitation personnel by first providing definitions of different types of cost-benefit analyses. This is followed by descriptions of potential applications of this research tool in rehabilitation programs.

Sara Wagner of JWK International describes a national follow-up study of rehabilitation clients. She discusses some of the methodological problems encountered and how they were handled. In addition, Wagner presents the advantages of a standardized follow-up questionnaire for the entire state-federal program.

Don K. Harrison from The University of Michigan discusses a model for competency evaluation in rehabilitation. The approach, based on a systems analysis framework is being developed by The University of Michigan Rehabilitation Research Institute

The ethical dilemma of (a) desiring a control group for validity purposes in program evaluation strategies, and (b) withholding needed services from eligible clients to achieve such control has limited the credibility of many evaluation efforts. Ken Reagles and John O'Neill suggest a potential resolution to this dilemma through the use of time series, single subject designs which have been developed from learning and behavioral research.

The project to revise the current RSA management information system (MIS) was not initiated until six months after the Syracuse conference. However, this effort is so vital to the future of rehabilitation program evaluation that it was deemed important to include in this monograph a paper describing this undertaking. The paper is a summary of the first stage of the Abt effort to revise the MIS. It is both informative and provocative.

CHAPTER VII

THE CASE REVIEW SCHEDULE FOR PROGRAM EVALUATION

Fred R. McFarlane, Ph.D.
Lana C. Brenes, M.S.

Introduction

The primary involvement of the Regional Rehabilitation Continuing Education Program (RCEP) in Region IX at San Diego State University in program evaluation is in training administrators, supervisors, and/or counselors in the Case Review process. Many thoughts arise as to what would be the most appropriate way to present the Case Review Schedule (CRS) and the Case Review (CR) process. As these thoughts were discussed, a graffiti statement came to mind: "Theories without facts are barren; facts without theories are useless."

How does this statement relate to program evaluation? Let's propose two changes in its wording. First, substitute the term "management goals" for "theories." Second, substitute the term "program evaluation" for "facts." The revised graffiti now reads, "Management goals without program evaluation are barren; program evaluation data without management goals are useless." This is a concept which must be addressed when considering program evaluation approaches. Essentially, why conduct program evaluation? An evaluator must first answer the question, "What is the management goal?" Based on determining the management goal, the evaluator must develop a method of evaluation to gather the data to determine if the goal has been attained.

Taking this approach, it is determined that the management goal for this conference (Program Evaluation for Rehabilitation Agency Personnel) is "to enhance the quality of rehabilitation services to persons with handicaps to employment." That is the management goal not only for this conference, but for the state-federal vocational rehabilitation (VR) system.

To analyze this goal, focus on the word "quality." The basis for obtaining the management goal must originate in the confines of the VR system from the point where the client enters the VR system to the point when the client terminates the VR system. Between client entry and exit is the only time a state VR agency has some control over the "quality of rehabilitation services." Finally, if one is going to look at the quality between these two points, where would one look first? The obvious place is the interaction between the counselor and the client. Therefore, when analyzing the overall management goal, administrators, trainers, program evaluation specialists, and supervisors can only facilitate the interaction between the counselor and the client to attain the goal of "quality rehabilitation services."

Using this deductive process of analyzing the management goal, it is necessary to the program evaluation system at a basic point - the interaction between the rehabilitation counselor and the client. How can this interaction be evaluated? First, examine what the counselor does with a client. There are numerous ways to look at the rehabilitation counselor-client interaction, but the only readily obtainable information available is the client's folder which includes the rehabilitation counselor's case documentation and related data about the client.

This is where the CR process of program evaluation starts. It is imperfect, obviously, because the rehabilitation counselor does not write everything down. It is based on his/her writing style and other variables. But when other options are examined for evaluation, this interaction adds burdens to the rehabilitation counselor. Further, these other options may not be applicable to obtain the necessary information. So why not use what is there? An underlying assumption is that a client's case folder will be used as the representative basis for what happens between the counselor and the client in the confines of the VR system from intake through closure. This is the theoretical base from which the CR process has been organized.

The framework for this paper proposes to cover three areas: (a) the development of the CRS and the CR process; (b) the structure of the CRS and the application format; and (c) the application of the CRS, not only as an evaluative tool but as a supervisory-training tool and a management tool.

Development of the Case Review Schedule and Case Review Process

Whenever a program evaluation procedure is put into effect, all parties must be involved. So when Region IX RCEP became involved in the CRS development, it was felt that right from the beginning there must be regional office and state VR agency input. Therefore, every revision of the CRS has been critiqued and/or reviewed by the states in Region IX. Further, the CRS has been field tested and applied in four of the state VR agencies in Region IX. Therefore, the CRS and the training that goes along with it have been developed jointly by the RSA regional office in Region IX, by the state VR agencies in Region IX, and by RCEP IX.

The original impetus for the development of the CRS was the VR program evaluation Standards. As the Standards were discussed with the states in Region IX, the general consensus was that there must be something developed which is more specific - some type of measure/process that will be responsive to the Standards and yield data on an ongoing basis. It should also assess the interaction between the rehabilitation counselor and the client.

The process chosen was the Case Review. The Case Review was the one process that was manageable. In no way was this considered a "cure all" to all questions that administrators and program evaluation personnel had in relation to the VR Standards. It was a starting point.

It was (is) an essential first step in developing a total program evaluation system.

The basic application of the CRS is evaluative. The intent is to evaluate the services provided by the counselor. The CRS can also be used for supervision. Evaluative data have limited application to making changes unless the results can be interpreted so that supervisors can assist rehabilitation counselors improve the type of services provided. Further, CR data are usable for some management decisions. What are the bases for the content of the CRS? The first basis is the Rehabilitation Act of 1973 and the subsequent amendments, the second is the Code of Federal Regulations (CFR), and the third is the Rehabilitation Services Manual (RSM). All items within the CRS come from one of these three source documents.

A recurrent statement has been made whenever the CRS is presented: "The size of the document is overwhelming." Another statement that will probably be made is that there is a duplication of items. The basis, though, that has been established for the CRS is compliance with the Rehabilitation Act, the CFR, and RSM. Some of the duplication and volume are a result of these sources. Therefore, a byproduct of the standardized application of the CRS on a national level would be to make recommendations to modify and/or eliminate questions that may not be essential and to minimize some of the redundancy of the three sources.

Structure of the Case Review Schedule

The structure of the CRS follows the rehabilitation process. Beginning with Section I on Client Information, the CRS deals with demographic characteristics, such as earnings at referral and closure, and length of time in status. Experience has shown that once this information has been gathered through the application of the CRS, it has been used by state and federal personnel in conjunction with the analysis of the results of the Case Review.

Section II, Evaluation of Rehabilitation Potential, is organized into two subsections according to the steps in the rehabilitation process. First, subsection IIA addresses the federal requirements for the preliminary diagnostic study. Second, subsection IIB contains the requirements for certification of acceptability for extended evaluation and the delivery of services within extended evaluation.

Section III, Eligibility, specifies the federal requirements for the determination of eligibility for VR services according to the components of the eligibility criteria. These include determination of a medically recognized physical or mental disability, substantial handicap to employment, and feasibility for employment.

Section IV, Evaluation of Rehabilitation Potential, contains the requirements for the thorough diagnostic study. The case documentation reviewed in this section is the evaluative and diagnostic information which the counselor gathers and uses to analyze the client's capabilities.

Section V addresses the IWRP requirements, including the issues of client involvement, client-counselor responsibilities, rehabilitation objectives and employment goals, and evaluation of the client's progress toward these objectives and goals.

The next section of the CRS, Section VI, deals with Delivery of Services. This section departs from the overall format of the CRS. Each service is evaluated in terms of whether or not the service is necessary, planned, and given to the client. Each service provided is evaluated in terms of its consistency with the evaluation and diagnostic study, with the IWRP, with employment at closure, and in terms of the quality of case management in the delivery of the service. The use of similar benefits and the economic needs test are also considered in relation to each appropriate service. Appropriateness in these areas is determined by the federal requirements and/or the individual state VR agency policy. This requires the integration of state policy into the case review training and the whole case review process.

Section VII, Termination of Cases, contains the procedural and justification requirements for closing a case in Statuses 08, 28, 30, or 26. Since this section is organized by subsections according to these statuses, the reviewer completes only the subsection which is appropriate to the status of the case at closure.

Finally, in Section VIII, the reviewer comments on the case which has just been reviewed or on the CR process. This provides the reviewer with the opportunity to note strengths or special concerns about the case record documentation which were not possible to indicate within the response format of the CRS.

One feature of the CRS which is important to discuss is its use as a conformance evaluation tool. Compliance and meeting the requirements set forth in the CRS and RSM have been discussed. The vast majority of questions in the CRS reflect these requirements and need to be completed for a conformance evaluation. There are some questions, however, which reflect federal "shoulds" rather than requirements, and these may be completed at the option of the review team. Whether the CRS questions denoting conformance are completed, or the total CRS is applied, is a policy determined at the beginning of the case review. The use of the CRS can be adapted according to the state VR agency evaluation needs and/or goals.

Medical consultation is an area in the CRS which does not reflect a conformance requirement and about which state VR agencies have the option to determine state policy. CRS questions pertaining to this area may be omitted from a conformance evaluation or they may be retained as part of the case review for additional feedback to state VR agencies regarding good practice.

"Skipping" is a structural feature of the CRS. The initial reaction to the CRS is that its length is extensive and, as a result, the time necessary to apply the CRS may outweigh the advantages of using such ar

evaluation instrument. In actuality, the number of pages completed in the CRS for each case is dependent on the status of the case being reviewed. For example, if a Status 26 case is being reviewed, 13 pages are completed. If a status 08 case is being reviewed, approximately two to three pages are completed. The total CRS is never completed for a single case.

The question-answer format of Sections II-V and VII basically involve two types of question-and-answer scales. A Type 1 question asks if the case record contains documentation; Type 2 questions ask about the quality of the counselor's documentation or the counselor's synthesis of client data. The area of counselor documentation based on counselor assessment is one that has been found to be exciting because it not only reflects the application of the CFR and RSM by the state VR agency, but it also focuses on counselor documentation skills and practices.

Specific Sections of the Case Review Schedule

To provide a perspective of specific sections of the CRS and the accompanying guidelines-instructions which are provided in the Case Review Manual, certain questions within two sections of the CRS will be highlighted. These two sections will be Section II on Eligibility and Section IV, which pertains to the Thorough Diagnostic Study.

Eligibility. The purpose of CRS Section III is to determine whether the client has met the eligibility requirements. To apply Section II to a case record, a reviewer assesses only the client data to which the counselor had access at the points of making decisions about the client's eligibility and the rehabilitation counselor's synthesis of the client data. The reviewer looks at the case documentation from the date of referral through the date of the eligibility determination.

The Eligibility section contains both types of question lead-ins and answer scales previously discussed: the Yes, No, N/A, and Less than Adequate, Adequate, N/A scales. An example of the Type 2 question-and-answer scale from the Eligibility section addresses the counselor's analysis of the specific ways in which certain factors contribute to the substantial handicap to employment.

Example:

How well does the counselor documentation in the case record...

5. analyze the specific ways in which the following factors, as appropriate to the client, impede the client's occupational performance by preventing the client from obtaining, retaining, or preparing for employment consistent with the client's capabilities and abilities?

   a. medical factors

   b. psychological factors

c. vocational factors

d. educational factors

The above factors are discussed in the following excerpt from the Case Review Manual: RSM Chapter 15.05.06 specified that the case record must include, as a minimum, the "agency's analysis showing the specific ways in which the medical, psychological, vocational, and other related factors impede the individual's occupational performance..."

The factors specified in 5.a-d pertain to:

a. medical-physical disabilities

b. psychological-mental and emotional disabilities

c. vocational-vocational adjustment including patterns of work behavior

d. educational-educational background in relation to vocational objectives, including preparation (e.g., training)

As the CR manual explains, CRS Section III, Question 5 requires the reviewer to make a judgment on specific criteria. In reference to the medical factors, adequate case record documentation, for example, must contain descriptive data such as medical or specialist reports, as well as counselor's synthesis of these data in terms of the impact on the client's occupational performance.

Question 6 in the Eligibility section also pertains to the substantial handicap to employment.

Example:

How well does the counselor documentation in the case record...

6. show that the related factors which bear upon successful vocational participation were considered?

The CR Manual (based on guidelines from the RSM) identifies, but does not limit the related factors to:

1. lack of marketable skills

2. low educational level

3. community and employer prejudices and attitudes concerning disability

4. Long-term unemployment

5. Unstable work record

6. poor attitudes toward work, family, and community

In applying the CRS to cases this is the type of information which a program evaluator would have access to, along with available references to the CFR, the RSM, and state policy, if applicable. The policies and procedures which are "state specific" are addressed in the Case Review training, since these policies may affect the application of certain CRS questions and the evaluation of documented case services.

Thorough Diagnostic Study. The second section of the CRS to be presented is Section IV, Thorough Diagnostic Study. This is the most important section of the CRS. The major professional difference between rehabilitation counselors and individuals called counselors in other agencies and settings, is the rehabilitation counselor's ability to synthesize information, an ability to articulate the information in a clear and concise treatment plan, and an ability to relate this information to the client's disability.

When the job duties of employment counselors, school counselors, etc. are examined, a significant difference between rehabilitation counselors and other counselors is the rehabilitation counselor's ability to complete a thorough diagnostic study. Unfortunately, it is also the area which presents the most difficulty. Diagnostic study is the area which needs the most attention for training and where supervision of case documentation seems to be most critical.

The purpose of the Thorough Diagnostic Study is "to determine the nature and scope of services." Rehabilitationists talk about developing an Individualized Written Rehabilitation Program (IWRP). Counselors cannot develop effective IWRP's unless they have done a thorough assessment of their clients. Unless counselors have good background information and have done a thorough analysis of that information and have articulated that analysis, the IWRP will be suspect.

The dates for the Thorough Diagnostic Study are from the point of referral through the IWRP development. It is all the information the rehabilitation counselor obtains from the time the client walks in: how the client sits and stands and dresses; the client's family relationships, prior work history, background; and the client's disability. The counselor often gets an "information overload" from the client. The counselor needs to assess the information and relate it to the client and the client's IWRP.

There are two answer scales in the Thorough Diagnostic Study. These are Type 1 - Yes, No, and N/A responses, and Type 2 - Less than Adequate, Adequate, and N/A responses. One of the questions using the Yes, No, N/A response format is:

Does the thorough diagnostic study...

3. contain the necessary specialty examination(s)?

If YES, does the Thorough Diagnostic Study document that the necessary specialty examination was given in the cases specified below:

a. a visual evaluation in all cases of visual impairment

b. a screening for hearing loss in all cases of blindness

c. an evaluation of the auditory system in all cases of hearing impairment

d. a psychological evaluation in all cases of mental retardation (including a valid test of intelligence and an assessment of social functioning, educational progress, and achievement)

d. other - please specify _____

The response to this question is based on whether the necessary information was in the case record. The CR manual instructions highlight the word "necessary." Was the examination that was provided a necessary examination for the blind?

Questions 8 through 16 use the Type 2 scales: "Less than Adequate, Adequate, and N/A." These questions raise the most concern in training and are of the most concern with regard to interpretation and application. These questions respond to the Describe and Appraise areas of the Thorough Diagnostic Study. Each of these questions is divided into point A and point B. Point A focuses on "to the degree necessary, how well do the data describe" and point B focuses on "to the degree necessary, how well does the counselor documentation appraise."

Point A focuses on how well the data describe. This does not require that the rehabilitation counselor generate all data. This information is descriptive, such as the medical examination, vocational examination, the psychological reports, etc. Basically, are there sufficient data in the case record? We would define "Less than Adequate" as being minimal or no data and "Adequate" as being test results, observations, reports, etc.

Point B for each of these items focuses on the rehabilitation counselor's appraisal. The difference between point A and point B is that point A focuses on all data, while point B focuses only on counselor documentation. This is the area where the rehabilitation counselor uses all the descriptive information in the case record. The rehabilitation counselor's documentation must show how the rehabilitation counselor takes the information in the case record, synthesizes the information, and documents how the information applies to the client and the client's vocational goal. The "Adequate" response relates the information, observations, etc. in the case record to the client and to the client's functioning.

One of the points stressed in the CR manual is "to the degree necessary." Obviously, the rehabilitation counselor could spend all his/her time gathering and analyzing information. The Thorough Diagnostic

Study is not judged by volume. It is judged by the quality of information. Therefore, one of the primary focuses of training in the CR process is to obtain consensus among all reviewers as to what is considered essential, what is important, and what is to the degree necessary? Training is critical in the use of the CRS. Further, it is critical that reviewers have actual field experience with the VR process so that they can relate to what happens during the day-to-day activities of the rehabilitation counselor.

Questions 8 and 9 are used to illustrate the Thorough Diagnostic Study. Question 8 is Personal Adjustment. The question asks if data describe the client's personal adjustment, and then if the counselor documentation appraises the client's personal adjustment. Personal adjustment, what does it mean? Personal adjustment means "the client's adjustment to his/her disability and his/her adjustment to significant others." In most cases, "Significant others" would be a husband, wife, or parent.

Question 9 is Vocational Adjustment. Descriptive data would include the client's background, work history, and type of adjustment toward a vocation. This adjustment would not necessarily be actual skills but adjustment to the work setting. In the CR manual, Vocational Adjustment addresses those factors considered essential to meeting the demands of work. "Does he/she get to work?" "Does he/she have trouble with supervisors?" Vocational Adjustment involves critical work-related behaviors. The descriptive data related to this question would be from vocational evaluation reports and contacts with prior employers. The appraisal aspect of this question identifies the implications for the client in relation to rehabilitation services.

The rationale has been provided underlying the information used in the Thorough Diagnostic Study. This section requires extensive training and orientation regarding criteria for responses and interpretation.

Case Review Manual

The CR manual contains the guidelines and instructions that each reviewer receives with the CRS. The CR manual is composed of five sections. The first section covers the training and provides an agenda outlining the training process which includes orientation to the CRS, CR manual instructions and participation in individual case reviews, and small and large group discussions.

The second section in the CR manual provides the guidelines and instructions for the CRS. We have highlighted, in the manual, examples of these instructions in relation to Section III.

The third section in the CR manual, the State Supplement, contains explanations, instructions, and references taken directly from state VR agency policies and procedures. The state-specific policy is provided in relation to those CRS questions where the state VR agency has the option of further defining the CFR and the RSM. A copy of the CRS is provided in the fourth section of the CR manual to serve as a reference

for the reviewer.

The fifth section of the CR manual provides a copy of the CFR which is one of the source documents for the CRS questions. The RSM is not included in this section of the CR manual due to the many chapters referenced to the RSM.

Sample Selection

Selection of the sample is the first stage in the application of the CRS. The selection of the sample revolves around the identification of certain sample characteristics. One characteristic that states have selected in the past is client disability. Another characteristic is case status. As an example, one of the state VR agencies in Region IX identified case status as a sample characteristic. They were interested in looking at the difference in documented case service practices between Status 26 and Status 28 closures.

Geographical origin represents a characteristic which has been identified consistently by state agencies in Region IX as a sample characteristic which needs to be considered. Since the resources may differ from one part of the state to another and employment opportunities vary geographically; case geographical origin has implications for case service practices.

Another characteristic is case funding source, which may include Supplemental Security Income, Social Security Disability Insurance, or Section 110 monies.

A possible sample characteristic is the dates for selection of cases from a certain period of time within a fiscal year. Finally, an option exists to select "other" characteristics. Other characteristics have not been used up to this point; however, that option is available.

Following identification of sample characteristics, it has been the role of the case review staff of RCEP IX to offer technical assistance to the state in the actual selection of the sample. Generally, the sample is selected from two characteristics by a proportionate random sample method. The percentage of cases chosen with each sample characteristic is selected according to the corresponding percentage of occurrence in the total number of state VR agency cases. Generally, the case review staff works closely with the states in the sample selection.

Administration

The administration of the case review process is a standardized process. The importance of this standardized process cannot be emphasized enough. It is one of the strengths of the CR process. Many evaluators can relate to the difficulty of using different case evaluation instruments within a district and then trying to draw conclusions or understand differences in case service practices among districts. Similar problems

occur on a state-wide or regional basis where different evaluation instruments and/or processes are used. The results differ but it is not possible to know the source of such discrepancies or if the results are reliable. Consequently, the case review team has focused on standardizing the process through training, controlling distribution of the cases, and introducing a reliability design.

Training consists of four parts. The first part of training requires five to six hours which focus on an orientation to the CRS sections and the corresponding instructions in the CR manual. Basically, the reviewer learns the CRS format and how to use the CRS and CR manual together.

In the second part of the training, each reviewer individually reviews the same state VR agency case and applies the CRS. The state VR agency case is selected for training according to the identified sample characteristics so that the reviewer will gain experience with cases which are representative of the sample of cases to be reviewed.

The third part of the training activity requires the reviewer to process the results of the individual review on a question-by-question basis in a small group. The focus of the discussion is on differences in interpretation, with the aim to develop consensus on the criteria to be used in the actual case review.

Finally, the fourth part of the training is to clarify any small group discrepancies in a large group discussion. This individual review and small group and large group processes are repeated with a second state agency case to further develop consistency among the reviewers in the criteria to be used for evaluation of case records. The development of consistency is an essential step in standardizing the CR process.

In addition to conducting the training, case review staff control the distribution of cases and monitor the flow of cases during the actual case review. Through a case distribution matrix, the administration of the review is standardized and integrated into a reliability design.

The reliability design allows the same cases to be distributed to all of the reviewers to determine the consistency in their responses to the same CRS question. The case review team has flexibility within the CR process to distribute six reliability cases among nine reviewers. As an example, if the review team is composed of a total of nine federal and state reviewers, the six reliability cases are distributed anonymously to each reviewer at different times. Therefore, a total number of 54 cases would be used as the basis to determine consistency among the reviewers.

Data Analysis

Data analysis is a critical stage in the application of the standardized CR process. Data analysis involves computer analysis, a standard format for presenting the raw data, and a preliminary data report.

As might be expected with a document as extensive as the CRS, the responses are keypunched and fed into a computer program, yielding a printout of the actual data. The CRS question is stated at the top of the computer data page and for each question in the CRS a page of data is printed out. Below each CRS question, the appropriate answer choices are given; either Yes or No, N/A, or Less than Adequate, Adequate, N/A. To the right, the frequencies of those answer choices are shown by total (T), state (S), and federal (F) reviewers. To the right on the printout, the frequencies are represented in percentages. The far right vertical column on the data page is a histogram which graphically represents the percentages of responses by T, S., and F according to the identified appropriate answer choices.

At the bottom of the data page is the printout of reliability information and results. The answer totals for the individual responses to each reliability case and the percentages are given. Reviewers are not identified by name, but they can be identified by state and federal affiliation to determine if there are any differences or patterns among groups of reviewers. Generally, it has been found that state reviewers are more critical than federal reviewers.

Reliability data are used to indicate a number of programmatic issues. In the early development of the CRS, the case review team looked at reliability data for indications of errors in the CRS questions. That is, were questions stated in such a way that they were difficult to understand or interpret? The case review team still looks for this kind of information, but also examined are indications of differences among reviewer groups or patterns in responses. If there is a lack of consistency among reviewer responses or an apparent pattern in responses with one to two reviewers, does this occur across one reliability case in particular? Is there strong consistency among reviewer responses?

Most recently, the case review team has been working with a systems analyst and a statistician to assist in developing a more appropriate measure. The new measure is a Kappa statistic and is based on majority agreement. This new statistical program is being field tested on data which were generated in one state agency to resolve any problems in the program before implementing it on a regular basis. The results of this Kappa statistic showed approximately a .87 reliability measure. This substantiated the basic statistic that had been used.

Preliminary Data Reports

Along with the standard printouts of actual data, the reliability results and measure of consistency, a preliminary data report is prepared. This report has been part of the developmental activities of the CR process.

The preliminary data report consists of a narrative which identifies the strengths and weaknesses reflected in the reliability results and highlights general trends in the unduplicated case review data. Currently, work is progressing on a reporting system which presents the data

in a graphic format by each CRS section. The data are presented across all of the sample characteristics for each CRS question. The guidelines for interpretation which are being proposed are used in conjunction with the graphic presentation of the data: (a) address the need for consideration of the sample size on a per question basis in interpreting the data, and (b) propose percentage categories for identifying the results as positive or as areas of concern and/or requiring some corrective action.

The intent is to present results which are usable to federal and state VR agency personnel. The interpretive process involves a brief written discussion based on the graphs and suggested guidelines, and the formulation of recommendations, if appropriate. Up to this point, the case review team role as trainers and consultants has not extended to the interpretation process. This part of the process has been the concern and responsibility of federal and state VR agency personnel.

Eligibility Guidelines

Finally, a recent addition has been suggested to the CRS in the Eligibility section. Concern has been expressed at the federal level about eligibility and the issue of clients being served who may not be eligible. In response to this concern, a procedure is built into the last revision of the CRS to have a case reviewed anonymously by more than one reviewer if eligibility is questioned by the initial reviewer. If a question of eligibility arises, the case is passed on to another reviewer anonymously to either substantiate the first reviewer's evaluation or to refute it. In some instances a case may be reviewed by more than two reviewers. If two or more reviewers agree that the client was not eligible and should not have been served, the case is pulled from the total review sample and this finding is reported to the federal and state VR agency personnel who are interpreting the data.

Case Review Schedule From the Management Perspective

The third part of this paper is a description of how the CRS can be used as a management tool. This includes data gathering, dissemination of information, and its application in a state agency as part of a management system.

When the CRS was being developed, it was felt that using it as a management tool would be appropriate. At this point, how a state agency integrates the CR data into existing management systems has not been clearly defined. Therefore, as a starting point, the following steps in the CR process have been developed to serve as a basis for its use as a management tool:

1. Awareness
2. Skills
3. Applied consistently

4. Applied systematically

5. Identification of training needs

6. Management needs

7. Follow-up

Step one is critical in the application of the CRS to develop an awareness on the part of the reviewers in the state. The awareness must focus on all sections of the CRS as to why we use it.

Step two in the application of the CRS as a management tool is to develop specific skills in the use of the instrument. Once the reviewer has the basic awareness of the CRS and CR manual, then he/she must develop actual skills in its application.

Step three is that the CRS and CR process must be applied consistently. There must be consistency among each reviewer.

Step four is that it must be applied systematically. We often take for granted the types of evaluation systems that are implemented and assume that they are going to be reliable, consistent, and accurate. This must periodically be brought into focus.

After step four and the CRS has been applied for the first time, the case review data will begin to identify trends in rehabilitation services. There will be some trends in services that are not desired or expected.

Step five is an identification of training needs. The data provide a basis for specific state VR agency training.

Step six is the area of management needs. Not all problems can be solved by training. Many changes need to be done through administrative and policy decisions. Therefore, step six provides an opportunity to analyze the data in relation to management needs. The application of data as a management tool facilitates training (step five) and management (step six). The data provide one form of assessment as well as an assessment for administrators and policy makers to make management changes, if appropriate. Management needs are not only specific to the state VR agency, the case review also yields information that could be provided to RSA with regard to management information needs.

Step seven provides for follow-up. It is essential that the CR process be continued so that longitudinal development is assessed. A critical part in the area of evaluation and its application is that it needs to be provided on a continuous basis. Continuous evaluation provides information which assesses how the agency has developed and how services have been provided over a specified period of time. These basic steps are critical in the application of the CRS and the CR process if it is going to be used as a management tool.

Nevada: A Model State System

How does the CR process work as a management tool? The following is an example of its use in the Nevada state agency. The administrator of the Nevada Rehabilitation Division decided that the agency would use the CR process with supervisors and administrators - not as part of a federal mandate, but as one element of the Nevada program evaluation system.

The first step was to select the sample characteristics. Nevada wanted to review each of their three district offices: (a) Carson City, which serves rural Nevada; (b) Reno; and (c) Las Vegas. Therefore, 36 cases were selected from the Carson City and Reno offices and 60 cases were selected from the Las Vegas office. These numbers were based on an approximate percentage of total cases for each district.

Nevada further wanted to examine selected case statuses. The statuses selected were: 08, 14, 16, 18, 26, and 28. The agency wanted to examine specific dates and time periods. These were open cases during the month of December 1977 and closed cases during the months of November and December, 1977. Finally, the agency wanted to select cases from each supervisor. There were three supervisors in Carson City and Reno and five supervisors in Las Vegas. From these characteristics, we selected the actual cases.

The second step in the Nevada training-review was to identify the procedure for the actual review. Table 1 illustrates that procedure. Instead of bringing everybody together with all the cases into one central area and completing the review, the case review team conducted the actual review in each district office. The supervisors and administrators were together for three days of training in Carson City. Over the next three-week period of time, the actual review was completed in each district office. The first week was in Carson City, the second week was in Reno, and the third week was in Las Vegas.

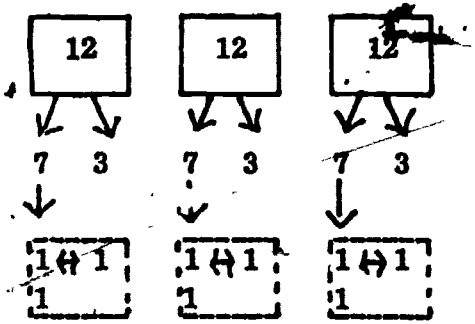
At each review site a staff person served as the facilitator. The process of the review required that each supervisor complete the review in his/her respective district office. Three administrators from the central office completed the review in each of the three district offices. To maintain consistency, the reviewers completed four reliability cases immediately after the three-day training.

The consistency ratio for each of the three supervisor groups and the administrators, as well as for the total group of reviewers, were calculated. The consistency ratio with regard to administrators immediately after training was .90; with supervisors across the three districts it was .88.

As reflected in Table 1, out of the twelve cases, seven were reviewed by a supervisor. Those are unduplicated cases. Three unduplicated cases were reviewed by administrators. That leaves two cases out of each of the supervisors in Carson City reviewed, and then out of those six, three (one from each supervisor) were reviewed by each of the administrators. In

Table 1

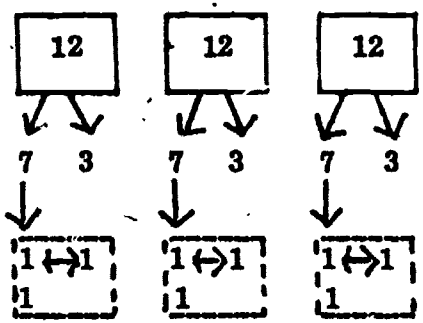
CARSON CITY



Each Supervisor — 13 cases – 3 supervisors x 13 = 39
Each Administrator — 6 cases – 3 administrator x 6 = 18

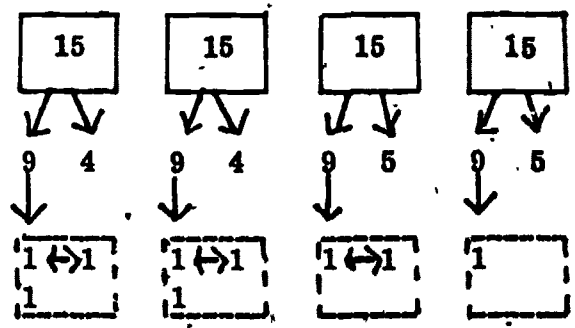
CRS Total = 57 for Carson City

RENO



CRS Total = 56 for Reno

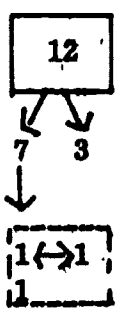
LAS VEGAS



Each Supervisor — 15 cases – 4 supervisors x 15 case = 60**
Each Administrator — 9 cases – 3 administrators x 9 cases = 27

CRS Total = 87 for Las Vegas

Explanation of Cases Reviewed By One Supervisor and One Member of Administrative Review Panel



(cases reviewed by supervisor -- unduplicated data)

(reliability cases reviewed by supervisor)

(cases rviewed by administrator -- unduplicated data)

(reliability case reviewed by administrator)

**60 cases were selected from 5 Las Vegas Supervisors but distributed across 4 Supervisors.

effect, a pyramid design was employed, allowing the supervisors who were closest to the rehabilitation counselor to do most of the review. Then the administrators reviewed a sample of the cases that the supervisors had done and completed their review.

The training was completed in February, 1978 and the actual review was completed in March, 1978. The final report will be completed in June, 1978.

The next step is analyzing the data. In the past, the interpretation has been done by federal and state VR agency personnel. In this instance, the case review team will be doing the interpretation and providing the state administrator and the program evaluation and training staff with specific recommendations.

If the process were carried one step further, the same procedure would be repeated with the same reviewers sometime between nine and twelve months later. The purpose would be (a) to see if there were any changes in case documentation, and (b) to provide some maintenance training.

The Nevada model, therefore, is one way in which to implement the CR process.

Summary

In summary, three objectives are indicated. The basic objective is that the CRS is an evaluative tool. The second is that the CRS can be a supervisory tool as used in Nevada. The third objective is that the CRS still needs to be tested as a management tool. The case review team will be focusing on this objective when the Nevada agency is given the final report.

CHAPTER VIII

REHABILITATION INDICATORS: A METHOD FOR ENHANCING ACCOUNTABILITY

Margaret Brown

Introduction

The Rehabilitation Indicators Project was first funded by the Rehabilitation Services Administration, Department of Health, Education, and Welfare, in October, 1974. The project, under the direction of Leonard Diller, Wilbert Fordyce, and Durand Jacobs, is housed at the Institute of Rehabilitation Medicine, New York University Medical Center, New York City.

During the first three years of the project, the conceptualization and initial development of Rehabilitation Indicators (RI's) were completed. To produce these results meetings were held involving over 75 rehabilitation professionals and consumers from across the United States. Present project activities are focused on initial testing of RI's in rehabilitation agencies and facilities. A demonstration phase has been planned to evaluate the extent to which several RI models of usage can address key purposes and needs within rehabilitation systems.

The Rehabilitation Indicators Project has viewed the development of a tool for program accountability as an extremely crucial process in which is it necessary to take into account the viewpoints of potential users of the tool. Not only is it important to get this viewpoint in the process of developing the tool, it is critical to take the tool out to field settings and see how the tool works. Project staff are currently in the process of shaking down this tool so that it looks like something where the bandaids aren't apparent. Then the tool will be taken out to a broad range of settings to find how it works. Until this lengthy process is further along the track the indicators will not be available for general use.

Rehabilitation Indicators constitute a new and basic tool, with much potential for use within a variety of rehabilitation settings and rehabilitation systems. Preceding the description of the nature of this tool, the needs and problems RI's were designed to address are enumerated and discussed.

Rehabilitation Indicators: Needs and Problems

Rehabilitation Indicators were developed in response to the continuing and long-standing difficulty of documenting the nature and scope of effective efforts within rehabilitation settings and systems. Adequate documentation is, of course, crucial both in terms of accountability (description of actual resources used in attaining expected client outcomes) and in terms of program evaluation (defining the impact of resources used).

Documentation of effective efforts is important to all participants in the rehabilitation process. Service providers receive adequate credit for their efforts, supervisors have valid information upon which to base supervision, and funding agents and Congress have a detailed picture to assist in their making resource allocation decisions. Also, clients get a clear understanding of their own goals and progress in order to enhance their participation in planning and monitoring their rehabilitation efforts.

Obstacles to Effective Documentation

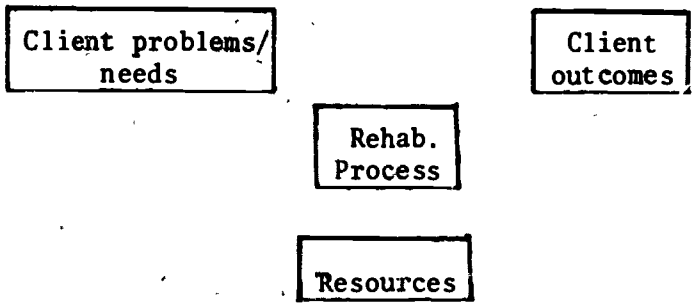
The difficulty in documenting effective efforts has four major components:

1. Difficulties in defining and measuring client change: What elements of the client's life are relevant in documenting change associated with rehabilitation? What approaches can be applied to measurement problems within the spectrum of relevant variables?

2. Difficulties in defining rehabilitation process variables that are associated with client change: What aspec of the rehabilitation system need to be documented? How?

3. Defining the range of resources used to attain rehabilitation outcomes: Rehabilitation is a process which over a lengthy period of time consumes resources of the client, the client's family, other support systems, public and private agencies, etc. In defining the resources relevant to documentation, problems exist in delineating the full range of direct and indirect resources and in determining the temporal relevance of resources to documentation. (For example, are funds used during early phases of medical rehabilitation part of the costs of later vocational rehabilitation?)

4. Difficulties in defining causal relationships within the rehabilitation process and resource variables defined above and depicted schematically below:



Problems here are numerous: What changes can be attributed to the rehabilitation effort and which are only associated with the effort. Which client needs are best addressed by which processes to produce specific desired outcomes? For what client problems are system efforts misplaced?

The Rehabilitation Indicators Project has focused its efforts on the first aspect of documentation: defining client change. Problems in documenting client change are discussed below.

Measuring/Assessing Client Change

Presently used systems of measurement are often inadequate to fully document changes in a client's life which are relevant to and associated with the rehabilitation process. In what ways is measurement inadequate?

First, measurement of client variables may be based on the premise that only a narrow range of the full spectrum of change is relevant to the documentation effort; e.g., only "employment status" is measured, with other important aspects of the client's life being ignored.

Second, when measurement is taken beyond "employment status," documentation may venture into areas of measurement that involve invisible changes where validity of measurement is difficult to determine. Measurement of "motivation," "self-concept," "self-esteem," etc., would fall into this category. Here documentation of change may be shown to be reliable, but the face validity of such measures is low (one doesn't know what has changed).

Third, measurement efforts may be misplaced in several ways. Two examples are discussed below:

1. Change may be effected in the client's environment; e.g., a ramp being built in the client's home. Instead of directly documenting this change in the environment, measures of the client may be substituted; e.g., a measure of the client's satisfaction with the rehabilitation system's performance.

2. Measurement of intra-client change may be substituted for direct measurement of behavioral outcomes; e.g., instead of measuring frequency of independent activities implemented by the client in day to day living, measures of activities of daily living (ADL) skills may be substituted.

Both ADL measures and indications of client satisfaction are of course useful for several purposes, but not as substitutes for documenting independent living activity patterns and environmental impact, respectively.

Misplaced measures derive from several possible courses: (a) the measurement may be traditional but not relevant for present use; (b) misplaced measures may be easier to apply than those that are correctly placed; (c) measures may be used only because they are available (as opposed to being relevant); and (d) limitations of the sensitivity of measurement instruments may not be sufficiently understood by the person selecting the measure.

The Rehabilitation Indicators Project has responded to these measurement problems with the development of a basic tool - RI's. However, the

development of RI's has also been affected by additional problems and needs within rehabilitation systems. Three additional problem areas associated with documentation of effective efforts are discussed below.

Additional Documentation Problems

Documentation activities are often simply added on to the day-to-day rehabilitation process and detract from, rather than complement service planning and monitoring. For example, measures designed to document effectiveness may be relevant to so narrow a range of client functioning that they contribute little information to the process of vocational goal selection, counseling, plan development, etc. Measures may not be applied to the variables of most concern to the client's and service provider's efforts; what is being "worked on" is not relevant to documentation measures and vice versa.

Whenever documentation is irrelevant in these ways, the time and effort the client and service provider devote to such measures will be wasted. Documentation efforts which are narrow and/or misplaced detract not only from time that could be used for other activities but also may distort the direction of planning and service provision. Such distortion occurs whenever service emphasis comes to mirror measurement demands rather than mirroring client needs (the latter defined within the context of system goals).

A second problem area is that of inflexibility in documentation systems. Measures may be adopted which reflect a single point of view and resultant data can answer questions formulated only within that perspective. The information generated by a documentation system should be sufficiently flexible to provide data relevant to information and accountability demands of clients, counselors, supervisors, administrators, evaluators, researchers, funding agencies, the Congress, etc.

A related problem area is defined by the need for documentation that may be easily communicated to different audiences. The language of many systems of documentation are too complex or jargonistic and translations are difficult or imprecise. Also, the level of measurement may be more refined than needed or understandable by a particular audience. Documenting effective efforts in the language of lay persons would serve to enhance accountability and would allow clients to participate more fully in the documentation process.

In summary, the Rehabilitation Indicators project has responded to a complex set of measurement and other problems associated with the need for documenting effective rehabilitation services. The Project has focused on problems in documenting client change with valid and sensitive measures that are useful within the processes of planning and monitoring services, and that provide information which is flexible and understandable. The project's response to this need has been the development of Rehabilitation Indicators.

Rehabilitation Indicators constitute a basic tool which responds to the needs defined above. Just as a tool such as a hammer has certain

characteristics (e.g., wooden handle), uses (e.g., hitting a nail), and potential purposes (e.g., building a house), so also do RI's. The characteristics, uses, and potential purposes of RI's are described below.

Characteristics of Rehabilitation Indicators

In brief, RI's are descriptors of observable elements of the client's life that can change during rehabilitation. Examples include: (a) walking up and down curbs; (b) being a homemaker; (c) frequency of social contact; (d) hours working/weeks; and (e) width of lavatory door in place of employment.

Descriptors of elements of the client's life that can change during rehabilitation are infinitely numerous and need to be sampled and organized before such a list of descriptors can be useful to the documentation of client problems and needs.

The RI list of sampled descriptors was developed by first applying the criteria of observability and meaningfulness. Non-observable elements (e.g., self-esteem, job satisfaction) were discarded, except those that could be directly operationalized in client behavior (e.g., "remembers own name"). Elements without direct functional meaning were also discarded; e.g., "flexes are" is a component of meaningful functional behaviors but is not meaningful in itself and was discarded.

This process of sampling from the list of descriptors does not imply that non-observable and/or non-meaningful elements are unimportant aspects of the clients; what is being implied is that documentation of change must be based on observability and meaningfulness to enhance validity and communication.

Further sampling was based upon a model of client change relevant to rehabilitation. The model defines sets of client variables that may change during rehabilitation and relationships between the variable sets; the model points to those classes of change which are crucial in the task of documenting effective efforts.

The first set of variables (person variables) are those that reside within the client and include beliefs, health, knowledge, motivation, skills, etc. Among these variable sets, only skills are observable and meaningful when directly operationalized in behavioral demonstrations.

The second set of variables in the model are environmental: physical elements (e.g., architectural barriers, availability of educational programs); social elements (e.g., employers' willingness to hire disabled persons); and personal elements (e.g., family support of vocational goals or independent living goals).

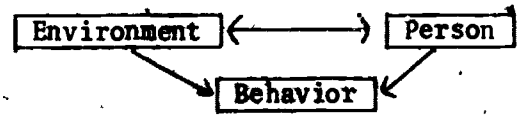
In the model, the third type of element that is potentially modifiable within rehabilitation is behavior. A client's behavior is a stream of actions tied to meanings and places; dividing the stream into meaningful units is an age-old problem within social science and also a

problem in developing an approach to measurement of client change.

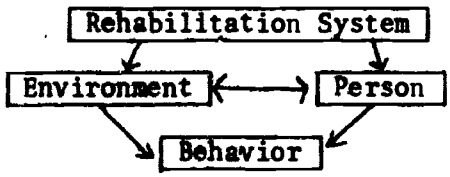
The problem of defining useful behavior units has four aspects. First, one can look at behaviors emitted by individuals within their natural context (e.g., "drives to work") or one can focus instead on behavioral demonstrations outside natural settings, in the laboratory or in a rehabilitation milieu (e.g., "walks ten feet, assisted by physical therapist"). Second, behavior can be categorized on a molar-molecular continuum. Thus, units can vary from small muscular movements to molar behaviors such as "dressing" to even larger units such as "attending college." Third, behavior can be categorized by context (or relevance to varying goals). Finally, one can choose to go beyond the surface and analyze molar behavior into factors or traits, e.g., aggressiveness, which are used to define units of study. However, the criteria of observability and meaningfulness point toward subdividing behavior into macro-units tied to rehabilitation-relevant goals (i.e., statuses, such as "being a homemaker" and activities, such as "eating a meal").

In summary, it is assumed that variables relevant to the client that may change during rehabilitation can be organized into three sets: person, environment, and behavior. Further, the client change model dictates a series of relationships between these sets. The relationships are based upon the concept that behavior is a function of person and environmental variables:

Further, it is clear that personal and environmental variables interrelate to each other directly, as well as to behavior:



Finally, if the rehabilitation system seeks to effect changes in the client's behavior (e.g., in employment status or in level of independent living), the model states that it must first effect changes in the client's environment and/or person variables:



This model of client change has helped organize the sample of descriptors that constitute RI's. The list of descriptors was sorted into four categories, reflecting each set of client variables in the model:

1. Observable, meaningful PERSON variables

   Skill Indicators

   Examples: "Walks on sidewalk"
   "Uses key to open door"
   "Uses verbs in correct tense"

2. Observable, meaningful ENVIRONMENTAL variables

   Environmental Indicators

   Examples: "Curb cuts"
   "Lavatory, in sports arena"
   "Unemployment rate, in area of residence"

3. Observable, meaningful BEHAVIORAL variables

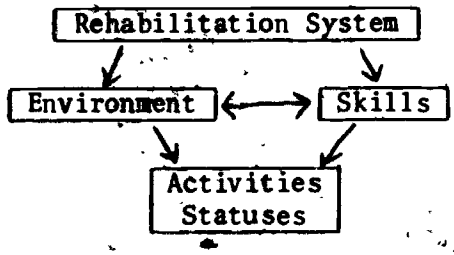
   Status Indicators

   Examples: "Client earned income"
   "Housing arrangement; hotel room"
   "Employment; homemaker"

   Activity Pattern Indicators

   Examples: "Buying a home"
   "Cleaning house"
   "Activities outside residence, duration"
   "Reading for pleasure"

The client change model suggests a logical arrangement or a hierarchy of client elements and an approach to effecting change among those elements. Behavioral change is pre-eminent in the model, with person and environmental change defined as instrumental in achieving behavior change.



In summary, rehabilitation indicators are constituted of four types of descriptors (skills, environment, statuses, activity patterns) of observable and meaningful client variables. Additionally, RI's can be characterized as being multi-dimensional and organized for flexibility of use.

In terms of multi-dimensionablity, within each type of indicators many areas or categories of client functioning have been sampled for development,

including vocational, educational, social, recreational, mobility, cognition, communication, self-care, independent living, etc.

RI's have been structurally organized for flexibility of use on the basis of four organizing strategies:

First, RI's are conceptually organized in terms of the four types of descriptors (skills, status, activity pattern, environment).

Second, within each type of RI, the descriptors are sorted into the content categories (e.g., vocational statuses, self-care statuses, etc.) mentioned above.

Third, the descriptors are being reviewed to obtain prior estimates of the relevance of each indicator to a variety of client disability categories: Which RI's are most relevant to psychiatric clients? Which to developmentally disabled clients? Which to spinal cord impaired patients? These estimates will be validated in empirical field testing. The result will be disability-relevant modules of RI's. The RI types and content categories may also be seen as bases for modularizing the total set of descriptors, to increase flexibility of use.

Fourth, RI descriptors have been organized into levels of detail. For example, within the "self-care" area of skill indicators, "eating/drinking" is an RI which constitutes a broad level of description. "Uses a fork and knife" and "butters bread" are two of the many detailed RI descriptors subsumed within the broad level. The organizing of RI's by level of detail of the content lends itself to the principle of indicators being used as "gates." If broadly defined content is judged to be irrelevant with a specific client, the refined content subsumed within may be skipped over in that instance. The "gating" operation is useful for saving time during data collection; the gate structure also lends itself to data reduction (from refined to broad categories) necessary to serve program evaluation purposes.

The final characteristic of RI's which needs to be stated is that of potential application: The RI content is relevant to many types of disabilities, including sensory, psychiatric, physical, and developmental. The flexibility and breadth of the RI content encourage use within many types of rehabilitation systems and settings, including state vocational rehabilitation (VR) agencies, vocational workshops, mental health centers, medical rehabilitation settings, etc.

To complete the discussion of the RI tool, the uses and potential purposes of RI's will be discussed, respectively, in the following two sections.

Uses of Rehabilitation Indicators

Indicators in general (e.g., social, economic) and RI's in particular have several basic uses:

First, RI's can be used to describe client variables at any one point in time. Examples: "employment status (at closure): sheltered employment;" "duration of passive reaction activities (at entry): 50 hours/week;" etc.

Second, RI's can be used to compare client variables at two or more points in time (i.e., assess change). Examples: "difference between duraction of socially isolated activities at entry into program and at present; eight hours/week decrease;" "decreased physical barriers in client's home: entry ramp;" etc.

Third, RI's can be used to state client goals and objectives. Examples: "employment status to be attained at closure: competitive/ wages;" "increased frequency of activities using a prosthesis."

Fourth, RI's can be used to describe problems, strengths, and needs. This usage combines description of the present level of a client variable and compares it to a standard or goal. Thus, "types 80 wpm" is a simple description taken out of context, but is a "strength" when compared to the goal of "employment status: competitive, secretarial position." Similarly, the fact that someone has not acquired the skill "washes dishes" may be termed a "problem" or "need" vis-a-vis the goal of the self-care status of "independent living."

The fifth use of RI's is to describe client outcomes. Here, RI's can be used to compare the level of a client variable at closure with a previously stated goal or objective. Thus, if "washes dishes" has been acquired as a skill and was a skill objective, the comparison defines a skill outcome.

Rehabilitation Indicators constitute value-free destriptors of client elements; values become attached to RI's when used to set goals, describe problems, etc. Such attributed values reside in the user and not in the RI's themselves. In other words, the relevance and value of each RI can only be determined in the context of each usage situation.

The multiple basic uses of the RI tool - description, assessment of change, goal formulation, problem/strength description, and documenting outcomes - constitute an additional dimension of RI flexibility. The high degree of flexibility of the RI tool and its basic uses allow RI's to be built into a wide variety of patterns or models of usage. The nature of such models will also be a function of the purposes which RI's are to address and the constraints of usage defined by the operational realities of rehabilitation settings or systems (see Figure 1). The remainder of this overview will discuss the purposes RI's can serve and the usage constraints that determine the design of usage models.

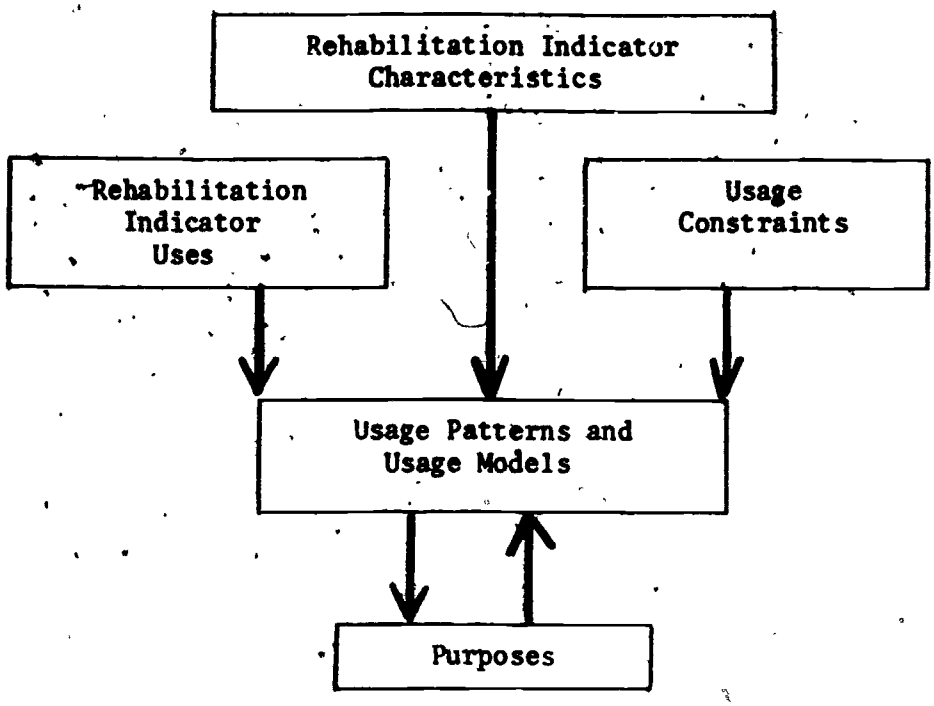


Figure 1: Model of RI Usage

Purposes Rehabilitation Indicators Can Address

Rehabilitation Indicators constitute a generic, descriptive language which is a basic tool for documenting client change. The following types of purposes can be addressed:

Enhancement of Program Evaluation and Program Planning

A key aspect of program evaluation is the documentation of the effectiveness of rehabilitation programs. In that RI's are descriptors of observable elements of the client's life, their use in documenting change enhances the face validity of evaluation data. Also, the defining and organizing characteristics of RI's and the multiple basic uses (description, change, goals, problems, strengths, and outcomes) allow flexible responses to a wide variety of evaluation approaches and foci.

With RI's, program evaluation can focus on an approximately wide or narrow range of content to fully reflect the rehabilitation program's areas of accountability - from vocational to independent living to an even more broadly-focused content. Finally, the RI tool's structure and organization allow for ease of data reduction to produce evaluation information suitable to differing audiences. For example, documentation can be reduced from detailed skill indicators to gross skill areas. Also, information can be obtained at relatively gross levels (status and activity pattern indicators) and/or at relatively specific levels (skills and environmental indicators), with data collapsed to appropriate levels, as program evaluation needs dictate.

Enhancement of Basic Research

The above discussion is also relevant to enhancement of basic research applied to rehabilitation clients and systems. Specific RI usage models will lend themselves to research analyzing causal relationships within rehabilitation processes. When RI usage models incorporate the client change model described earlier and document changes accordingly, they lend themselves to studies that in lieu of control groups allow partial attribution of client change to rehabilitation system efforts, narrowing ambiguities of interpreting results.

Thus, the model of service delivery and documentation rather than a research design based on random assignment of clients to treatment and non-treatment conditions forms the basis for exploring cause of outcome. The RI tool also has potential utility for addressing areas of rehabilitation research. These include:

1. Developing a method for functionally defining severity of disability
2. Providing a method for weighting case closures on the basis of difficulty of attaining goals, or some other functionally defined rationale
3. Providing a basic tool to be used in benefit/cost research (RI's defining multidimentional benefits)
4. Providing a basic tool for identifying the level of utility of specific rehabilitation inputs in contributing to client outcomes among specific client groups

Improve Service Quality

Many usage models will have a direct impact on quality of services; others will have only indirect effects. The direct effects will be referred to in this discussion.

Certain RI usage models will enhance client participation in service planning and service provision. Clients can contribute more fully through the generic language aspect of RI's and through a usage system that encourages clients to state their goals and problems in terms of observable RI descriptors.

Some usage models will assist counselors and other service providers by increasing their effectiveness in gathering and organizing relevant diagnostic data, interacting with clients in a common language, developing a plan of services which evidence high internal consistency (goals correlated with objectives which are then correlated with service) and in documenting results in a style that complements planning and monitoring.

Supervision of counselors could be enhanced within those usage systems that call for documentation closely paralleling the diagnosis-planning-implementation-monitoring process of counselor and client interactions.

The supervisor and counselor both benefit when the information base on which their interaction rests becomes a more valid reflection of service process.

Enhanced Resource Allocation Decisions

Through an information base that is more valid, multi-dimensional, and more relevant, decisions by the Congress and program administrators regarding allocation of resources should be improved. This purpose, thus, directly derives from enhancement of program evaluation and basic research. Resource allocation decisions will be improved to the extent such decisions are based on documentation of effectiveness in contrast to political and other sources of input.

Rehabilitation Indicators Usage Models

In designing a usage model that incorporates RI's to serve a specific purpose, the designer needs to consider several sets of constraints. These include (a) the subset of purposes being addressed, (b) constraints within the rehabilitation system, (c) the multiple basic uses of RI's, and (d) the characteristics of RI's.

In considering the purposes to be served, the designer of a usage model first needs to differentiate purposes that focus on directly enhancing service quality from those which do not. Whatever the purpose of RI usage, a model must then be designed so that RI's are integrated into the planning, provision, and monitoring processes of rehabilitation settings. In general, RI's could be integrated into the client-service provider interaction in several ways, some of which are described below:

1. RI's could provide a "standard list" of client elements (statuses, activities skills, and environmental elements) that may prove useful as a checklist of crucial sets of variables that need to be considered while developing a rehabilitation plan. These include gathering diagnostic data, organizing and translating diagnostic and other data into a form useful for communicating with the client and with referral agencies.

2. RI's can provide the behavioral, observable content from which suitable descriptors may be drawn and formulated as goals and objectives.

3. The selected RI's can be used as indicators of change, and of rehabilitation outcomes. The client's progress can be recorded using relevant RI's during the process of receiving services. This usage lends itself both to program evaluation and to enhancement of tracking and monitoring.

4. The selected RI's help focus the efforts of the counselor in contacting referral agencies to obtain diagnostic data and feedback regarding progress. Client-counselor interactions can also become more focused, enhancing client participation. Finally, follow-along services and follow-up studies can be focused onto

key client elements defined by RI's.

When integrated into rehabilitation processes, RI's will enhance the delivery of services. However, the RI tool will lend itself to many models of usage, including being used as a data collection device outside the planning, delivery, and monitoring process coordinated by the counselor. This latter model could be applied to a sample of clients or to all clients who enter a rehabilitation setting. Thus, a "data specialist" could collect RI data to be fed back (or not) into the client-counselor planning process and/or to be used for basic research and for program evaluation.

Usage models will be further modified by who is defined as the primary utilizer(s) of RI information: clients, counselors, or those outside the client-counselor interaction (supervisors, program evaluators, etc.). When more than one primary utilizer is defined, the usage model must work to obtain and provide more flexible RI information to suit the varying needs and communication capabilities of different groups.

Most usage models are focused on obtaining, encoding, recording, and utilizing RI information within the boundaries of rehabilitation settings. (A simple exception would consist of RI usage by "outsiders" to define and formulate system goals or standards in RI terms.) When usage occurs within a setting, setting constraints must be taken into account in designing the usage model. For example, presently used methods of information management and of service provision need to be clearly reviewed in developing an RI usage model, to complement present practices or to modify them in a less disruptive fashion. Time constraints within the setting also need to be considered; the benefits in information versus costs in time associated with various sized RI modules.

The type of client and type of setting will largely determine the specific method(s) of data collection that will provide reliable and valid RI information: the client, via oral or written self reports; staff or others, via observation; referral sources, through observation and reports; or a combination of all three sources. Finally, the purpose of RI usage and type of setting will largely determine the timing of data collection. This can be at entry, shortly after entry, at one or more points during rehabilitation service delivery, at exit (or closure), and after closure. Similar concerns will also dictate whether data collection is on a repeated measure, longitudinal basis, or on a cross-sectional basis, where different clients are measured at different points in time.

Usage as a Rehabilitation Indicator Function

The usage system design also will be a function of RI characteristics. A usage system may incorporate modules of one or more, or all types of RI's, broad or specific content areas, and disability relevant dimensions. This is displayed in Figure 2.

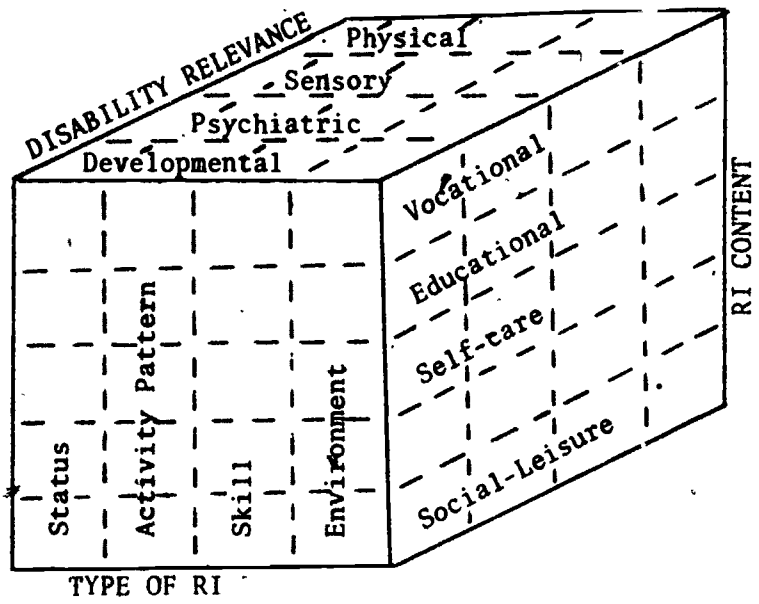


Figure 2. Potential RI moduler

The usage system also needs to address the RI characteristic of level of detail (or sensitivity) of RI's and select an appropriate level of measurement. This will depend on the purpose, who the user of the information is, and the source of the information.

The usage model must also address the basic uses of RI's. A usage model may incorporate pure description only; for example, describing client functioning levels at entry into a rehabilitation setting. A usage model may integrate only the goal setting aspect of RI use; for example, a comparison of client-determined goals to counselor-determined goals within a basic research study. Usage models may only focus on the description of change aspect of RI's; for example, a comparison of client employment status at closure with that at one year's follow-up.

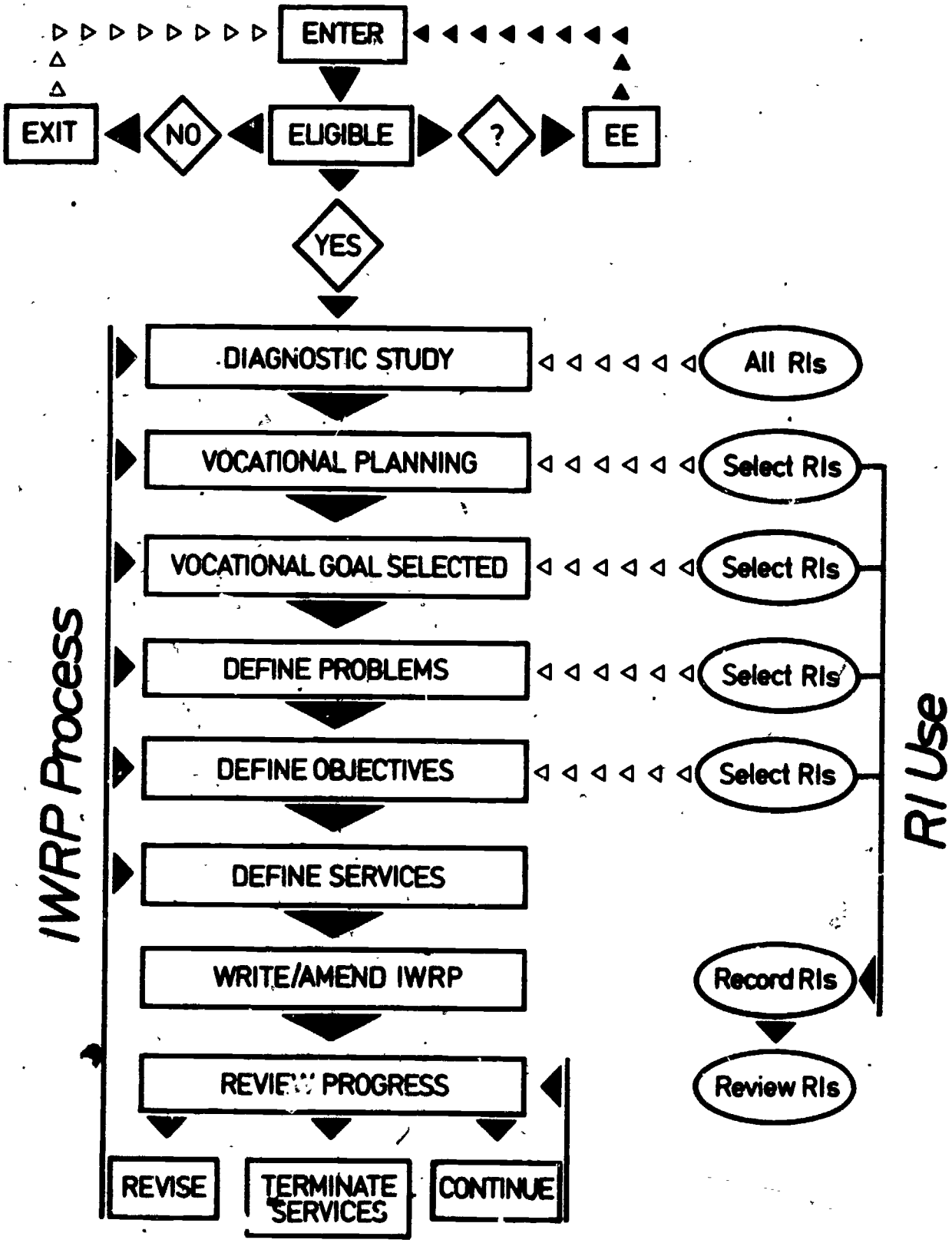
More complex usage models could incorporate more than one basic use; using RI description of relevant client variables to formulate RI goals and RI problems, as well as using RI's to monitor change, in terms of goal attainment or other types of outcomes. RI usage systems can be extremely simple or complex, depending on the purposes to be served, the users, sources of information, present systems, and setting constraints.

Rehabilitation Indicators can be implanted in the development of Individual Written Rehabilitation Programs (IWRP). In Appendix A is a schematic representation of the process. At the point of gathering diagnostic data, a counselor could use RI's as a way of organizing data and of making sure that the kind of data he/she is interested in obtaining from referral agencies, from the client, etc., is received. In this instance, RI's can be employed as a checklist and as an organizing mechanism.

When one gets into vocational counseling and vocational plan development, one way that RI's could be used is to selectively focus in on client

problem areas in terms of the kinds of RI variables that have been developed, defining first of all a goal or set of goals. "Competitive employment, earning wages" is an example of a goal (selected from status RI's) indicated on the second page of Appendix A. Thus, RI's can be used to state a goal. RI's can be used to state the problems the client has in reaching this goal. Four examples of problems in reaching the vocational and other goals are listed. There are also examples of RI objectives and examples of the kinds of services to address some of the problems.

Basically, RI's form a catalog, which in this usage model is used to help formulate goals and logically related objectives. Such goals and objectives are observable elements of client behaviors, skills, and the environment. These selected RI's are used to describe initial functioning and to track and assess client change during rehabilitation. Each selected RI can be used to describe a problem, establish goals, and indicate attained outcomes. Such usage should address many of the purposes described in this paper.


ENTER
EXIT
NO
ELIGIBLE
?
EE
YES
DIAGNOSTIC STUDY
All RIs
VOCATIONAL PLANNING
Select RIs
VOCATIONAL GOAL SELECTED
Select RIs
DEFINE PROBLEMS
Select RIs
DEFINE OBJECTIVES
Select RIs
DEFINE SERVICES
WRITE/AMEND IWRP
Record RIs
REVIEW PROGRESS
Review RIs
REVISE
TERMINATE SERVICES
CONTINUE
IWRP Process
RI Use

EXAMPLE OF REHABILITATION INDICATOR'S 121
BEING USED AS PART OF AN INDIVIDUAL WRITTEN REHABILITATION PROGRAM

VOCATIONAL GOAL : ----------S101 Competitive employment: wage/salary

Progress: (T_1)----------S116 Unemployed, not seeking work

(T_2)----------S113 Rehabilitation client

. .
. .
. .

(T_N)----------S101 Competitive employment: wage/salary

PROBLEM	(Activity patterns and Statuses)
1. Living with parents	S414 Parental family (don't support emp. goal)
2. Inadequate education	S305 H.S. diploma
3. Infrequent social contact	AP- increase social content (inadequate support system)
4. Dependent for self-care	A603 bathing (dependent)

RE PROBLEM # 1

OBJECTIVES--

1a. Acquire skills to establish independent household:
Skill RI's 46.01 - 46.10

Progress: (T_1)- Skills Null

(T_2)- Skills 46.01, 46.02 O.K.

. .
. .
. .

(T_N)- 46.01-46.09 O.K.; 46.10 Null

1b. Acquire Household skills:
Skill RI's 38.01, 39.03, 40.01-40.15; 41.01-41.08

Progress: (T_1)- Skills Null

(T_2)- Skills 40.01-40.15 O.K.

. .
. .
. .

(T_N)- ALL O.K.

SERVICES

RE Objective 1a: Counseling (by 5-27-79), VR Counselor
--Discuss skill objectives
--Discuss developing a better social support system

RE Objective 1b: -Counseling (by 6-30-79), VR counselor
- Johnson County Independent Living Skills Training Program (by 7-31-79)

CHAPTER IX

BENEFIT-COST ANALYSIS AS A PROGRAM EVALUATION TECHNIQUE FOR REHABILITATION AGENCIES

Robert M. Thrall
Larry Glass

Introduction

The question addressed in this paper is the usability of benefit-cost models for the state-federal rehabilitation program. The senior author is working on two projects sponsored by the Rehabilitation Services Administration (RSA). One of the projects, supported for several years by a grant to the Texas Institute of Rehabilitation and Research, is called "Analytic Aids for Research Project Selection" (AARPS). The second project, funded by a contract to Rehab Group, Inc., has a charge to study the applicability of benefit-cost models to the physical restoration aspect of rehabilitation.

Considerable skepticism has been encountered about benefit-cost, or as some people call it, cost-benefit analysis. Also, since there is a considerable amount of confusion in the terms used, it is important to clarify the meaning. There are three concepts: cost-effectiveness, benefit-cost difference, and benefit-cost ratio.

Understanding Cost-Effectiveness, Benefit-Cost Difference, and Benefit-Cost Ratio

Cost-Effectiveness. The term is used in the way it arose in the Defense Department. Suppose that the Defense Department anticipated the need for a new weapons system. Once this need was accepted, the first question would be: "What's the cheapest way to get the system?" The concept of cost-effectiveness formalizes this question. The cost-effectiveness approach either does not consider the benefits of the weapons system at all, or else the benefits are considered to be infinite. Thus, it is not necessary to set up any measures of benefits; just look for the cheapest way to achieve the goal.

Benefit-Cost Difference. For many years, benefit-cost difference has been used in public construction programs such as waterways, dams, and bridges, most of which come under the aegis of the U.S. Army Corp of Engineers. The first requirement is to conduct a benefit-cost analysis which estimates all the benefits and the costs. If there are no positive differences (if benefits do not exceed the costs) the program cannot go further.

A positive benefit-cost difference does not mean that the program is

accepted, but a positive difference is a necessary condition for acceptance. This seems fairly reasonable since many things in our lives are not done unless the benefits are expected to exceed the costs. Of course, if one is driving a car in the rain and the car's brakes fail, benefit-cost consideration is reduced to a question of whether to go over an embankment and get killed or crash into a wall and get a broken leg. Neither alternative is very pleasant, but action is taken so that the difference between the benefit and cost is the more favorable, even though it may be negative.

The requirement for a benefit-cost difference analysis may result in some "phony" statistics. For example, one discount rate may be used for benefits and another for costs; certain benefits may be overstated and some related costs may not even be included. However, since the analysis must be made available for public review, its deficiencies can be brought out in the open.

For example, a few years ago there was a proposal for an interbasic transfer of water from the mouth of the Mississippi River to the high plains of Texas to provide irrigation for cotton when the ground water was used up. Arguments based on contents of the benefit-cost statement were an important factor in a voted defeat of the proposal.

Benefit-Cost Ratio. The benefit-cost difference is quite appropriate for evaluating a single project. The benefit-cost ratio is applicable when there are limited resources, e.g., manpower or time, and several projects or alternatives are under consideration. For example, the final examination period at Rice University was modified and students were observed using the benefit-cost ratio. Students' limited resource was the amount of study time before the end of the examination period. Benefits were measured in terms of grades, and costs in terms of time. Each student wished to allocate time so as to maximize the total grade record. For each course, the student would estimate the grade change per hour of study, i.e., calculate a benefit-cost ratio, then give attention to courses in descending order of the ratios.

More generally, suppose that you have seven alternatives, A_1 - A_7, with respective benefits and costs B_1 - B_7 and C_1 - C_7, and have numbered them so that B_1/C_1, B_2/C_2.....B_7/C_7. The benefit-cost decision precedent is to select, in the order given as many of the alternatives as resources permit (provided that all selected alternatives have a positive benefit-cost difference, i.e., a ratio greater than 1). This procedure provides the most benefit to be obtained per unit of resource. The phrase "the biggest bang for the buck" is sometimes used to describe use of the benefit-cost ratio.

In summary: (a) cost-effectiveness refers to achieving an already selected objective in the best way possible; (b) the benefit-cost difference is an indicator of whether or not a course of action is desirable; and (c) the benefit-cost ratio applies when selecting one or more alternatives from a set of alternatives when you have a limited budget or some other limiting factor.

In this process some general principles of modeling are applicable. One principle is that for any kind of model to be effective it has to be logically sound. Models have been found in which two conditional probabilities were added rather than multipled; this is clearly not logically sound. It is the modeler's responsibility to see that the model is logically correct. That doesn't necessarily mean completely correct, because no worthy physical scientist will make such an assertion. Rather, the physical scientist says that if one assumes this, one can prove that. No physical scientist in physics, chemistry, or biology has been able to guarantee that assumptions are correct and certainly assumptions in the social sciences and in social situations are even more elusive. However, it is important to make sure that the analysis based on the assumptions is sound.

Scientific Completeness

It is important to make sure that a model has taken into account all the important and relevant factors. One way to achieve this is to overcomplicate the model and then discard factors which turn out to have little impact. One is more likely to get a sound simple model if one starts with a complex model and eliminates things than if one builds a model from the bottom up. It is easy to overlook some very important factors in the latter instance. Scientific completeness is an important challenge and charge in model-building.

Models have variables and constants. For example, fixed costs in rehabilitation would be constants. Flexible costs, or costs that change depending on who the vendor is would be variable costs. In general modeling, things must be measurable although some quantities are very difficult to measure. It is hard to measure how well you feel, but by contrast, measuring how many pennies there are in your pocketbook is extremely easy.

Most of the quantities in health-related models are hard to measure, but one must make sure that each model makes use of variables and constants that in some sense can be measured. The model builder is responsible for assuring measurement, but the help and cooperation of the people who are going to be using the model is needed. One of the ways help is enlisted is through Delphi-type exercises.

A model which comes from "on high" and descends on people already loaded with heavy day-to-day activities is not likely to be successful. To obtain staff acceptance and involvement, it is necessary to listen to the questions raised by the first person who tries to use the model and then be willing to reformulate it. A model that has just been designed and sent out from Washington, D.C. or somewhere else is not going to have a very good reception. The model needs to go through a stage of back and forth interchange. Everyone recognizes that utilization is a very important part of the whole process. Utilization will not be discussed in detail because RSA has been one of the pioneers in accepting the importance of utilization procedures.

Obstacles to Model Construction

There are constraints in building models. Some of these are:

1. Legal constraints. The laws under which rehabilitation operates say certain things shall be done and certain things may not be done.

2. Monetary constraints. Only a certain amount of money can be spent.

3. Personnel constraints. Only so many people are available to do the job.

All of these constraints have to be in the model in some form, or it is worthless. In practice, a great deal of the time, there is only one constraint that is binding - a principal constraint.

Suppose all of the possible rehabilitation benefits and costs can be counted. Take an example: Suppose a rehabilitation counselor was acquiring a wheelchair that would make it easier for someone with a spinal cord injury to be mobile and perhaps a gadget was being acquired so an individual could use a typewriter. If 10,000 people were going to use this type of wheelchair, the cost would include 10,000 times the cost of each wheelchair. The benefits would be 10,000 times the average benefit to each of these individuals. But there would also be the costs - capital type costs - of putting the chair into production in the first place and there might be some other benefits. The difference between all of the benefits and all of the costs is the expected net benefit.

Suppose there is some single scarce resource which turns out to be the limiting factor. Using an example of research project funding, the scarce resource is going to be manpower rather than dollars. Enough people are not available and authorizations for extra places are difficult. Whatever the scarce resource, the benefit-cost ratio puts the cost of the scarce item into the denominator. Thus, the expected net benefit is taken and divided by the amount of the scarce resource that is to be used; the higher this quotient, the more favorable the project.

Variables in Modeling

Probability of success, P_s. This depends on a number of components. These include the capabilities of the principal investigator and his staff, and whether the problem is feasible to work on at all. For example, if you're going into some type of cancer project, you might have the most skillful investigator available and make no progress because of the state of the art.

Probability of utilization if successful, P_u. The research may be successful, and (as happens with most academic research) the results published, the investigator gets a promotion or recognition, and the results just sit on the shelf. No one ever uses them and the investigator doesn't

even worry about whether the work can be used or not. That is one of the deficiencies of much research; there is no attention to utilization.

Individual benefits, B_I. This number is intended to cover benefits to a single recepient.

Size of impacted population, N. If there are N individuals in a class, each of whom has expected benefit B_I, then the expected benefit to the class is the product NB_I. For example if Project I deals with spinal cord disabilities with a class size of 100 and expected individual benefit B_{12}, then we must compare the products 100,000 B_I, and 100 B_{12} to appreciate the relative impacts. We expect B_{12} to be much greater than B_I, but is it enough to overcome the population size factor? How do you choose allocating your resources between these? This is the case where a legislative requirement exists that special attention be given to the severely disabled. A constraint such as this has to be taken into account in any model.

Non-personal benefits, B_s. Some benefits do not relate to individuals. Non-personal benefits are especially true where the contribution of research is to increase knowledge. The term B_s must, of course, be multipled by the probability of success.

Benefits of funding, whether the project is successful or not, B_f. Benefits of funding is politically oriented. Sometimes it is very important to just do something to show some sensitivity to a problem. An illustration can be made with a non-health-related example. Several years ago a great public outcry was heard that something should be done about gas distribution so people would not freeze in their homes. When the members of the House of Representatives would go out onto the capitol steps, reporters would meet them and say, "What are you doing about this now?" So the Congress introduced a bill saying something could be done, the president signed it, and everyone felt the problem was being handled. The chairman of the subcommittee involved in the House of Representatives reported that action had been taken, knowing that the action taken hadn't really changed anything.

Nothing in the bill existed that couldn't have been done before it was passed, but it was politically necessary to go through the process just to reassure the public that something was happening. In health care, there will be cases where some action has to be undertaken, knowing that it is likely to be unsuccessful because the lack of action results in political trouble. It is not being stated that it is a good thing in some general ethical sense to respond to these pressures, but since it is important to have a model that is realistic, a B_f term is needed. To summarize, one model takes the form:

$$B = P_s [P_u N B_I + B_s] + B_f$$

This is an interesting model which involves only six terms. How does one measure those individual terms? That is difficult. However, some ideas about the probability of utilization stem from research done under the sponsorship of RSA by Havelock and his colleagues at The University

of Michigan. Also, additional suggestions are provided by a study on the probability of rehabilitation success by a group in Minneapolis using retrospective analysis. How does one find the number of individuals impacted? This is a real stickler. Census data do not really apply to the precise data of interest. More research is needed in the area. If someone from Michigan is asked, "How many people in Michigan have severe spinal injuries?", does anyone know or believe that the number is known? The authors don't believe it is known in Texas, although one could make a pretty shrewd guess.

Interpreting Benefit Terms

The benefit terms in the model turn out to be very sticky, so much so that the AARPS group gave up on this form of the model because it turned out that the benefits cannot be measured be a single number. For example, the individual benefit term B_I cannot account for both economic benefits and quality of life benefits. Because of such difficulties, the AARPS model was revised into what is called a vector model, using a multidimensional utility method.

Another feature of the AARPS model is that each term requires a submodel. For example, the probability of utilization is the product of the probability of administrative utilization and the probability of individual utilization after it has been administratively accepted. Consider a possible new treatment for low back pain that one wishes to have implemented. First of all, at the state level, administrative action is needed: "We are going to implement this." Next, consideration must be given to the number of eligible individuals who will actually undergo the treatment. The population is not uniform; some people are impacted much more than others, so the population must be subdivided into a number of sub-populations, and for each of those sub-populations, different probabilities of utilization and different benefits will exist.

A basic problem is the different categories of data. There are hard data, which in this case is interpreted to mean economic data, and there are soft data, such as measures of the quality of life and its improvements. Coming from an academic background, a preference is seen for hard data which will always win out in academia. If a department chairperson takes a resume to the Dean for a person who did a lot of measurement work in a very difficult soft area, the Dean will look at it and smile sweetly. But when a resume is taken to the Dean for someone who has actually measured some physical constant, the Dean beams and says, "We should give this person a big raise." Academians love hard data, so models constructed in academia tend to emphasize hard data, despite the fact that in real life most decisions are made on the basis of soft data.

So, academia is inclined in one direction and real life in another, and then one wonders why models aren't better utilized. The merging of hard and soft data needs to be dealt with effectively. Soft data are terribly important but they tend to change rapidly with time. Soft data are evaluated differently from individual to individual, and agreement on single measures is not always expected. In fact, a Nobel

prize was awarded to Kenneth Arrow for his celebrated impossibility theory. Essentially what he said is that you cannot get any social measure which is universally acceptable.

Differentiating Among Value Profiles

One way out of the dilemma of the absence of universally accepted measures is to postpone questions of combining hard and soft data. Instead of trying to use a single number, as the original AARPS model attempted, vectors could be used. A vector refers to a collection of numbers to measure standings. A collection of numbers, however, may be evaluated differently by different individuals.

For example, consider the administrator of a health care facility. The administrator will probably be very sensitive to cost factors although patient pain relief is an objective; performance may also be appraised on the basis of how well the budget is balanced. Such an administrator may have an entirely different value profile from a researcher. Thus, three recognizable points of view may emerge: from the financial administrator, from the researcher, and from the client. Agreement among the three should not be expected. A utopia does not exist under which they are going to agree unless they're all dead; then the agreement doesn't mean anything!

If one wishes to use a vector measure it means a whole string of quantities. What is the financial impact? What is the quality of life impact? Is one number enough to measure quality of life? What about advancement of knowledge and political factors? The current AARPS model (for research prioritization) has settled on five dimensions. Each dimension should be narrow, thus reasonable agreement can be expected on its measurement. It should be possible to establish an index on which agreement may be obtained. Differences in agreement may relate to how the index should relate to other indices. If the indices are contracted for narrow measurement, the chance increases for obtaining reasonable agreement. On the other hand, the smallest number of components possible should be used that are compatible with the first requirement.

The "Analytic Aids for Research Project Selection" Model

The AARPS model for research prioritization involves several stages. First objectives are identified. After objectives are determined, it is necessary to make some kind of measurement of the progress of each objective. The measure may not be a simple measure, such as: "Now you're two units better off than you were before," because the objectives themselves are complex. Each objective may have quality of life components, economic components, and political components. These components are the dimensions of a multi-dimensional utility space and each objective must be scaled on each dimension. For whatever project or course of action under consideration, it is desirable to end up with a value in the form of a vector in the utility space.

At the final stage, the responsible decision maker synthesizes the terms used into a single number.

Other modelers have come to somewhat similar procedures. For example, Edwards and Snapper have something very similar, using the following process:

1. Identify the organization
2. Identify the issues (objectives)
3. Identify the entities to be evaluated
4. Identify the dimensions of value
5. Rank the dimensions in order of importance
6. Rate dimensions in importance
7. Combine the dimensions into a single scale

Edwards' and Snapper's work has been used fairly extensively in child welfare bureaus.

To summarize, a global model is desirable which, taking into account the objectives of the organization, will lead to measurement on a number of different dimensions. It should be realized that this global model is something which one can probably never use to generate numbers. The global model needs to be simplified for each individual application.

The following is a hypothetical example which contains three classes or benefits or objectives:

Patient Centered

1. Restoration of earning power
2. Relief of pain
3. General enhancement of the quality of life
4. Increased capacity for self care

Physician Centered

5. Increased scientific understanding of the basic medical problem
6. Improvement in methods of treatment

Facility Centered

7. Enhanced cost-effective handling of inpatient care
8. Enhanced cost-effective handling of outpatient care

Suppose three value or utility dimensions are considered - economic value, patient well-being value, and scientific value - and relative weights for the three value dimensions are determined by three classes of evaluators: patient, physican (researcher), and administrator. The patient looks at the economic value and says that it is worth 3 as compared with 7 for patient well-being and 2 for scientific value. (The patient may be concerned with the arthritis pain in his little finger.) The physician (researcher) puts weights of 3 each on economic value and patient well-being value but assigns a major weight of 6 on scientific value. The administrator assigns a big weight of 7 on economic value and puts weights of 3 and 2 respectively on patient well-being and scientific value.

It should not be believed that the administrator is wrong or the patient is wrong or that the doctor is wrong; agreement should not be expected. The difficulty with one-dimensional models is that they force people to agree when they don't.

The dilemma that is apparent provides the central point and reason for using multi-dimensional evaluation. The modeler should not develop weights in advance and say "We have to live with this." In other words, modelers should not prescribe weights for the value dimensions although modelers will, as part of their responsibility, obtain relative weights from a collection of representative evaluators. Values will be presented to the managers for use in determining their own weights. After all, managers will reign or fall by the decisions made. So if at one time managers say that economic value is the most important, they might give a weight of 0.5 on the administrator and 0.25 on each of the others.

For example, if some project was given values respectively of 7, 4, and 12 on the economic, patient well-being, and scientific utility dimensions, then the total benefit measure would be:

$$0.5 \quad (7 \times 7 + 4 \times 3 + 12 \times 2)$$

$$+0.25 \quad (7 \times 3 + 4 \times 3 + 12 \times 6)$$

$$+0.25 \quad (7 \times 3 + 4 \times 7 + 12 \times 2)$$

$$+0.5 \quad (85) + 0.25\ (105) + 0.25\ (73)$$

$$+87$$

If two weeks later the manager decided to place more emphasis on patient well-being, the weights might change from 0.5, 0.25, 0.25 to .3, .6, 1, and a (large) benefit measure of 95.8 would be attained for the same project.

Another sub-model is related to the problem of research project selection. RSA had advertised for proposals, each to be directed to what is called a "project concept." Each project concept described something

to be done in a request for bids. In response to each project concept, a number of project proposals were received. RSA wanted to evaluate these proposals using a peer review process. In its original form, each peer reviewer was given a list of 14 items, some relating to probability of success, some to probability of utilization, some to degree of realization of the objectives of the proj ct concept, and some dealing with financial management.

Each of the items was assigned a weight by RSA. The peer reviewer was instructed to check "zero" or "one" for each item for each proposal. The sum of the weights corresponding to the "ones" was the project score for that reviewer. The total project score was the sum for that project of those given by individual reviewers. All proposals for a given project concept were rated by the same reviewers (usually three or four) and each reviewer rated several project concepts.

A first reaction to this procedure was to propose a seven-point rating scale to replace the simple zero or one alternative. This modification was accepted and put into effect. However, some problems still existed with this evaluation model. First, the procedure involved adding "probability of success" to "probability of utilization," whereas probabilities should be multipled. Clearly, if a zero probability of success exists, it doesn't matter how useful the tester under consideration would be, the value is still zero, rather than zero plus something. In a like manner, the whole thing is worthless if it doesn't do anything for you, if it has nothing to do with your mission. In other words, these three numbers representing various probabilities need to be multipled and not added.

A detailed examination of the situation led to the conclusion that almost all of the multi-dimensional features of project evaluation were imbedded in the selection of project concepts and that for the proposals themselves, a simple six-term model, with some modifications, would suffice. Moreover, even in this simple model, terms common to all proposals to a given project concept might not need to be evaluated. The result was a model of the form:

$$B = P_s P_u P_r + FM$$

where P_s = probability of success, P_u = probability of utilization if successful, P_r = proportion of objectives expected to be realized if successful and utilized financial management. FM was scaled to give it appropriate total effect.

This simple multiplicative model had several uses. First, it was logically superior to the additive procedure being used, although the outputs were determined to be highly correlated. Next, a study of the model results suggested the hypothesis that "the face-to-face portion of the peer review process would have little net benefit." If this hypothesis was substantiated, substantial cost savings could be realized in the evaluation process, thereby freeing funds for more productive use in other areas.

Application Characteristics of Benefit-Cost Analysis

A model of benefit-cost analysis or a measure of program effectiveness has been discussed. The application characteristics of the model will be discussed in more detail. Benefit-cost approaches have been associated with vocational rehabilitation (VR) for quite so time. Although many techniques of benefit-cost analysis have been veloped in fields other than vocational rehabilitation, their pot for rehabilitation use has not been fully explored.

It would be difficult to directly apply a theoretically-complete benefit-cost model of vocational rehabilitation to a particular rehabilitation program. It would be even more difficult to develop a practical program evaluation model without an all-inclusive benefit cost model. During and after development of a complete model, sub-models may be derived which are immediately applicable to specific programs, issues, or populations. These sub-models do not try to describe or evaluate an entire service delivery system, but to answer specific questions. Because the questions will inherently take different forms, the sub-models will take different forms.

The important point about sub-models is that they will be determinably and logically related to the general model and to each other. Because this "model system" is initially derived from input by the users (in this case, program evaluators) and refined with additional input from the users, it will tend to be intrinsically more relevant and usable than many other approaches. Much of the resistance to benefit-cost modeling and modeling in general has resulted from the apparent arbitrariness of the model design strategy.

Often, the modelers came into a program evaluation effort with a rudimentary understanding of the system being evaluated and imposed a rigid and sometimes arbitrary framework onto it. The results were frequently uninterpretable and, when they were interpretable, were frequently off the mark. In the type of modeling presented in this part, the users define the benefits and costs and participate in the operationalization of them; thus, it is unlikely that they will not recognize the results.

Elements in the Benefit-Cost Model

The model system being proposed can be compared to a menu in a Chinese restaurant. One looks it over and knows pretty much what can be done in the kitchen without having to eat everything on the menu. Because the sub-models are built as quasi-independent modules (meals, to continue the analogy), when a decision is made to try one module for a particular question, one knows what one will get and, equally important, one knows what one won't get. This scheme allows benefit-cost analysis to be individualized for each state or program - if one doesn't need or want sweet and sour pork, one doesn't have to order it. For major evaluative efforts, the family plan also exists - one from Column A, two from Column B, etc. Each module or sub-model in the system would be available at a determined price to the user agencies depending on the extent and level of sophistication desired.

This model system approach permits the application of sophisticated analytical tools at minimal cost but with high specificity or resolution. The modular nature of the system implies that more uniform reporting of evaluative analyses to central agencies and more consistent reporting from year to year is possible than would be with a less unified approach.

Examples of Sub-Models

The sub-models fall into two general classes: descriptive models and decision models. The descriptive models are typically constructed for the purpose of providing information such as the characteristics of the population being served, the types of services being provided, and the current allocation of financial and human resources. The necessity for this type of information is obvious. The form which the information takes is usually straightforward but it is often more difficult to obtain than one might expect. The decision models tend to be somewhat more variable and complex and more often take the form of contrasts or comparisons between programs over a period of time.

In reality, the distinction between decision and descriptive models may be more frame of reference than analytical framework. Either class of model may be focused on populations, on programs, or on resources. Populations are most frequently described on the basis of disability chracteristics, demographic characteristics, or outcome. Resources are usually viewed as personnel, facilities, or economics, although time may also be considered in certain instances.

Input-Process-Outcome Model

Within each class of sub-models, there are a myriad of possible special cases, limited only by the users to define relevant questions or problems. One special case which has been applied extensively in both the descriptive and the decision models is the "process model" or, as it is sometimes referred to, the "input-process-outcome model."

This approach describes the universe of factors which apply to a client or client population in terms of input variables, process variables, and output variables. Input variables include demographic characteristics, health status at entry to the service milieu, work status at entry, and pre-onset factors, etc. Process variables are those which characterize the nature, quality, quantity, and costs of services and service delivery. Outcome variables include closure status, income, costs of maintenance, etc. From a research or experimental perspective, outcomes can be equated with dependent variables while the input and process variables are conceptualized as either independent or intervening variables, depending upon the focus of the model. Variables can be added or deleted at the discretion of the users. Another important feature of this type of model system is that the users operationalize the variables and specify the source.

Although process models have not often been associated with theoretically complete benefit-cost models and especially not with user-oriented

multi-criteria models, process models and certain other types of sub-models have been extensively used. They appear sporadically in research and more frequently in program evaluation. VR services have been evaluated with process models by a number of researchers. Vocational rehabilitation studies have been based on econometric process models with limitations to consideration of economic factors.

Summary

To reiterate, the theoretically complete multi-criteria model and the various sub-models which are derived from it are distinguishable on one or more levels from most other benefit-cost analyses in six primary ways:

1. The users - whether federal, state, direct providers, researchers, or consumers - define both the benefit and the cost factors according to their biases and needs.

2. Because the relative weights of the factor categories are determined by the users, they are explicitly stated and understood.

3. Evaluation costs are contained because users only buy or pay for what they need and because the [illegible] may often be collected from pre-existing or readily modifiab[illegible]ources - the R-300 for example.

4. The general model system is highly flexible yet consistent.

5. The models will help to indicate what type of data should be collected.

6. The operational models are directly related to each other and to the general model.

Several examples of the types of questions to which the approach we have been talking about is especially amenable include:

1. What are the characteristics of the population receiving VR services?

2. For which sub-populations are the services most effective?

3. What factors determine the effectiveness of services?

4. Why does Program A cost more than Program B?

5. Why is Program A more effective than Program B?

6. What is the most cost-effective mix of clients and services?

CHAPTER X

A NATIONAL FOLLOW-UP STUDY OF FORMER REHABILITATION CLIENTS

Sara Wagner

Introduction

In fiscal year 1978, the federal government contributed approximately $870,200,000 to the Vocational Rehabilitation (VR) program. A commitment of that amount, to be continued, justifies an assessment of the extent to which former VR clients have benefited from participation in the VR program. The current national VR follow-up study was designed to help make this assessment by collecting detailed information about client and program characteristics and the long-term health and economic benefits which have resulted from clients' participation in the VR program. The study was conducted by JWK International Corporation and Opinion Research Corporation.

During the early phases of the study, a number of methodological issues were raised directly related to state agency evaluation efforts. The following discussion, after identifying the study objectives and reviewing the program benefit measures, contrasts data collection objectives of state agencies with those of national data collection efforts. Three data collection strategies are described. The discussion concludes with a brief assessment of possibilities for cooperative state-federal collection of follow-up data.

Follow-up Survey Objectives

During the follow-up study, personal interviews were completed with a national probability sample of 6,000 former VR clients. The major objectives of the VR follow-up survey included the collection of information to assess clients' benefit retention and to evaluate indirect benefits or participation in the VR program. This information will be made available to the Rehabilitation Services Administration (RSA), to Congress, and constituents of the rehabilitation program.

While state agency follow-up studies are required to provide information to fulfill federal requirements and to provide information to state legislators, state agency studies often evaluate program strengths and weaknesses. Feedback from such studies to agency managers encourages program modification to improve agency performance. Data collected at the state level need not be complete enough for national program justification purposes.

VR Program Benefit Measures

An immediate concern when work on the survey began was the identification of VR program benefit measures. The selected measures were identified following consultations with VR personnel and a review of available information. The following categories of information were identified as relevant to the analysis of VR program benefits: (a) demographic and personal characteristics; (b) economic/vocational benefits; (c) health status; (d) client attitudes toward VR; and (e) indirect benefits of VR services.

An analysis of issues pertinent to each of these categories resulted in a series of questions addressed by the study:

1. What is the full range of benefits which emanate from the VR program, considering not only those which accrue to those clients successfully rehabilitated, but to those closed as not rehabilitated as well?

2. What is the retention rate of improved health status achieved through the program?

   a. How does health status relate to client characteristics?

   b. How does benefit retention differ in cases where a physical or mental impairment was corrected as opposed to those cases where the condition was only ameliorated?

   c. How does health status relate to services provided?

3. What are the relationships between different employment closures, program benefits, and client characteristics?

   a. How do these relationships vary over time?

4. In what ways are vocational training and other VR services effective?

   a. Do they lead to upgrading skills?

   b. Do they lead to changes in occupational standing?

   c. Do they lead to vertical mobility?

5. Is the client satisfied with VR services?

   a. Is the client satisfied with VR staff with whom there was contact?

   b. Would the client recommend VR services to a disabled friend?

6. What has been the impact of VR services on members of the client's family and friends?

An instrument was designed for use with former VR clients which examined a wide range of program benefits and identified the relation of benefit retention with other factors. The resulting lengthy and detailed questionnaire was reviewed several times by professionals from the VR community before being submitted for Office of Management and Budget approval. After initial pretests were completed, the decision was made to split some portions of the questionnaire into two forms to reduce respondent burden. Forms were also prepared for self-administration, with interviewer assistance, by deaf and inarticulate respondents.

Limitations of VR Standards Data Requirements

Since state agencies are required to collect follow-up data on former clients to meet the current VR program evaluation Standards, the need for the VR follow-up study has been questioned. However, studies conducted by state agencies and follow-up studies vary in the amount and type of information collected. State agencies are now required by the Standards to collect information about clients closed rehabilitated. However, a state could comply with federal Standards and provide only a fraction of the information sought by the VR follow-up study. For instance, Standard 6 specifies that information should be collected on the following items:

1. Percent of rehabilitated clients still employed at time of follow-up, specifying one year, two years, or three years after closure

2. Percent with earnings at follow-up, mean earnings at follow-up, and mean earnings for all with or without earnings at follow-up

3. Percent increase or decrease of earnings at closure to earnings at follow-up

4. Percent of rehabilitated clients (Status 26) unemployed at follow-up for: less than one month, one to three months, four to six months, seven to twelve months, more than twelve months

To insure that the client is satisfied with vocational rehabilitation services as developed with the counselor, Standard 8 specifies that the following information be collected:

1. Percent of clients rehabilitated throughout the fiscal year (Status 26) and not rehabilitated (Statuses 28 and 30) throughout the fiscal year who express satisfaction with the following, specifying one year, two years, or three years:

   a. Counselor's willingness to listen to client's ideas and suggestions in developing the individualized written rehabilitation program

b. Adequacy of information provided by the counselor to clients for understanding their disability

c. Promptness in the delivery of services

d. Kind of training received

e. Benefits of training received

f. Assistance in job seeking and final employment

g. Results of physical restoration services

2. Percentage of clients contacted during the follow-up period who stated they would recommend vocational rehabilitation to a disabled friend 1/

State agencies are not required to collect data about retention of any improvements in health status. Indirect benefits to family members often result from successful rehabilitation, yet data are not often collected by state agencies to evaluate such benefits. In addition, no data collection is required to fulfill Standard 6 information on unsuccessful closures. The VR follow-up study will collect this information from both successful and unsuccessful client closures and will permit comparisons of the outcomes of both types of client groups.

Sampling Issues

A two-stage stratified sampling plan was selected as the most efficient design for the survey. The 126 primary sampling units (Standard Metropolitan Statistical Areas [SMSA] and counties) include the 20 largest SMSA's and the six largest counties outside the SMSA's. The remaining counties were selected with probability proportional to size using 1970 census data as the measure of size. Since the VR follow-up study is a national study, cluster or area sampling was used to achieve high precision at reasonable cost. States have expressed concern that area sampling may produce unreliable state estimates, and this is true. However, simple random sampling or sampling based on a large number of clusters within each state would be prohibitive costwise at the national level, though feasible in some states for state follow-up surveys.

A total of 6,000 former VR clients were interviewed. These included:

1. 4,000 closed rehabilitated (Status 26)

2. 1,000 closed, other reasons, after initiation of the individualized written rehabilitation program (Status 28)

3. 1,000 closed, other reasons, before initiation of individualized written rehabilitation program (Status 30)

1/ Federal Register, Vol. 40, No. 245, Friday, December 19, 1975

Oversampling allowed interviewers to select only appropriate replacements for clients who were not available. Clients were selected to represent former VR clients six months, and one, two, three, and four years after closure. Appropriate statistical weighting procedures were used during data analysis to reflect the outcome of cases handled under the VR system.

Data Collection Methods

Three different types of data collection strategies should be evaluated in designing any follow-up survey. The most widely used methods for collecting data from hard-to-reach populations are mailed questionnaires, personal interviews, and telephone interviews. Each method has its strengths and weaknesses, just as each has its ardent supporters and detractors. The following section briefly reviews some of the advantages and disadvantages of each of the three methods appropriate for use in a survey of VR clients.

Mailed Questionnaires

Mailed questionnaires are used by many state agencies to collect information and are frequently used in large national surveys when budgets prohibit the use of personal interviews. Questionnaires require little skill to administer. They are either mailed or distributed directly to intended respondents. They are inexpensive to distribute over a wide geographic area and to widely divergent types of respondents.

Questions are posed uniformly to all respondents and the client is under no pressure to respond immediately. Questionnaires without identifiers pose little threat to the anonymity of the respondent, and can reach subgroups of the population not otherwise easily contacted. In an early experiment designed to test the validity of information provided using different data collection methods, Edwards (1957) found that information obtained using questionnaires that insured individual anonymity was more accurate than information obtained through personal interviews.

Mailed questionnaires are known to present significant problems. Mailed questionnaires seldom produce high completion rates. Thoughtful attention to design and format, careful pretesting, an adequate follow-up effort, and the use of monetary incentives may increase the response rate. All too often, survey conclusions are based on information provided by less than half of the target population sampled, with no information provided about the characteristics of non-respondents. The often used technique of comparing information provided by early respondents with that provided by late respondents fails to provide any information about the non-respondent group.

However, intensive efforts to reach non-respondents and interview them personally can sometimes provide adequate information to justify valid conclusions about the similarity of the respondent and non-respondent group. Mailed questionnaires cannot be answered by individuals who are illiterate, and are seldom answered by those with low levels of education.

They are unsuitable for use with blind clients and for those with disabilities that make writing difficult, since information provided on the questionnaire might reflect opinions of the surrogate respondent rather than the former VR client.

Personal Interviews

Personal interviews are considered by some researchers as the only valid method of data collection. A personal interview offers opportunities for greater care in communicating questions and in eliciting information. The interviewer is able to observe both the subject and the entire situation in which he or she is responding. Thus, this method is appropriate for acquiring information from persons who are illiterate (as at least ten percent of the adult population is today).

While people are often willing to speak fully when responding to open-ended questions during a personal interview, written responses to open-ended questions on questionnaires tend to be brief and incomplete. Surveys conducted by personal interviews have an additional advantage over other methods in that they yield a better sample of the general population since more people may respond when they are not required to provide length, written replies. Another advantage of the personal interview is its greater sensitivity to misunderstandings of those being interviewed.

For these reasons, the personal interview using suitable probes is considered a more appropriate technique for collecting information about complex subjects. Personal interviews may be administered to deaf members of the community through use of interviewers skilled in the use of sign language or through use of a trusted third person in the interview situation. Blind members of the community who would be forced to rely on others to complete a questionnaire can also participate directly in the personal interview. For example, a refusal that might be expected from a disabled client who writes only with great effort may be eliminated through use of personal interviews.

Finally, carefully trained, experienced interviewers (a crucial element in this method of data collection) can locate respondents often written off as "unable to locate" by the novice data collector. Skilled interviewers also develop the ability to gain entry into homes where inexperienced interviewers would be refused admittance. As a result, surveys using experienced interviewers as data collectors are seldom plagued by low response rates.

The disadvantages of this method of data collection are serious and cannot be overlooked by the evaluator planning to conduct a follow-up survey of former VR clients. Although skilled interviewers seldom command fees greatly exceeding the minimum wage, the cost per completed interview, including travel costs and the trips required for follow-up calls, may easily range from $50 to $100. The cost per interview, which will be determined by the length of the interview, the distances that must be traveled, and the difficulties encountered in locating and gaining

access to respondents, is still escalating and is rapidly become prohibitive. The unwillingness of many population groups to admit strangers into their homes has recently become a serious problem; this disadvantage can often be overcome by mailing intended respondents a notification of interview plans.

Response bias is more difficult to overcome in the personal interview situation. Respondents tend to provide favorable or socially acceptable replies when the interviewer is present, particularly when the interviewer is known to the respondent. Colombotos (1969), Hyman (1975), and Selltiz, Wrightsman, and Cook (1976) have provided excellent discussions of the effects of response sets on the validity of responses obtained during personal interviews. The tendency of inexperienced, poorly trained interviewers to interpret questions for respondents and to record responses inaccurately may also invalidate data. Invasion of privacy is often a concern of the respondent during a face-to-face interview. Since respondents have been located by name and address, they may fear that their personal identifiers will remain on completed forms, and that unfavorable responses may later affect their ability to obtain further services.

Although personal interviews do not appear to represent a feasible alternative for most VR state agency studies, they are the preferred method for use in large scale national surveys when a high response rate is required. Several large organizations maintain networks of highly skilled, experienced, mature interviewers who are able to interview locally without biasing the data.

Telephone Interviews

Telephone interviews combine advantages and disadvantages of both mailed questionnaires and face-to-face personal interviews. Several large survey organizations as well as many state VR organizational units are turning to this rapid method of data collection to avoid the prohibitive costs of personal interviewing. The cost of telephone interviewing is lower than for personal interviewing, interviewers can more readily be supervised, rapport with the respondent can frequently be established, personal admittance to the home is no longer a problem, and the threat to privacy is less problematic than with personal interviews.

The most often mentioned disadvantage of interviewing by telephone is the sampling bias introduced when 20 to 40 percent of the population either have no telephone or have telephones with unlisted numbers, or numbers listed under someone else's name. This problem, however, can be reduced in agencies where clients' telephone numbers are available. Interviews with clients should be short, and threatening questions avoided or placed at the end of the interview to prevent termination of the conversation. Since telephone interviews are difficult to complete with the majority of hard-of-hearing clients, face-to-face interviews should be planned for deaf or hard-of-hearing clients and only

highly skilled interviewers should be used with this client group.

Although the relative advantages and disadvantages of telephone interviews have been the topic of studies by Colombotos (1967), Sudman (1967), and Leuthold and Scheele (1971), the technique is just now gaining acceptance and additional controlled studies are needed in which the telephone approach is compared with other methods of collecting data.

Privacy Legislation

When conducting state agency studies in-house, evaluators may have relatively free access to individual client records. Such access is limited when conducting a national survey because of the compliance requirements of the Privacy Act, the Freedom of Information Act, and the Department of Health, Education, and Welfare's human subject regulations. To insure full compliance with these requirements, national survey interviewers must read the following paragraphs to each respondent:

> The information you give me in the interview will be turned over to the Department of Health, Education, and Welfare after all names and addresses and social security numbers have been discarded.
>
> We will destroy all of the questionnaires after making our report to the Rehabilitation Services Administration, that is, RSA. We will not report your identity to RSA with the information we give them. Unless required by law, neither we nor RSA will release any other information in a form which will permit anyone to identify you.
>
> Your answers are important in making the program more helpful to people who need rehabilitation services. Anything you tell me will be kept strictly confidential and we will never reveal your name or any information about you personally. However, please keep in mind that you may, if you choose, refuse to answer any question.

These statements assure that each participant will have an accurate statement of the confidentiality pledge and its limits. Under the Freedom of Information Act, an individual could request access to data tapes. Therefore, to further protect individuals, the phrase "unless required by law" is inserted. It is extremely unlikely that a law would be passed to require RSA to release data tapes to an individual.

To further protect the individual's anonymity, each questionnaire contains an arbitrary code number assigned by the contractor and none contain the interviewee's name, address, social security number, or any other personal identifier. A code sheet is used in the verification process which lists the assigned code number for each questionnaire, in

numerical order. Opposite each code number appears the name and address of the interviewee corresponding to that questionnaire. The supervisors select questionnaires for verification either randomly or on the basis of incomplete responses. They then use the code sheet to identify the appropriate interviewees and recontact them to check or to supplement their reported responses. There is no other code sheet (with or without code numbers) listing interviewees in order alphabetically, by address, or by any personal identifier. As soon as the contractor transfers needed reports and records, project staff destroy any records (including questionnaires) remaining in their possession which contain information on individuals.

The right to privacy is a fundamental right protected by the Constitution of the United States. The Privacy Act and the Freedom of Information Act also affect data collection procedures used in VR agencies and many states now have similar legislation regulating release of information.

Cooperative State-Federal Data Collection Efforts

Why not, then, compile data which are available as a result of state follow-up studies and avoid the need for this additional survey, a task that directly benefits few VR clients? Cooperative planning between state and federal agencies for such an effort is required to prevent inundation of either level with too many data of little value. Aggregation of essential data collected in each of the vocational rehabilitation organizational units is indeed an appealing prospect and would be possible if (a) uniform sampling procedures are followed, (b) data collection instruments are identical, (c) data collection techniques are identical, (d) all interviewers are provided with similar training and skilled supervision, and (e) careful control is exercised over transfer of data from the data collection instrument to the computer.

An analysis of several follow-up procedures currently in use reveals that none of these conditions is currently being met. At least one federal agency - the Administration on Aging - has initiated several contracts to recommend sampling procedures and to identify instruments that will permit collection of data that can be compared from agency to agency. The projects have proved more difficult than anticipated and the agency has not yet released any information related to this major effort.

The instruments prepared for use in the VR follow-up study have been suggested as possible data collection instruments for use in state follow-up studies. The questionnaire requests information that could be collected by skilled interviewers at the agency level to meet all of the requirements of a good follow-up study. The instrument was designed, however, for personal interviewing. It contains a number of show cards bearing response categories and its length and extensive use of skip patterns would prohibit use of the form as a mailed questionnaire. In its current form the instrument would not be suitable for telephone interviewing. However, it could be modified and adapted for use in this increasingly popular method of data collection.

A follow-up questionnaire specifically designed for use in state VR agencies will soon be released by RSA for review by potential users. The instrument could be modified for use in collecting data from former VR clients by means of any of the procedures outlined in this paper.

Summary

A review of one of the major data collection efforts in vocational rehabilitation, the VR follow-up study, has been presented. Differences between data collection efforts at the agency level and the national level have been discussed. The advantages and disadvantages of personal interviews, mailed questionnaires, and telephone interviews as data collection instruments for use in follow-up studies have been briefly summarized. The feasibility of aggregation of data across all agencies was considered and modifications of the instrument planned for use in the VR follow-up study was proposed as one step toward collecting comparable data in all the VR organizational units.

References

Colombotos, J. Personal versus telephone interviews: Effects on responses. Public Health Reports, 1969, 84 (9), 773-782.

Edwards, A. L. Techniques of attitude side construction. New York: Appleton-Century-Crofts, 1957.

Hyman, H. H., Cobb, W. J., Feldman, J., Hart, C. W., & Stimmer, C. H. Interviewing in social research. Chicago: University of Chicago Press, 1975.

Leuthold, D. A. & Scheele, R. Patterns of bias in samples based on telephone directories. Public Opinion Quarterly, 1971, 35, 249-257.

Selltic, C., Wrightsman, L. S., & Cook, S. W. Research methods in social relations. New York: Holt, Rinehart, and Winston, 1976.

Sudman, S. Reducing the cost of surveys. Chicago: Aldine-Atherton, 1967.

CHAPTER XI

COMPETENCY EVALUATION IN REHABILITATION (CEIR): TOWARD A COMPETENCY-BASED CLIENT-OUTCOME SYSTEM*

Don K. Harrison

A literature review suggests that rehabilitation program evaluation efforts are fragmented, scattered, and disorganized (Miller, Lee, Wargel, & Won, 1977). Although nine Standards for program evaluation have been established for the state-federal Vocational Rehabilitation (VR) program (U.S. Government, Code of Federal Regulations, 1975), an underlying conceptual framework is still needed for conducting comprehensive program evaluation, framing program evaluation questions, and guiding, developing, and organizing program evaluation research (Crystal, 1978). The case status system provides an indication of the type of services to be provided within each status, but it does not clearly specify the client outcomes intended (Hawrylúk, 1974) except employment (Status 26), an ultimate program objective.

The case status system focuses more on client processing and administrative tracking than on assessing competencies in terms of client performance during the rehabilitation process, upon which subsequent employment and adjustment may be contingent (Gay, Reagles, & Wright, 1971). Multiple client outcomes (Walls & Tseng, 1976) and intermediate client outcomes are given little or no attention. A comprehensive program evaluation system is needed which will provide for determining clients' needs at the initiation of service, measuring clients' achievement of intermediate objectives during the rehabilitation process, and assessing clients' gains at closure (employed and nonemployed) and thereafter. Such a system would provide information to guide management decisions related to programs and counseling developments related to individualized written rehabilitation plans (IWRP).

This paper discusses competency evaluation in rehabilitation as a proposed framework for rehabilitation program evaluation and covers the rationale, development criteria, a systems perspective of the VR program, and implications for future development.

Rationale and Assumptions

Although the concept of competency development has been discussed as to its efficacy for use in preservice preparation of rehabilitation counselors (Diamonti & Murphy, 1977; Anthony, Dell Orto, Lasky, Power, Shrey, & Spaniol, 1977), it has far wider potential and promise in rehabilitation, with staff other than counselors, and particularly with clients. Competency evaluation in rehabilitation (CEIR) assumes that the purpose of evaluation is to assess change in the competency level of a target

*Reprinted with permission from the Journal of Applied Rehabilitation Counseling.

population as a result of exposure to program and/or treatment conditions. CEIR assumes further that rehabilitation programs exist for the purpose of modifying one or more client competencies (behavior, knowledge, or attitude) for a target group exposed to services, training, or treatment. For example, rehabilitation counselors, after preservice education, are expected to demonstrate competencies in counseling skills at defined levels of proficiency with certain types of clients under described conditions. Rehabilitation supervisors, after inservice training in management-by-objectives, are expected to demonstrate management-by-objective skills at defined levels of proficiency with employees under prescribed conditions.

In a like manner, rehabilitation clients, after receiving rehabilitation services, should be expected to demonstrate competencies in predetermined areas at defined levels of proficiency at intermediate points during, at completion of, and at specific time intervals subsequent to the rehabilitation process.

Linking Client Change and Tasks Performed

The reference to client competencies in rehabilitation program evaluation centers on the intermediate and terminal behaviors of clients and on tasks performed by rehabilitation workers only as those tasks relate to client achievement of certain outcomes. For example, a rehabilitation counselor who, when asked by the supervisor to justify budget expenditures, reports on tasks (e.g., contacting employers, arranging psychological testings, interviewing clients, purchasing artificial appliances) is communicating what is being done with clients. The report does not specify what clients are accomplishing or what clients are able to do as a result of the tasks performed by the rehabilitation counselor. Several tasks and competencies have been identified for rehabilitation counselors (Wright & Frazer, 1975; Tripp, 1975; Muthard & Salomone, 1969; Harrison & Zawada, 1975), but empirical data are needed about the effects of rehabilitation tasks performed and their relationship to client outcomes (Rubin & Reagles, 1976; Bolton, 1978).

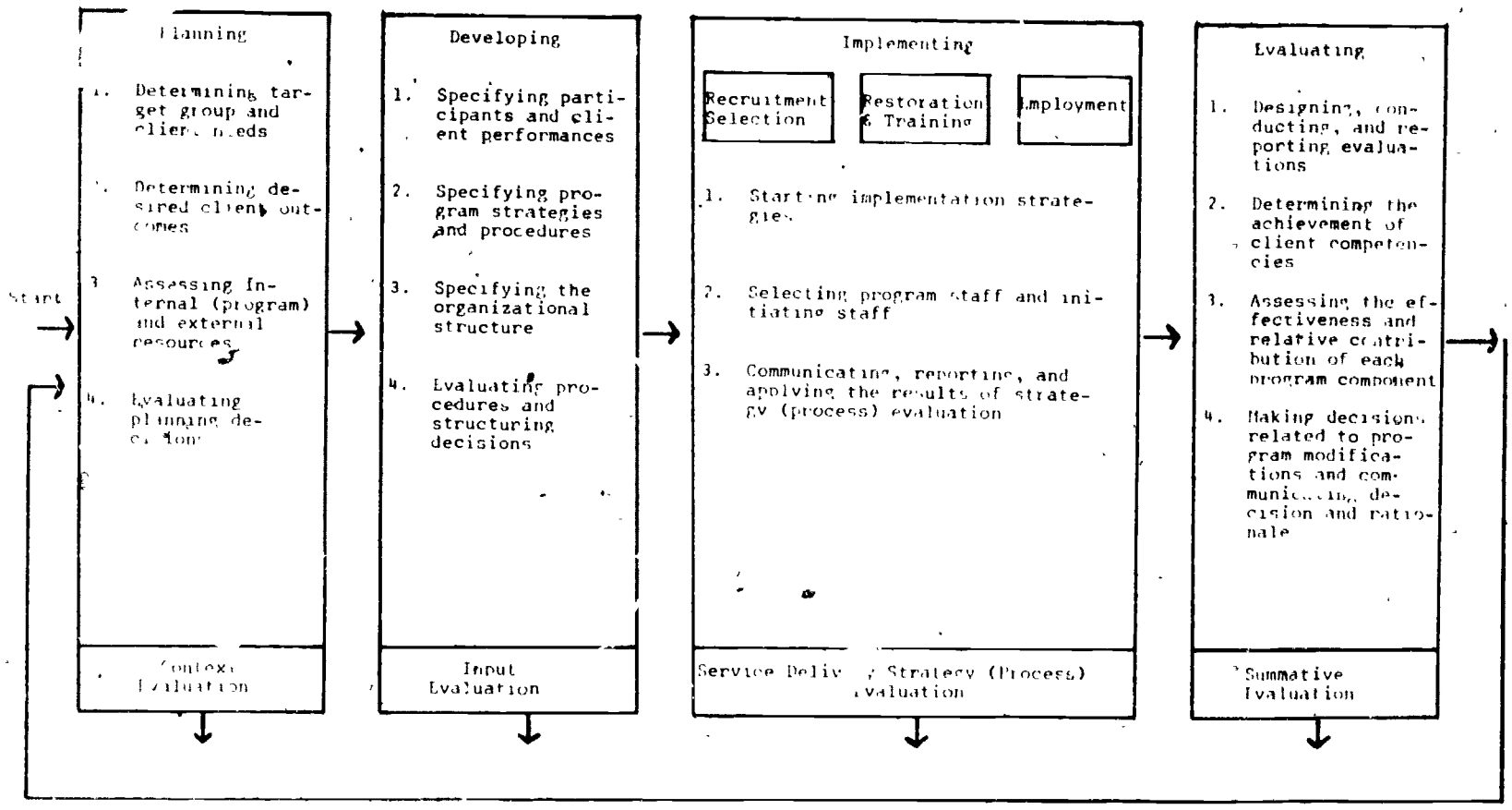
Research is needed to validate tasks of rehabilitation personnel to ascertain what rehabilitation worker competencies, when paired with specific client problems will produce what kind of client outcomes under specific conditions. When attention is focused on client competencies, an attempt is made to establish a "rehabilitation connection" between the tasks performed by rehabilitation personnel and the competencies achieved by clients.

The Cycle of Planning, Developing, Implementing, and Evaluating

A competency-based client-outcome evaluation system reviews program evaluation as an integral part of program planning, program development, and program implementation (Jones, Dayton, & Gillatt, 1977). Assessing changes in the competency levels of rehabilitation clients of the state-federal VR program suggests that (Figure 1):

Figure 1

A Comprehensive Approach to
Competency-based Client-outcome Evaluation
in
Rehabilitation



Adapted from Jones, H. B., Dayton, C., & Gellatt, H. B. (1977)

151

1. The desired competencies required by clients should be identified through a needs assessment evaluation which should form the basis for establishing the agency's goals (program planning)

2. The desired competencies required by clients should form the foundation for the design of service delivery strategies (processes) needed to achieve specific client outcomes (program development and design)

3. The desired competencies required by clients require that services and intervention strategies be provided timely and appropriately to achieve goals (program implementation)

4. The desired competencies required by clients, which should form the underlying rationale for the agency's goals, should be assessed by measuring the level of competencies achieved and/or maintained by clients: (a) at intermediate points during the rehabilitation process; (b) at completion of the rehabilitation process; and (c) at points subsequent to the rehabilitation process (program evaluation)

Assessing Client Needs to Develop Goals

Although goal oriented, the VR program has been criticized for its lack of specificity in statement of goals and objectives (Sussman, 1966), which has been considered almost mythical in nature (Spaniol, 1975). An investigation which tested the efficacy of an approach for implementing needs assessment evaluation to determine client-centered goals (Duguay, 1978) also highlighted the importance of identifying client changes which clients and other rehabilitation constituents desire from rehabilitation services.

Agency goals should be based upon an assessment of client needs. Alternative measurement instruments of rehabilitation service gains (First Institute on Rehabilitation Issues, 1974) are being developed, but they are limited since their development was lacking a broad concensus among the rehabilitation constituency of what should be measured, a problem which also exists between the agency and the client (Bolton, 1978).

A competency-based client-outcome program evaluation system requires a clear understanding of client needs so that program interventions can be evaluated in relationship to satisfaction of client needs, consistent with agency goals and mission. It is assumed that when client needs are met, need satisfaction may manifest itself in specific client behaviors. A primary goal of the state-federal VR program is to achieve client participation in gainful employment. However, with the passage of independent living rehabilitation legislation (Public Law 602, 95th Congress), other client needs upon which employment may be predicted (physical, social, and psychological) may begin to receive the attention deserved even though employment may not be the client objective. Although clients may or may

not achieve gainful employment, client gains in areas such as physical, social, and psychological functioning (Westerhide, Lenhart, & Miller, 1973) should be measured, properly attributed, and credited to rehabilitation intervention.

The specification of both vocational and non-vocational needs of clients, and the translation of these into multiple client outcomes, is a fundamental consideration in the evaluation of client competencies. Rehabilitation gain is broadened to include measurements in physical and affective dimensions; independent living rehabilitation is consistent with assisting clients to improve their non-vocational competencies (Reagles, Wright, & Butler, 1973).

Design Criteria

The system should provide information which can be used by counselors in developing and evaluating IWRP's and by agency managers and supervisors in assessing the needs of various disability groups, monitoring service delivery, and making program decisions. At least seven criteria which the CEIR system should address are reliability, validity, usability, flexibility, generality, ethicality, and contextuality. Most of these criteria have been addressed in taxonomy development in the related field of manpower training (Bates, Harrison, & Gordon, 1973).

Validity

The competency evaluation system should be based on goals of the VR program designed from a needs assessment of the constituency (clients, counselors, employers, etc.) of the rehabilitation program.

Reliability

The system should include criteria for evaluating clients reliably. Different rehabilitation workers, for example, should make the same assessment if the same rules are followed.

Usability

The evaluation system and its components should not require training, judgment, and resources beyond those that could be reasonably integrated into the VR program. This implies a preference for a conceptual structure that is as close as possible to the case status flow system to avoid having the evaluation scheme seem foreign and requiring additional and excessive work to implement. The system should be based on efficiency in terms of staff time and not require delays in clients receiving services to achieve goals.

Generality

The evaluation system should be comprehensive enough to provide evaluation data that would cover a range of client outcomes at various statuses during the rehabilitation process, at the end of the process, and subsequent thereto.

Ethicality

The evaluation system should not require information that would invade conditions of client privacy as a precondition beyond that which is normally permitted. Only minimal confidential information should be needed to be reliably and effectively used.

Contextuality

The evaluation system should make provisions for considering the impact of external factors in achieving goals in the VR program. Some context variables include socio-political conditions, financial resources, economic and labor market conditions, and the state of rehabilitation technology.

A Systems Perspective for Rehabilitation

In discussing measurement of rehabilitation outcomes, Walls and Tseng (1976) meaningfully present the VR program as an input-intervention-output system. A slight variation from that paradigm is to view the state-federal VR program from a systems perspective by using the Rehabilitation Services Administration (RSA) subgoals, which are related to direct service provision to clients, as the program system components. The system components are (Figure 2):

1. Recruitment and Selection Sub-system

   The process of outreach, referral, eligibility determination, and individualized client planning so that handicapped individuals receive appropriate services

2. Restoration and Training Sub-system

   The program content and use of physical and mental restoration, vocational training, and other supportive rehabilitation services

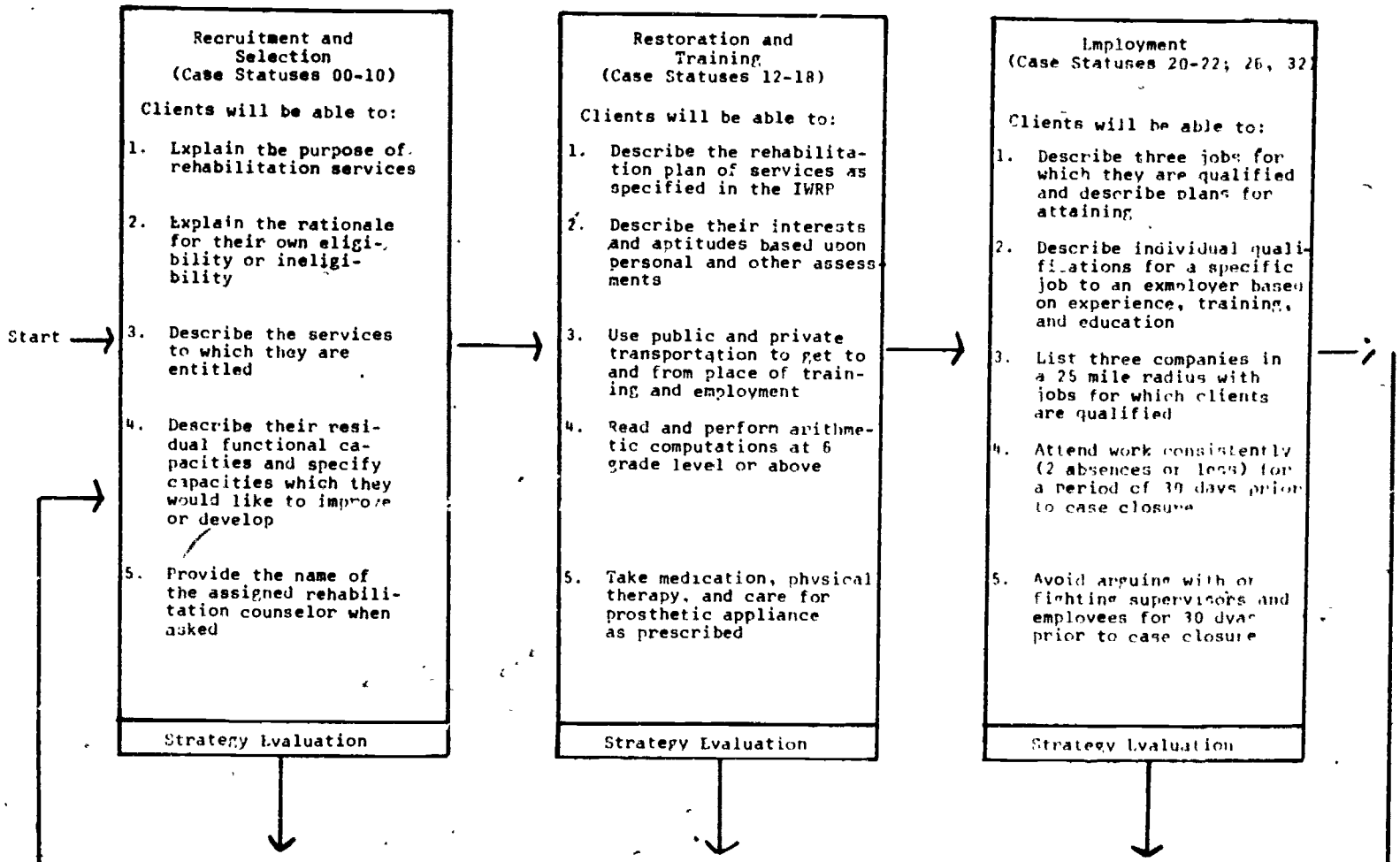
3. Employment Sub-system

   The process of job identification, placement, and follow-up support and assuming increased opportunities for optimal employment related to the present and potential job market, including homebound employment and the self-employed (Figure 2)

Each RSA subgoal, as a program system component, has the case statuses and examples of possible client outcomes to be achieved by each component. Each component has the function of achieving certain intermediate client outcomes for which it is designed. The lack of satisfactory achievement, or arrested competency development, at an earlier point in the system may result in non-achievement or inadequate achievement of competencies at subsequent levels. In terms of program

Figure 2

*REHABILITATION SYSTEM COMPONENTS



*Adapted and Based on RSA Subgoals

evaluation, some advantages of viewing the state-federal VR program from a systems perspective are provided.

1. Within each component, a checkpoint may be established and measures taken of the adequacy of that component in achieving the intended intermediate client outcomes. Thus, program components or intermediate outcomes may be redesigned or modified.

2. Relatively immediate feedback is available for use by managers for program component redesign. Clients may not move completely through the vocational rehabilitation system without information that certain competencies were not achieved earlier in the system.

3. Breaking the VR system into component parts, for evaluation purposes, may be more manageable since the relative contribution of each system component to rehabilitation may be ascertained.

4. More variance may be accounted for in evaluation research than is possible in using the terminal outcome of employment alone without intervening measurements. For instance, a considerable amount of time may pass between entry and exit of a client from the rehabilitation system, with many impacting intervening non-programmatic variables that are not explained. Under such conditions an explanation of the impact of the rehabilitation programming on client competencies and outcomes may be highly speculative.

5. Rehabilitation program evaluation is placed in a prospective dimension with the understanding that the objective of a designated program component is to develop specific client competencies. Poorly focused program goals and objectives may be avoided and program evaluators and managers may not have to retrospectively construct the intent of programs for purposes of evaluation.

6. Multiple measures of rehabilitation outcomes may be specified, and their achievement may be measured in relation to a specific program system component.

7. The establishment of intermediate client competencies for program system components is consistent with a rehabilitation philosophy of assisting disabled persons to achieve the highest level of independence of which they are capable.

Summary and Implications

The acceptance of competency evaluation in rehabilitation for the state-federal VR program suggests that rehabilitation may be viewed from a systems perspective, whereby system program components have a specific

function of achieving specific competencies in clients within described ranges and levels. Since evaluation is predicated on an assessment of client outcomes (knowledge, attitudes, and behaviors), an underlying assumption is that program goals will be determined as a result of clear understanding of client needs which are specified in terms of functional competencies to be achieved.

Although the goal of the VR program is client employment, this goal needs to be translated into muliple dimensional intermediate client competencies which are achieved by each program system component (in-process) and the terminal competencies to be attained at the conclusion of rehabilitation services (end of process). The need exists to identify (a) a group of intermediate client competencies for each system component, (b) a group of terminal competencies which clients should be able to functionally demonstrate upon termination of the rehabilitation process, and (c) a group of measures acceptable for assessing multiple client outcomes.

References

Anthony, W. A., Dell Orto, A. E., Lasky, R. G., Power, P. W., Shrey, D. E., & Spaniol, L. J. The realities of competency-based rehabilitation counselor education: A response to Diamonti and Murphy. Rehabilitation Counseling Bulletin, 1977, 21 (1), 58-62.

Bates, P., Harrison, D. K., & Gordon, J. A systems approach to a taxonomy of disadvantagement. Manpower Science Service, Ann Arbor, 1973.

[illegible]. Methodological problems in the evaluation of rehabilitation counselor performance. Rehabilitation Counseling Bulletin, 1978, 21, 190-193.

Crystal, R. M. A survey of the current status and program evaluation needs in the state-federal rehabilitation program. Michigan Studies in Rehabilitation, Series I, Monograph II, D. K. Harrison and Juliet V. Miller (Eds.). The University of Michigan Rehabilitation Research Institute, Ann Arbor, 1978.

Diamonti, M. C. & Murphy, S. T. Behavioral objectives and rehabilitation counselor education: A critique. Rehabilitation Counseling Bulletin, 1977, 21 (1), 51-57.

Duguay, A. R. A model for implementing a needs assessment evaluation to determine client centered goals and objectives of Michigan Bureau of Rehabilitation. Unpublished doctoral dissertation, The University of Michigan, 1978.

First Institute on Rehabilitation Issues. Measurement of outcomes. Institute, W. VA: Research and Training Center Press, 1974.

Gay, D. A., Reagles, K. W., & Wright, G. N. Rehabilitation client sustention: A longitudinal study. Wisconsin Studies in Vocational Rehabilitation. Monograph XVI. University of Wisconsin Regional Rehabilitation Research Institute, Madison, 1971.

Harrison, D. K. & Miller, J. V. (Eds.). Rehabilitation program evaluation: Problems, objectives, and projects. Proceedings of a participatory planning conference. The University of Michigan Rehabilitation Research Institute, Ann Arbor, 1977.

Harrison, D. K. & Zawada, M. A. Michigan competencies for rehabilitation counselors. The University of Michigan, Ann Arbor, 1975.

Hawryluk, A. Rehabilitation gain. Rehabilitation Literature, 1974, 35 (11), 322-341.

Jones, G. B., Dayton, C., & Gellatt, H. B. New methods for delivering human services. New York: Human Sciences Press, 1977.

Miller, J. V., Lee, C. C., Wargel, J., & Won, H. Program evaluation approaches for state rehabilitation agencies: Current status and future directions. Michigan Studies in Rehabilitation, Series I, Monograph I. D. K. Harrison and R. C. Riggs (Eds.). The University of Michigan Rehabilitation Research Institute, Ann Arbor, 1977.

Muthard J. E. & Salomone, P. R. The roles and functions of the rehabilitation counselor. Rehabilitation Counseling Bulletin, 1969, 13 (1), 81-168.

Reagles, K. W., Wright, G. N., & Butler, A. J. Human service scale. Human Service Systems, Madison, 1973.

Rubin, S. E. & Reagles, K. W. Assessment of counselor performance. In Bolton, B. (Ed.). Handbook of Measurement and Evaluation in Rehabilitation. Baltimore: University Park Press, 1976.

Spaniol, L. Program evaluation models for rehabilitation: A review of the literature. Wisconsin Studies in Vocational Rehabilitation, Series 3, Monograph XVIII, The University of Wisconsin Regional Rehabilitation Research Institute, Madison, 1975.

Sussman, M. B. Sociology and rehabilitation. Published by the American Sociological Association in cooperation with the Vocational Rehabilitation Administration, U.S. Department of Health, Education, and Welfare under VRA Grant No. RD-1684-G, 1966.

Tripp, J. E. Perceptions of state vocational rehabilitation counselors, supervisors, and administrators of core competencies for professional vocational rehabilitation counseling. Dissertation, University of Northern Colorado, 1975.

U.S. Government, Code of Federal Regulations. Revised general standards. Title 45, Chapter 13, Part 1370. Washington, DC, 1975.

United States Senate and House of Representatives. The rehabilitation, comprehensive services, and developmental disabilities amendments of 1978, Public Law 602, 95th Congress, 1978.

Walls, R. T. & Tseng, M. S. Measurement of client outcomes in rehabilitation. In Bolton, B. (Ed.) Handbook of Measurement and Evaluation in Rehabilitation. Baltimore: University Park Press, 1976.

Westerheide, W. J., Lenhart, L., & Miller, M. C. Field test of a services outcome measurement form: Case difficulty. Monograph II. Department of Institutions, Social and Rehabilitation Series, Oklahoma City, 1973.

Wright, G. N. & Grazer, R. T. Wisconsin studies in vocational rehabilitation: Task analysis for the evaluation, preparation, classification, and utilization of rehabilitation counselor-track personnel. Madison: University of Wisconsin, Regional Rehabilitation Research Institute, 1975.

CHAPTER XII

SINGLE-SUBJECT DESIGNS FOR CLIENT GROUPS: IMPLICATIONS FOR PROGRAM EVALUATION*

Kenneth W. Reagles
John O'Neill

Introduction

The Rehabilitation Act of 1973 mandated that the Secretary of HEW report annually to the Congress about the status of the state-federal vocational rehabilitation program, especially with regard to the primary intent of the act that gives priority to the severely handicapped. One impact of the act was a burgeoning of program evaluation efforts on the part of the individual state agencies. The capability to conduct such evaluations was lacking, however. Trained program evaluators, especially persons familiar with the Rehabilitation Act, were not available. In addition, methodological inadequacies have made creditable program evaluations difficult at best. As a result, program evaluation efforts and results have lagged far behind the intent of the act.

Perhaps one of the biggest dilemmas confronting program evaluators is the one which historically has been the nemesis of evaluation researchers. The dilemma concerns, on the one hand, the methodological desirability of having a control group, and on the other hand, the ethical problem of withholding assumably essential services to persons selected as "controls." One approach, typified by the efficacy studies in counseling and psychotherapy, has been to "delay" services to certain individuals in need of such services under the pretext that the agency was incapable of handling any more clients. The result was a quasi-control or comparison group. The methodological limitations of such an approach have been well documented (Campbell and Stanley, 1963), to say nothing of the persistent ethical problem. But without the use of control groups, program evaluators, like their evaluation research predecessors, will lack the necessary credibility to truly document the impact of the services they are attempting to evaluate.

Since the ethical consideration will no doubt persist, the escape route from the dilemma must necessarily be a methodological one. An approach that appears to hold considerable promise is the arrangement that allows a population, sample, or group of subjects to serve, in effect, as their own controls. The design considerations which relate to such a methodological feature emanate from the research surrounding the efficacy of behavioristic approaches to behavior change with single subjects. Recently, in the writings of Mitchell (1969), Guralnick (1973), Schmidt (1974), Thoresen and Anton (1974), and Miller and Warner (1975), the utility of such designs has been discussed, although the design features have existed for some time (Skinner and Ferster, 1957).

---

* Reprinted with permission from the Rehabilitation Counseling Bulletin.

The Designs

A number of research designs have been suggested which have particular intuitive appeal for use by either program evaluators or, more importantly, by rehabilitation counselors who are concerned with the progress of their clients. Additionally, the same designs lend themselves nicely to use in rehabilitation workshops, especially with the systematic monitoring of specific client behaviors in need of remediation and concomitant interventive alternatives. For each of the designs presented below, the design is first explained and examples are then provided to illustrate the use of the designs in rehabilitation settings.

The Withdrawal Design

The "withdrawal design" is characterized by first establishing a consistent level of the criterion, or a "baseline" (Leitenberg, 1973). The criterion may be the frequency of specific behavior or the rate of occurrence of some event. Treatment is then applied during the second stage of the design to affect the criterion which was measured during baseline sessions (see Figure 1). The third stage of the experimental procedure involves a withdrawal of the treatment condition and a return to the baseline procedure; the intent is for the criterion which was affected during the treatment stage to return to the baseline level. This is the initial validation of the effect of the treatment. The fourth and final stage of this design is to reinstate the treatment conditions with the intention of, once again, affecting the criterion. Evidence of replication may be considered as validation of the effect of the treatment, even though no control group has been used.

An example of how the withdrawal design might have been useful was in the 1955 Connecticut crackdown on speeding (Campbell, 1972). This well-read investigation was an example of an interrupted time-series design where baseline observations (traffic fatalities) were made over an extended period of time. A treatment condition (crackdown on speeding violators) was then instituted with the intention of affecting (decreasing) the number of traffic fatalities. According to the withdrawal design, a second baseline period would have been established by withdrawing the crackdown; this would have been followed by a second treatment (crackdown) period. If the same effect was produced, then the crackdown would be confirmed as the cause of the observed decrease in traffic fatalities.

As an example of how the withdrawal design may be used in a rehabilitation agency setting, let us imagine that an administrator of a certain agency suspects that the agency has a relatively high rate of client "dropouts," i.e., individuals who are closed as "not rehabilitated" for a variety of reasons. Such persons are referred to the agency, but do not reply to counselor correspondence and phone calls, or they fail to keep appointments with physicians, psychologists, and others who conduct pre-rehabilitation evaluations for the agency.

After talking to a few counselors and clients, the program evaluator suspects that clients do not fully understand the intent of the agency, nor their responsibilities in the process. The evaluator devises a standardized orientation to the agency consisting of an audio recording accompanying a series of 35-mm slides. While perfecting the orientation, a baseline rate of number of "dropouts" to number of referrals each week is established over a period of three months (see Figure 1). In the subsequent three-month period the new orientation procedure is used with all new referrals; the new rate is observed and recorded. To validate the impact of the new orientation procedure, it is not used for the clients entering the system in the next three-month period; the resultant rate of dropouts-to-referrals is observed and recorded. The new orientation procedure is implemented a second time, and the resultant change in the rate of dropouts is observed. If, indeed, the new orientation procedure is producing the desired effect, the orientation procedure is maintained as standard operating procedure.

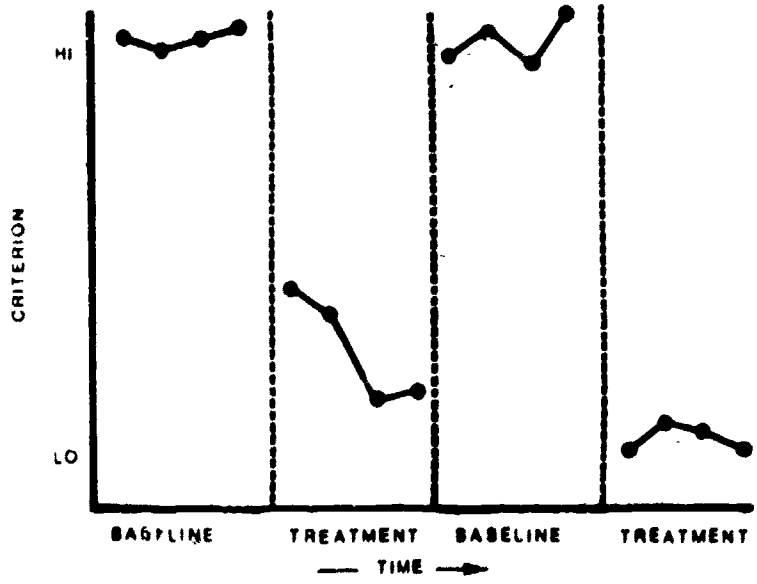


Figure 1. Withdrawal Design

The Reversal Design

The "reversal" design is similar to the withdrawal design in that there are four distinct phases to each. Each begins with the establishment of a baseline, followed by the initiation of treatment or intervention. In the third phase the treatment is withdrawn and a return to baseline occurs; in the final phase the intervention is reinstated. If the effect is desirable, it is continued as standard operating procedure. Unlike the withdrawal design, however, the reversal design (Leitenberg, 1973) is intended for two incompatible behaviors, one of which the

treatment procedure is intended to increase while the other decreases, for example, walking and sitting (see Figure 2).

Viewed within the context of a rehabilitation workshop setting, this design could be used to assess, for example, the effect of immediate feedback on the two incompatible behaviors of production rate and task distractibility (e.g., inappropriately leaving the worksite, disruptive talking to other workers, daydreaming, and the like). According to this design, base rates would be established for the criteria of "production rate" and "distractibility;" each worker would then be presented with an automatic counting device which would provide immediate visual feedback of their production rate. The intention of this treatment would be to increase production rate while decreasing distractibility; the resultant changes in the criteria are observed and recorded.

The third phase would be the removal of the counting device (the treatment) while observing and recording the changes in the criteria. The fourth and final phase consists of reinstituting the treatment procedure by reinstating the automatic counting device. If the effect observed during the second phase is replicated, the effectiveness of providing immediate and continuous feedback as a means of increasing production while decreasing distractibility would have been demonstrated, again without the use of a control group.

The reversal design lends itself to use in the rehabilitation agency setting and should be useful to program evaluators who are often confronted with situations in which two variables may be in opposition to one another. Consider an age-old debate: What is the relationship between the size of a counselor's caseload and the quality of rehabilitation services received by clients? Although there are a variety of criterion measures which might be selected to reflect the "quality of services," let us consider just one - client satisfaction. Using an instrument such as the Client Satisfaction Scale (Wright, Reagles, & Butler, 1969), the relative degree of client satisfaction may be measured.

Procedurally, a program evaluator would select a sample of rehabilitation counselors whose caseloads were considered "high," e.g., 250 - 300 clients at any one time. For a period of four months, the client satisfaction of all clients closed in two-week intervals by the counselors included in the study is measured. The results are recorded, as are the sizes of their respective caseloads in the same period of time. For the next four-month period, the sizes of the caseloads are reduced dramatically, for example, in half. Presumably, the client satisfaction would increase as the size of the caseloads decreased; the data would be collected and plotted to verify the assumption. Without a control procedure, however, one could not conclude unequivocally that the reduction in caseload size produced the desired increase in client satisfaction. Thus, much to the chagrin of the counselors, the caseload sizes are increased to their previous levels; the resultant change in client satisfaction is recorded. Again assuming the previous relationship of high caseloads to low satisfaction, the relative level of client satisfaction would be expected to decrease. The program evaluator collects such data for the four-month

period and then reinstitutes the low caseload situation. If the desired effect is replicated, documentation of the desirability of lower counselor caseloads may be assumed. The evaluator can then "fine tune" the system to determine the optimum caseload size relative to optimum client satisfaction within the constraints of available resources.

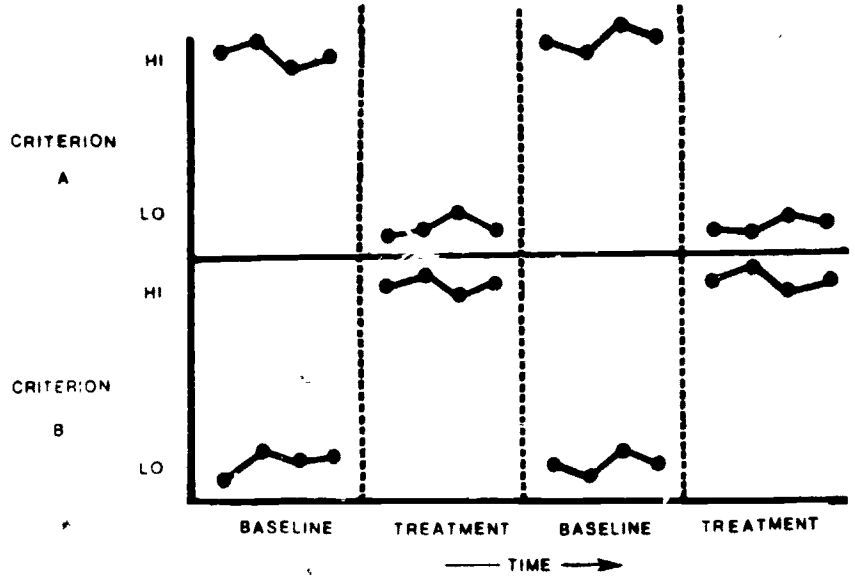


Figure 2. Reversal Design

The Multiple Baseline Design

In the multiple baseline design, baseline data are simultaneously collected on a number of independent behaviors. The same treatment or intervention procedure is then applied consecutively in three phases to each behavior (see Figure 3) while continuing to collect data on all the behaviors for which baselines were established (Leitenberg, 1973). The intent of this design is to demonstrate experimental control by either increasing or decreasing the rate or frequency of each behavior, while monitoring those behaviors not targeted for treatment. A necessary condition is, of course, for the behaviors or criteria to be functionally distinct or independent from each other.

Such a design could be used in the rehabilitation workshop setting to assess the effects of selective reinforcement of workers for (a) proper appearance, (b) punctuality, and (c) distractibility. According to the procedures suggested by the design, baseline observations would be established in the first phase for the targeted behaviors. In the second phase the first targeted behavior, proper appearance, is selectively reinforced; the resultant changes of appearance are noted and recorded, as

are the frequencies of the other targeted behaviors. In the third phase the second targeted behavior, punctuality, is selectively reinforced; the frequencies of desired reinforced behavior, punctuality, and the third targeted behavior are observed and recorded. The final phase of the design calls for reinforcement of the third targeted behavior, distractibility. The effect upon the frequency of the desired behavior is observed and recorded. The expected result is, as diagrammed in Figure 3, that there will be a systematic change in the desirable behavior as each is selected for reinforcement while the nonreinforced behaviors remain at baseline levels. If that occurs, it would be considered evidence of the impact of selective reinforcement.

An example of the utility of the multiple baseline design in a rehabilitation agency setting may be found in the instance of the relationship of client advocacy with case velocity, i.e., the rate of speed with which clients flow through the rehabilitation process. The Rehabilitation Services Administration's evaluation Standards specify three specific segments within the rehabilitation process in which the case velocity is of concern: (a) time from referral to acceptance; (b) time in extended evaluation; and (c) time from acceptance to successful closure (criteria A, B, and C, respectively, in Figure 3).

One of the assumed impacts of client advocacy is that clients will move through the rehabilitation process much more smoothly and, thus, the case velocity should increase. Procedurally, the base rate for each of the three rehabilitation process segments is determined. Client advocacy is then given to clients in referral status to the status of "acceptance" for a three-month period; the case velocities in all segments are observed and recorded. If the hypothesis is accurate, the velocity should increase for those clients receiving client advocacy services (criterion A) and not for the others. In the next three-month period the "treatment" is applied to group B (those clients in extended evaluation) and continued with group A if desired; the resultant rates are observed and recorded. Again, if the hypothesis is correct, criterion B should exhibit an increase in velocity, while the case velocity of criterion C remains relatively constant.

Finally, those clients who have been accepted and are receiving services (criterion C) are given advocacy services; the resultant change in case velocities of such clients is observed and recorded. One could conclude rather unequivocally that client advocacy services do indeed result in increased case velocity - clients move through the system faster. The conditions of the design have been met; the treatment (client advocacy) was the same in each instance and the criteria (case velocities) were demonstrated to be functionally distinct. Again, the need for a control group has been obviated.

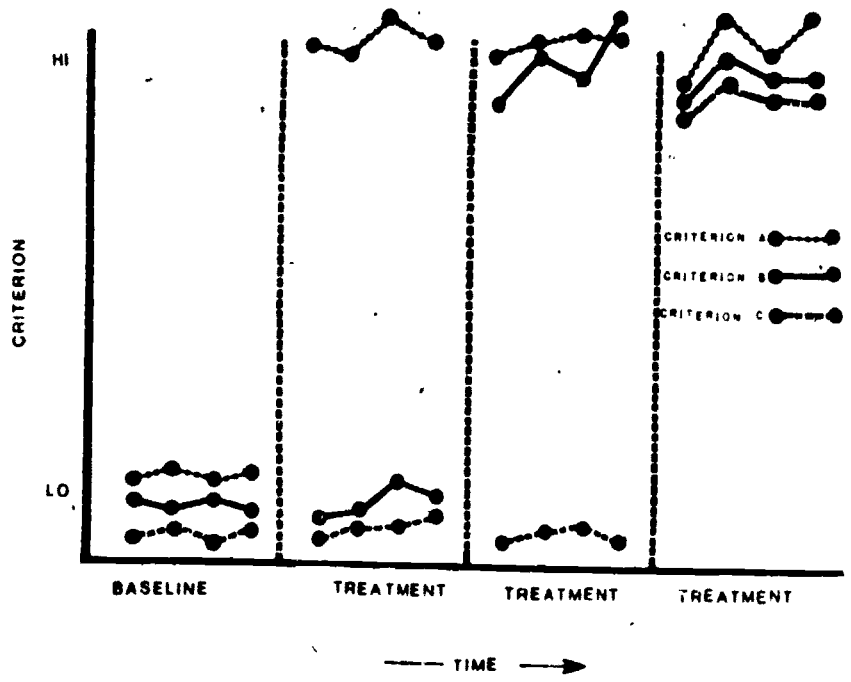


Figure 3. Multiple Criterion Design

Discussion

The single-subject designs offer new opportunities to enhance the quality of program evaluation in human service agency settings. The ability to establish causality with reasonable certainty without the use of traditional comparison or control groups is an attractive feature of these designs. They are not, however, without their own limitations, although such limitations are not serious when viewed relative to their unique strengths.

The first concern is related to an ethical issue. Although ethical concerns led to the identification of these designs, another ethical dilemma has emerged. The new ethical dilemma is especially apparent in the withdrawal design, wherein the withdrawal of some treatment condition will apparently result in some undesirable outcome. The example used by Campbell (1972) of the crackdown on highway speeding in Connecticut is useful to illustrate the dilemma. It would be ethically difficult to reinstate the condition of "soft" enforcement after "tough" law enforcement was apparently demonstrated to have resulted in a reduction of traffic fatalities. A similar parallel could be drawn for rehabilitation-related research or evaluation conditions. Reasonable certainty of effect can only be assured by replication, however, A potential escape from the new ethical dilemma is offered by the multiple baseline design in which it may be possible to obviate the withdrawal of apparently desirous treatments or conditions.

A second potential limitation concerns the interpretation of findings generated by such design strategies. The greatest potential threats to internal validity are the historial events which compete with the treatment conditions as an alternative explanation for observed differences. "History" has been defined by Campbell and Stanley (1963) as "the specific events occurring between the first and second measurement in addition to the experimental variable" (p. 5). Although the repeated observation of apparent causation may be revealed via methodologies using such designs, the treatment of systematic historical events cannot be ignored. The prudent evaluator will be aware of this design limitation and be sensitive to any competing events as potentially confounding of apparent findings.

A third limitation concerns the statistical analysis of data collected within the context of the single-subject designs in group circumstances. An assumption underlying many statistical techniques is the "independence of observations;" with the repeated measurement of the same individuals as is required with single-subject design conditions, one cannot assume such needed independence. Thus, there has been considerable debate in the literature concerning this issue. Gentile, Rodin, and Klein (1972) presented a case in defense of the use of statistics that require the assumption of independence of observation, on the basis that the violation of the assumption was not critical. Their case was rebutted rather dramatically by Kratochwill, Alden, DeMuth, Dawson, Panicucci, Arntson, McMurray, Hempstead, and Levin (1974); it appeared that the dilemma would persist and the value of the designs would be jeopardized by their statistical weaknesses. However, a recent article by Keselman and Leventhal (1974) offers a new insight. They have presented a model which accommodates mathematically the aforementioned limitation, especially with analyses of variance. Holtzman (1967) noted that the problem of dealing with multiple time series (i.e., collections of individuals) can be dealt with by mathematics (i.e., time-series analysis) which rely on modern computing facilities and, thus, are capable of resolution. Regardless, attention should be drawn to the dilemma and researchers and users of research are advised to keep abreast of current developments when considering sophisticated statistical analysis of data generated by these designs.

References

Campbell, D. T. & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Campbell, D. T. Reforms as experiments. In C. H. Weiss (Ed.) Evaluating Action Programs: Readings in Social Action and Education. Boston: Allyn and Bacon, 1972.

Gentile, J. R., Roden, A. H., & Klein, R. D. An analysis-of-variance model for the intra-subject replication design. Journal of Applied Behavioral Analysis, 1972, 5, 193-198.

Guralnick, M. J. A research-service model for support of handicapped children. Exceptional Children, 1973, 39, 277-282.

Holtzman, W. Statistical models for the study of change in the single case. In C. Harris (Ed.) Problems in Measuring Change. Madison: University of Wisconsin Press, 1967.

Keselman, H. J. & Leventhal, L. Concerning the statistical procedures enumerated by Gentile, et al.: Another perspective. Journal of Applied Behavior Analysis, 1974, 7, 643-646.

Kratochwill, T., Alden, K., Demuth, D., Dawson, D., Panicucci, C., Arntson, P., McMurray, N., Hempstead, J., & Levin, J. A further consideration in the application of an analysis-of-variance model for the intra-subject replication design. Journal of Applied Behavior Analysis, 1974, 7 (4), 629-634.

Leitenberg, H. The use of single-case methodology in psychotherapy research. Journal of Abnormal Psychology, 1973, 82, 87-101.

Miller, E. & Warner, R. W. Single subject research and evaluation. Personnel and Guidance Journal, 1975, 54, 130-133.

Mitchell, K. R. Repeated measures and the evaluation of change in the individual client during counseling. Journal of Counseling Psychology, 1969, 16, 522-527.

Rehabilitation Act of 1973 (Public Law 93-112), 93rd Congress, 2nd Session, H. R. 8070, September 29, 1973.

Schmidt, J. A. Research techniques for counselors: The multiple baseline. Personnel and Guidance Journal, 1974, 53, 200-206.

Skinner, B. F. & Ferster, C. B. Schedules of reinforcement. New York: Appleton-Century-Crofts, 1957.

Thoresen, C. E. & Anton, J. L. Intensive experimental research in counseling. Journal of Counseling Psychology, 1974, 21, 553-559.

Wright, G., Reagles, K., & Butler, A. Final project report: The Wood County project. Madison: The University of Wisconsin Regional Rehabilitation Research Institute, 1969.

CHAPTER XIII

A REVISED REHABILITATION SERVICES ADMINISTRATION MANAGEMENT INFORMATION SYSTEM

Charles B. Cole
Bruce A. Maloof
Ralph R. Turner

Background

This paper is part of the final product of a two-year effort leading to preliminary design of a "Comprehensive Management Information System for the State-Federal Vocational Rehabilitation Program." The project was initiated by the Rehabilitation Services Administration (RSA) in October, 1978. Abt Associates of Cambridge, Massachusetts has been the federal contractor responsible for this effort. This report includes activities undertaken during Phase I of the project.

The project's background includes a series of RSA-funded studies, extending from the mid-1970's to the present, in which various management and related client-oriented studies had suggested the need for a synthesizing information system to help improve the empirical basis of management activity within RSA, if not within other components of the vocational rehabilitation (VR) system.

A description of the background would not be complete without noting that the environment of the project was less than stable. First, the VR system is in the process of adjusting to Congressional mandates to serve the severely disabled more effectively. This includes implementing a major thrust in the field of independent living. Also, at the beginning of the project RSA was in DHEW's Office of Human Development Services, and was engaged in pursuing that organization's policy initiatives. In April of 1980, the agency became part of the Office of Special Education and Rehabilitation Services in the newly-formed Department of Education and found itself in need of defining ways to implement cross-cutting objectives with former Office of Education components. Finally, this period was characterized by intense introspection by senior RSA staff concerning the most appropriate role for the agency to play vis-a-vis the state VR agencies, public and private sector service providers, and the rehabilitation research community now largely under the fiscal and programmatic wing of the National Institute of Handicapped Research.

Objectives of the Abt Associates Study

The factors indicated above rendered impossible a management information system (MIS) concept which relied on fixed agency objectives and fixed notions as to which management functions were to be carried out

most assiduously. Rather, the concept which has evolved is that of creating a flexible, but nonetheless systematic, information-producing environment for RSA managers which can be useful in both present, more or less "known" circumstances, as well as in undefined future ones.

From the standpoint of the study efforts conducted by Abt Associates, Inc. (AAI), there has remained the necessity of initiating and completing specific tasks for the purposes of:

1. Developing and maintaining working relationships with the RSA staff who will form the core of the MIS user network

2. Re-stating the mission, goals, objectives, and management functions of RSA in ways which could be, at least theoretically, linked to needs for information

3. Describing the state of information availability and use within the agency

4. Identifying the substantive information elements which RSA managers need to carry out their functions and describing the minimum conditions of access which would meet those information needs

5. Drafting a design for the analytical capability to be built into the MIS, including a revised set of recurring reports

6. Describing an MIS design which integrates all of the earlier work, and specifies the major hardware and software subsystems

While these were indeed the milestones of the initial phase, it was necessary to extend substantially the time frame necessary to accomplish them and to modify their content in very substantial ways. It was also true that the processes which the AAI staff followed to obtain the information varied widely from initial expectations. None of these circumstances, however, are unusual in the experience of MIS design and development.

Major Data Collection Strategies Followed

Interviews With Key Personnel

There were, in effect, three waves of interviews. The first targeted RSA staff in an effort to learn their views of their respective roles and the function which information was playing or could play in helping them to carry out those roles.

The second wave involved non-RSA personnel in the legislative branch, OMB, OHDS, the HEW Audit Agency, DMC, and state agencies. A limited number of RSA Regional Office staff were involved at this point. The need for a third wave involving RSA staff reactions to prospective

information elements was obviated by the convening of the RSA MIS Workgroup in April of 1980.

The actual third wave combined two objectives. First, a contract modification calling for coverage of discretionary programs, as appropriate, in all elements of Phase I necessitated a round of staff interviews in RSA's Central and Regional Offices as well as at some 33 discretionary program projects themselves. In the latter, AAI was to gain an understanding of the data which could feasibly be expected to flow from such projects in ways which would produce information of practical use to RSA managers. Additionally, it had become very apparent that Regional Office staff had very explicit needs for access to data which existing arrangements had not facilitated. Thus, five Regional Offices were visited to allow for unrestrained discussions of how MIS design concepts should be shaped so that their concerns could be incorporated effectively.

Invitational Symposia

Three invitational symposia were designed to bring together important intellectual contributors to MIS related topics and to address jointly general and specific concerns.

The first symposium was general in its focus and was held June 11-12, 1979. The purpose of this symposium was to provide a forum for the presentation and discussion of issues concerning RSA's current experiences with information management. The gathering attracted a diverse group of participants representing various levels of RSA's operations. The discussion was structured around eight presentations, each concentrating on some aspect of the state-of-the-art of current management information systems in RSA. These presentations were as follows:

1. The VR System: Past and Present
2. HDS Information Systems Improvement Strategy
3. Policy Management Uses of Information in RSA
4. Goal Setting in Relation to the Program and Financial Plan
5. The Regional Mission and the Need for Management Information Systems Data
6. Information Management in the Oregon Vocational Rehabilitation Division
7. The effect of MIS Target Group Descriptors on Case Management
8. Sampling and Sample Studies

A second symposium was convened on December 6 and 7, 1979 with a view to bringing RSA and CSAVR representatives together with those contractors, grantees, and RSA Regional Offices who were involved in projects that included attention to matters of (1) financial control, (2) financial decision making, and (3) the financial reporting necessary to satisfy #'s 1 and 2. This meeting dealt with the need for improving the quality of financial reporting and financial management strategies. However, participants were united in their doubts that a uniform system of financial practice could be developed across state lines, because of unique state-specific financial management requirements. The symposium concluded with agreement on some major questions which participants hoped might be approached in subsequent gatherings, as well as through other means. These questions were:

1. What are some of the specific elements that constitute a good financial management and purchasing system, and a financial reporting system?

2. What changes need to be introduced in current source documents to support planning (including budgeting), monitoring, and evaluating?

3. What kind of information about financial practices and transactions needs to be transmitted to RSA Central and/or Regional Offices as a subset of that information needed by state agencies and VR facilities for the purposes of sound budgeting and accounting?

4. What kind of strategies can be used to introduce necessary changes in financial practices and reporting procedures?

While other meetings on this topic have been infeasible to date, the commitment of participants to continue to have informal contacts has been realized to the genuine benefit of the MIS project.

The third symposium actually combined two scheduled symposia. On April 7 and 8, 1980, RSA staff, members of CSAVR, a representative of the Executive Office of the President, Abt Associates, and others met with leading researchers in the field of functional assessment measurement and the use of weighted case closures and other innovative outcome measures. In terms of its effect on subsequent MIS developments, this was certainly a pivotal conference. This symposium proceeded through a number of well-developed presentations of the results of applied research in these topics. At the conclusion of the symposium, there emerged a powerful consensus of researchers and VR administrators alike that this technology had advanced to a point where large scale implementation was possible.

On April 9, the group was joined by some additional presenters for a lively discussion of evaluation and monitoring issues affecting independent living (IL) programs. During this session, participants evinced strong apprehensions concerning the strategies RSA might adopt regarding

measurement of client outcomes in IL projects. There appeared to be considerable sentiment to the effect that IL projects (particularly Part B projects) should be left free to develop outside of formal evaluation structures for some years more. While several very well prepared papers suggested that process and outcome measurement in Independent Living was a reasonable prospect with sold empirical underpinnings, the major outcome of this meeting may have been the formation of a network of project leaders pledged to "coordinate" with each other.

The RSA MIS Workgroup

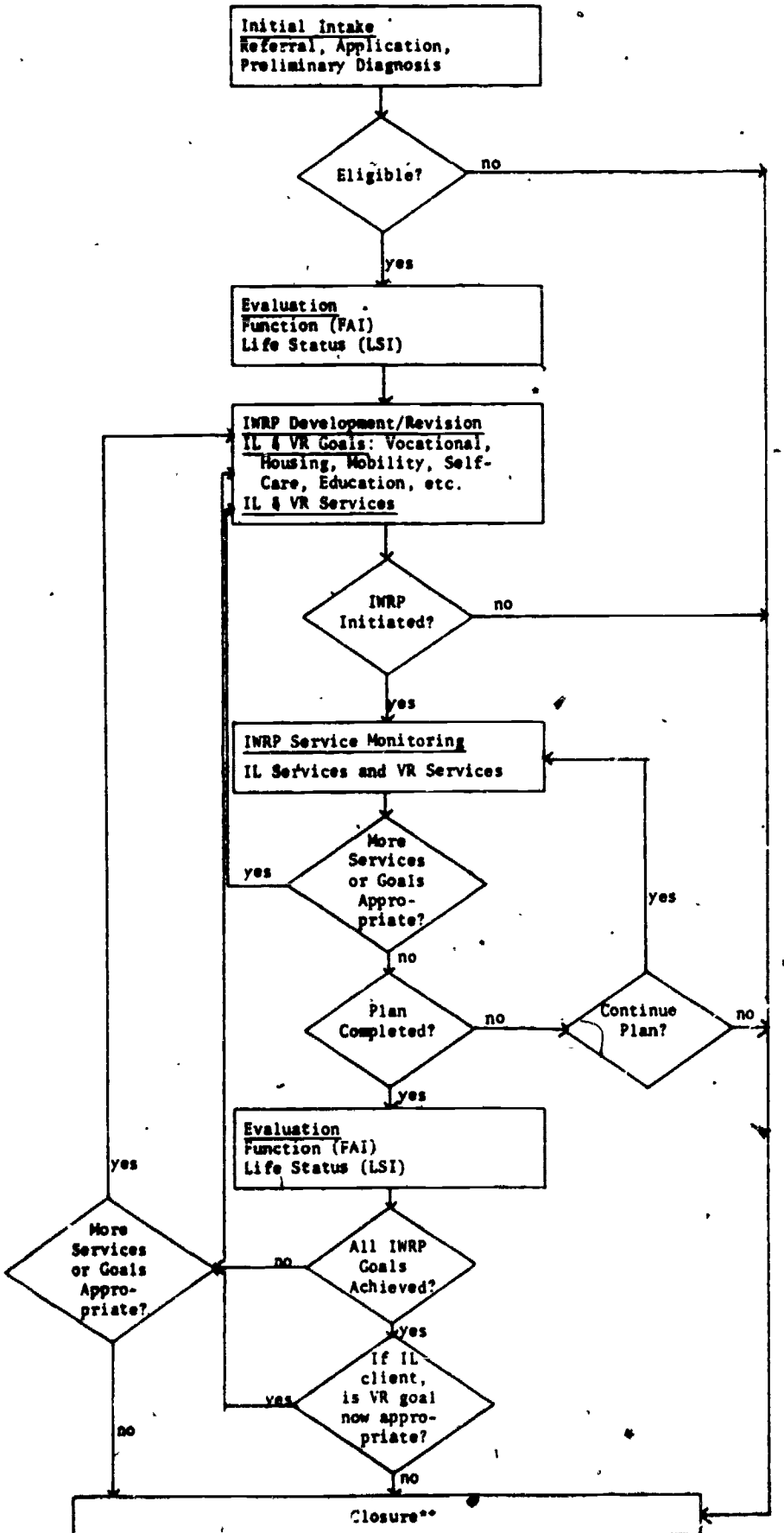
In April, 1980, after a delay of several months, RSA Commissioner Humphreys announced the appointment of a workgroup which would have responsibility for definitively articulating the information needs of RSA. The workgroup met without AAI for approximately three weeks, after which it unveiled a sweeping revision to the conceptual framework of the project. When AAI rejoined the process in mid-May, the workgroup had also received top management approval to recommend adoption of program-wide scales for assessing clients' functional abilities, implementation of a system of life status indicators as IWRP goal and client change measures, and the synthesis of an integrated VR and IL Part A client flow model.

The client flow model permitted the accommodation of independent living goals and the incorporation of functional assessment measures and Life Status Indicators in the VR process. The current version of the model is presented in Exhibit 1.1; each of the major components and decision points is represented. An Initial Intake component provides the determination of Eligibility. Once established, an Evaluation component provides information needed for Plan Development, which includes IL and VR goals and services. Once the plan has been developed, the client has the option of accepting the plan and receiving Services. (The client's eligibility status may also be reviewed by the agency.) The IL and VR services are Monitored, during which decisions about their appropriateness and plan completion are made. If the plan is completed, an Evaluation process is undertaken to provide change data. Finally, the model permits the assessment of goals and a decision concerning whether a VR goal is now appropriate for IL clients. If it is, then a new IWRP can be initiated.

The workgroup also articulated an extensive list of information elements as the core of the agency's needs with respect to the Basic 110/SSI/SSDI program. It then became the task of AAI's staff to give technical dimension to the workgroup's specifications and to incorporate them into subsequent reports. The expertise of workgroup members was very extensive and their interest and enthusiasm for the project converted the project into a highly collaborative effort involving almost daily contact between workgroup members and AAI staff assigned to specific tasks. This relationship made nearly certain that reports developed through the remainder of Phase I would be consonant with the most active concerns of the agency.

Exhibit 1.1

RSA VR/IL Client Flow Model (Draft)



NOTES:

* A VR client (110 Support) is a client whose goals include a vocational goal; an IL client (Part A) is a client whose goals do not include a vocational goal. There is one plan. In this model, the plan becomes the important baseline point for data. Some standardization of descriptors of plan contents is called for.

** In this model, closure for the VR client implies achievement of the vocational goal plus any other goals in the plan. For a client with a single vocational goal, this is the equivalent of the current 26. Successful closure for the IL client means plan completion and achievement of all IWRP goals; since some goals may relate to establishment of maintenance or life support resources, this does not necessarily imply termination of services. For IL clients, "closure" may have a different meaning. For resources management, data management, and integration into the VR case flow model, closure implies IWRP goal achievement.

In August, 1980, the workgroup's mandate was extended for the life of the MIS project and was re-titled the "MIS Implementation Executive Task Group." Thus, the MIS's prospects for successful implementation improved immeasurably as it moved toward Phase II - initiation.

Major Phase I Reports

Abt Associates has submitted five major and three secondary* reports to RSA under this contract. The major reports are as follows:

1. State-of-the-art Review
2. Conceptual Framework
3. Information Needs Assessment
4. Draft Statistical Analysis Plan
5. Preliminary Systems Design

The Conceptual Framework

The purpose of developing a conceptual framework was to map the programmatic and functional features of RSA in a way which would assist AAI in working with policymakers, planners, program monitors, and evaluators in identifying their information needs. To succeed in this purpose, the conceptual framework attempted to:

1. Reveal the scope of RSA's formula and discretionary grant programs, and how these individual programs are intended to contribute to the accomplishment of organizational goals and mission
2. Describe the general performance criteria that RSA and other executive agencies generally acknowledge should apply in the management of programs
3. Depict RSA's major management functions and the classes of information needed to discharge these functions
4. Outline the criteria that would subsequently be applied in selecting specific data elements to support information needs

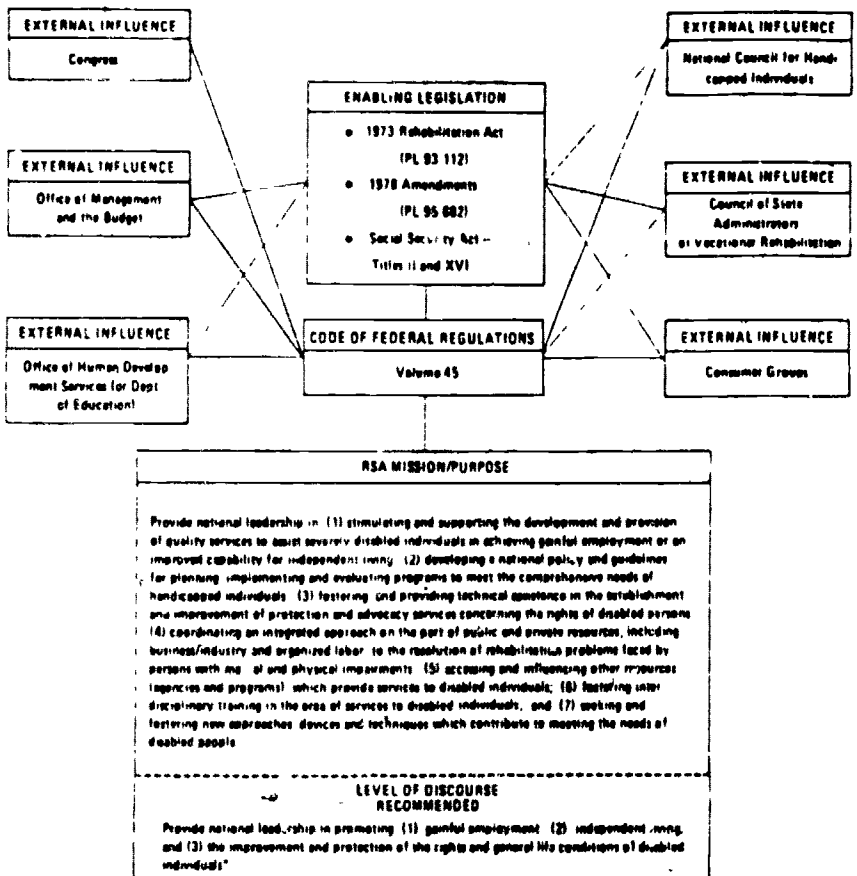
Furthermore, to serve its intended purpose successfully, the conceptual framework strove (a) to abridge goals, programs, functions, and the like to their fundamental parts without losing the essence of RSA, and (b) to avoid being so concrete that even moderate changes in RSA's

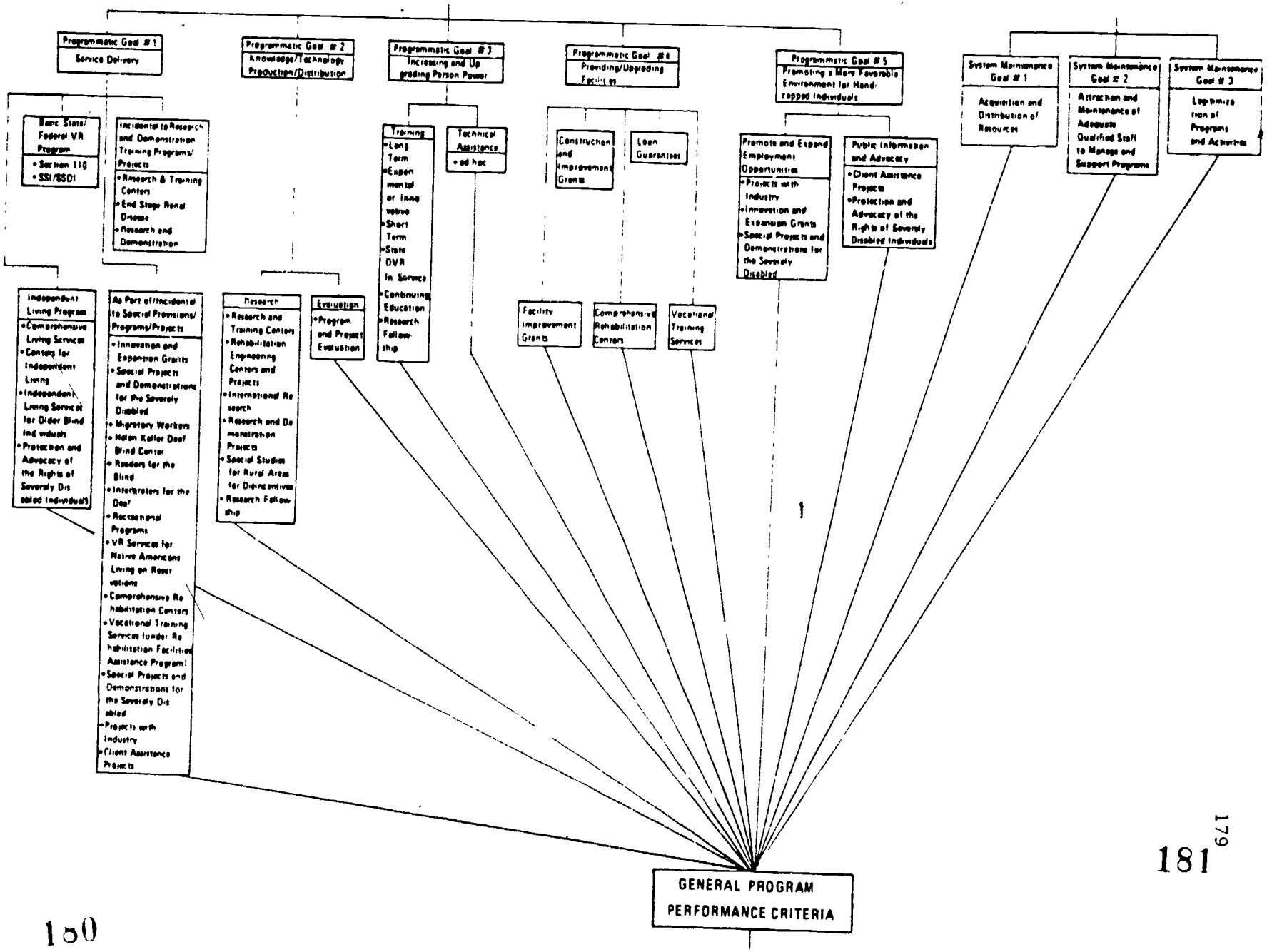
---

* The secondary reports are summaries of the three symposia, much of which is reflected in the major reports.

Exhibit 2.1

GRAPHIC SUMMARY OF THE CONCEPTUAL FRAMEWORK




Programmatic Goal # 1
Service Delivery
Basic State/ Federal VR Program
• Section 110
• SSI/SSDI
Incidental to Research and Demonstration Training Programs/ Projects
• Research & Training Centers
• End Stage Renal Disease
• Research and Demonstration
Programmatic Goal # 2
Knowledge/Technology Production/Distribution
Programmatic Goal # 3
Increasing and Upgrading Person Power
Training
• Long Term
• Experimental or Innovative
• Short Term
• State DVR In Service
• Continuing Education
• Research Fellowship
Technical Assistance
• ad hoc
Programmatic Goal # 4
Providing/Upgrading Facilities
Construction and Improvement Grants
Loan Guarantees
Programmatic Goal # 5
Promoting a More Favorable Environment for Handicapped Individuals
Promote and Expand Employment Opportunities
• Projects with Industry
• Innovation and Expansion Grants
• Special Projects and Demonstrations for the Severely Disabled
Public Information and Advocacy
• Client Assistance Projects
• Protection and Advocacy of the Rights of Severely Disabled Individuals
System Maintenance Goal # 1
Acquisition and Distribution of Resources
System Maintenance Goal # 2
Attraction and Maintenance of Adequate Qualified Staff to Manage and Support Programs
System Maintenance Goal # 3
Legitimization of Programs and Activities
Independent Living Program
• Comprehensive Living Services
• Centers for Independent Living
• Independent Living Services for Older Blind Individuals
• Protection and Advocacy of the Rights of Severely Disabled Individuals
As Part of/Incidental to Special Provisions/ Programs/Projects
• Innovation and Expansion Grants
• Special Projects and Demonstrations for the Severely Disabled
• Migratory Workers
• Helen Keller Deaf Blind Center
• Readers for the Blind
• Interpreters for the Deaf
• Recreational Programs
• VR Services for Native Americans Living on Reservations
• Comprehensive Rehabilitation Centers
• Vocational Training Services (under Rehabilitation Facilities Assistance Program)
• Special Projects and Demonstrations for the Severely Disabled
• Projects with Industry
• Client Assistance Projects
Research
• Research and Training Centers
• Rehabilitation Engineering Centers and Projects
• International Research
• Research and Demonstration Projects
• Special Studies for Rural Areas for Disincentives
• Research Fellowship
Evaluation
• Program and Project Evaluation
Facility Improvement Grants
Comprehensive Rehabilitation Centers
Vocational Training Services
GENERAL PROGRAM PERFORMANCE CRITERIA

programs or functions would render the framework obsolete.

The explication of the conceptual framework can be viewed schematically in Exhibit 2.1. It reveals RSA's mission, five programmatic goals and the projects and programs related to each, three system maintenance goals, and finally, five performance criteria which interact with RSA's management functions of planning, monitoring, and evaluating to project needs for management information related to:

1. Program availability
2. Program appropriateness
3. Program utilization
4. Program effectiveness
5. Program efficiency

This framework helped to structure inquiries regarding the information element needs of RSA managers. Not unexpectedly, actual experimentation with this framework by the RSA MIS workgroup led to some collaboratively proposed amendments intended to streamline the framework and to structure it with an issue-oriented vocabulary with which members of the VR system would be more conversant. With such changes, the classes of management information called for by the conceptual framework included: (a) program coverage, (b) program efficiency, (c) program impact, and (4) program compliance (C-E-I-C). These reflect RSA's management activity directed at insuring that certain general program qualities are achieved as follows:

1. Coverage (C) - that RSA-administered programs have a scale with a knowable relationship to measures of the scale of need and that they are available on an "equitable" basis
2. Efficiency (E) - that these programs utilize physical, fiscal, and personnel resources efficiently
3. Impact (I) - that the program activities result in some favorable changes among those in need
4. Compliance (C) - that the processes involved in executing the programs comply with a variety of mandatory provisions, general plans and targets, precepts of management practice, and with RSA policy initiatives

The final contribution of the conceptual framework to the eventual task of identifying information elements was the specification of technical criteria for selecting data elements.* Three such criteria were

* In the Information Needs Assessment Report, AAI extended the concept of technical criteria to apply to the quality of the information elements that passed the screening criteria proposed here, and discussed how these qualities would be defended by the proposed system design.

of interest in deciding whether or not an expressed need for an information element should be satisfied. They were: (a) face validity, (b) accessibility, and (c) cost. Each is explained briefly below:

Face validity was intended to mean two things: (a) the data element bears a logical relationship to the information class it is intended to support, and (b) the data element supports one or more of the statistical analyses which will be the basis for the MIS outputs.* At first glance, it might not appear to have been necessary to state the first requirement for face validity; after all, who would even consider recommending a data element that bore no relationship to an information class? The concern was not that completely extraneous recommendations for data elements would be made, but that recommendations might not carry the weight of agreement. For example, it is reasonable to assume that for the Section 110 program, one of the interests to support information about program coverage would be the "proportion of the potentially eligible target population that is actually served." While there would probably be a general consensus regarding this data need, there might be far less agreement about the data element supporting the need; i.e., would we use the Ridge/Worrall, JWK, or U.S. Census approach to estimating the size of the target population?

The second requirement for face validity was intended to emphasize our recommendation that data elements should not be included unless they have near term applicability to the MIS outputs. Warehousing data for some as yet unforeseen use would increase the costs, complexity, and burden of information handling and use, resulting in decreased efficiency in MIS management and probably decreased utilization of the MIS by program managers.

Accessibility refers to the likelihood that data can be obtained and that they will be reliable. A management information system cannot thrive if the data it relies on are unobtainable and/or of questionable quality because program managers will either fail to use the system due to a lack of confidence in its outputs, or they will use the system in spite of its deficiencies, thereby running the risk of committing grievous decision errors. Several conditions must be met in order to insure accessibility.

1. Data elements must be operationally definable, i.e., there must be a means to measure the variable of interest. It might be of considerable interest to determine the extent to which vocational gains can be attributed to VR services. However, it may be impossible to collect this information on individual clients, and even extremely improbable that it should be collected on a regular basis for groups of VR clients because of

---

* This second point reveals that final specification of data elements will be an iterative process, i.e., draft specification → statistical analysis plan → final specifications.

the need for very large samples and a true experimental design to control for an imponderable number of covariates. It may be possible to do this as an occasional large-scale social experiment, however, rather than as periodic input into an automated MIS.

2. Not only must variables be operationally definable, but these definitions must enjoy a great deal of spatial and temporal permanence (reliability). If definitions are not uniformly understood and applied by counselors, facilities, state directors of VR or whoever the data source may be at different sites, and by the same sources at different points in time, it would be difficult, at best, to interpret the results at any particular point in time, and also meaningless to aggregate or compare data across sites or over time.

3. Finally, RSA must have the authority to obtain the data. Variables may be operationally definable and it may be possible to collect data reliably, but RSA may simply not have the authority to require that these data be submitted by those who are the sources of the data or to fund alternative means of collecting the data.

The final major criteria for selecting data elements to be discussed is cost. We use the term "cost" not only to refer to money, but also to time and the related concept of the ratio of benefit to burden. Unlike the other two criteria, cost is impossible to evaluate adequately for individual data elements independent of a view of the entire information system. Cost can be determined on a data element by data element basis; however, without a cosmic view of the system, the question of how much is too much (money or burden) cannot be answered except perhaps for cases of outrageously expensive and very burdensome items.

Three other milestones were also reached during Phase I of the project: (a) a review of the state-of-the-art in information handling and use in RSA; (b) an information needs assessment; and (c) a statistical analysis plan for the data elements proposed for the system.

A Look Toward Phase II

Phase II's major milestones will include the following:

1. Preparation of a final, detailed system design which specifies all software and hardware approaches and follows administrative approval of system components

2. Completion of data acquisition and analysis pretests relative to new or altered data requirements, their practical utility, and the burden they place on states and other organizations involved in providing input to the MIS

3. Completion of systems and applications programming required to run a large-scale system test involving all MIS subsystems and modules

4. Completion of the systems test and conveyance of the software on a "turnkey" basis to RSA and the Department of Education

These milestones will leave RSA with a highly usable source of technical support for its planning, program implementation, monitoring, evaluating, and reporting activity in the coming decade.

"PROGRAM EVALUATION FOR REHABILITATION AGENCY PERSONNEL"

List of Participants (Trainees)

James Agre
Div. of Vocational Rehabilitation
Room #1005 Labor & Industry Bldg.
Trenton, New Jersey 08625

Mary Arginteanu
Comm. for Visually Handicapped
3003 Parkwood Avenue
Richmond, Virginia 23221

Richard W. Ault
Maine Bureau of Rehabilitation
32 Winthrop Street
Augusta, Maine 04330

Mario G. Barillas
Rehab. Education & Services
507 10th Street - 5th Floor
Des Moines, Iowa 50309

Bill Beasley, Jr., Ph.D.
Rehab. & Crippled Children Services
P.O. Box 11586
2129 East South Boulevard
Montgomery, Alabama 36111

Elizabeth J. Bilheimer
Arkansas Rehabilitation Services
1401 Brookwood
P.O. Box 3781
Little Rock, Arkansas 72203

Lana C. Brenes
San Diego State University
Rehabilitation Center
6363 Alvarado Court
San Diego, CA 92120

David Brown
State Commission for the Blind
1100 Raymond Blvd.
Newark, New Jersey 07102

Bill Brownfield
Dept. of Vocational Rehabilitation
4901 Fitzhugh Avenue
P.O. Box 11045
Richmond, Virginia 23230

Gloria Burger
Jewish Vocational Service
1 South Franklin
Chicago, Illinois 60606

Franklin E. Campbell
Region X RSA
1321 2nd Avenue
Seattle, Washington 98101

Wendell H. Carter
Commission on Accreditation of
Rehabilitation Facilities
2500 North Pantano Road
Tucson, Arizona 85715

Charles S. Chandler, Ph.D.
Vocational Rehabilitation Dept.
P.O. Box 4945
Columbia, South Carolina 29240

Anthony Cobb
Iowa Commission for the Blind
Fourth and Keosauqua Way
Des Moines, Iowa 50309

John Collins
Div. of Vocational Rehabilitation
623 East Adams
Chicago, Illinois 60603

Paul G. Cooper
Arkansas Rehabilitation Research
and Training Center
Fayetteville, Arkansas 72701

Justino Correa
Puerto Rico State Vocational
Rehabilitation Program
Box 1118
Hato Rey, Puerto Rico 00919

Dennis Cox
Div. of Vocational Rehabilitation
P.O. Box 1830
Santa Fe, New Mexico 87503

G. Wendell Cox
Rehabilitation for the Blind
P.O. Box 4872
Jackson, Mississippi 39216

Sam R. Crawford
Bureau for the Blind
619 East Capitol
Jefferson City, Missouri 65101

Steve Daggett
Regional Rehabilitation Institute
Portland State University
P.O. Box 751
Portland, Oregon 97207

Robert Darnell
Division of Rehabilitation
1575 Sherman
Denver, Colorado 80203

Donald Dellario
University of Scranton
Scranton, Pennsylvania 18510

Jane Douglas
Rehabilitation Division
505 East King
Carson City, Nevada 89710

Sandra Drabik
Div. of Vocational Rehabilitation
Room 1005 Labor & Industry Bldg.
Trenton, New Jersey 08625

Joseph A. Farrell
Vocational Rehabilitation
40 Fountain Street
Providence, Rhode Island 02903

Richard E. Field
Visual Services Division
Social & Rehabilitation Services
Box 4210
Helena, Montana 59601

John Folk
Div. of Vocational Rehabilitation
P.O. Box 8717 BWI Airport
Baltimore, Maryland 21240

Wilma T. Fountroy
Rehabilitation Services Commission
4656 Heaton Road
Columbus, Ohio 43229

Billy Fox, Ph.D.
Vocational Rehabilitation Division
P.O. Box 1698
Jackson, Mississippi 39205

Edith C. French
Division of Blind Services
2571 Executive Center Circle East
Tallahassee, Florida 32301

D. Ray Fuller, Jr.
Div. of Rehabilitation Services
1401 Brookwood Drive
P.O. Box 3781
Little Rock, Arkansas 72203

Gary F. Gaeth
Dept. of Social Services
Office of Services for the Blind
300 South Capitol Bldg.
Lansing, MI 48926

Ralph W. Gant
Utah State Board of Education
Div. of Rehabilitation Services
250 East 500 South
Salt Lake City, Utah 84111

David L. Gardner
Government of Guam
Dept. of Vocational Rehabilitation
P.O. Box 10-C
Agana, Guam 96910

Dennis A. Gay
University of Northern Colorado
Greeley, Colorado 80639

Marion German
Div. of Vocational Rehabilitation
Room 1005 Labor & Industry Bldg.
Trenton, New Jersey 08625

Jim Hall
Bureau of Rehabilitation
P.O. Box 30010
Lansing, MI 48909

Ray Halverson
Commission for the Blind
State House
Boise, Idaho 83720

Jim Harper
University of Arkansas
Box 3781
Little Rock, Arkansas 72203

David Hollingsworth
Syracuse University
Division of Special Education
and Rehabilitation
805 South Crouse Avenue
Syracuse, New York 13210

Carroll Hostetter
Rehabilitation Services Division
Dept. of Institutions, Social &
Rehabilitation Services
P.O. Box 25352
Oklahoma City, Oklahoma 73125

Gerald Hunter
Indiana Rehabilitation Services
1001 Willinois Bldg., Room 338
17 West Market
Indianapolis, Indiana 46204

Arloine M. Hutcheson
Texas State Commission
for the Blind
314 W. 11th Street
Austin, Texas 78753

K. A. Jagannathan, Ph.D.
Administration on Aging
Room 3655 North Building
330 Independence Avenue, S.W.
Washington, DC 20201

Betty Jo Jensen, Ph.D.
Div. of Services for the
Visually Handicapped
309 East 1st South
Salt Lake City, Utah 84121

Mary Joyce
R.I. Vocational Rehabilitation
40 Fountain Street
Providence, Rhode Island 02908

Orv Karan
Waisman Center on Mental Retardation
1500 Highland Avenue
Madison, Wisconsin 53706

Bob Lange
Colorado State University
Gaggenheim Hall
Ft. Collins, Colorado 80521

Jack E. Larson
Div. of Vocational Rehabilitation
3523 North Ten Mile Drive
Jefferson City, Missouri 65101

Alfred H. Letourneau
Dept. of Social and Rehab. Services
State Office Building
Montpelier, Vermont 05602

Allan G. Levine
JFK Federal Building
Boston, Massachusetts 02203

Perry Levinson, Ph.D.
Office of Rehabilitation Services
Department of Education
101 Marietta Tower, Suite 903
Atlanta, Georgia 30323

Jeffrey McCarthy
College Regional Rehabilitation
Continuing Education Center
500 Salisbury Street
Worcester, Massachusetts 01069

Mike McGill
Dept. of Institutions, Social &
Rehabilitation Services
P.O. Box 25352
Oklahoma City, Oklahoma 73125

Don McLaughlin
Research and Training Center
509 Allen Hall
West Virginia University
Morgantown, West Virginia 26506

Kathleen Medred
Vocational Rehabilitation
Pouch F. 0581
State Office Building
Juneau, Alaska 99811

Mary Jane Meehan
Bureau of Rehabilitation
122 C Street, NW
Washington, DC 20202

Audrey L. Miles
50 U.N. Plaza, Room 455
San Francisco, California 94102

David L. Miller
Program Evaluation & Statistics
830 "K" Street Mall, Room 124
Sacramento, California 95814

Juliet V. Miller, Ph.D.
ERIC Clearinghouse for Adult,
Career & Vocational Education
National Center for Research
in Vocational Education
The Ohio State University
1960 Kenny Road
Columbus, Ohio 43210

Jeffrey L. Moran
NYS Education Department
Office of Voc. Rehabilitation
99 Washington Avenue
Albany, New York 12230

Seymour Mund
Div. of Vocational Rehabilitation
614 Asylum Avenue
Hartford, Connecticut 06105

Richard Murray
Syracuse University
805 S. Crouse Avenue
Syracuse, New York 13210

Richard A. Nida, Ph.D.
Research & Training Center
One Dunbar Plaza
Dunbar, West Virginia 25064

Toshio Nishioka
Vocational Rehabilitation and
Services for the Blind
Honolulu, Hawaii 96809

Richard W. Norcross
Mass. Rehabilitation Commission
296 Boylston Street
Boston, Massachusetts 02116

George W. Parrish
Vocational Rehabilitation Services
P.O. Box 26053
Raleigh, North Carolina 27611

Paulette Pellani
Commission for the Blind
1100 Raymond Blvd.
Newark, New Jersey 07102

Victoria Placzek
Services for the Visually Impaired
1047 South Street
Lincoln, Nebraska 68502

Gary Prazak
Walker & Associates
123 East Grant Street
Minneapolis, Minnesota 55403

Stephen B. Preston
Div. of Vocational Rehabilitation
State Capitol
Charleston, West Virginia 25305

Charles R. Raeke
Commission for the Blind
P.O. Box 12866
Austin, Texas 78735

Carl Rennewitz
Vocational Rehab. Services Division
Dept. of Social & Health Services
MS OB-31C
Olympia, Washington 98504

Mary Rider
Div. of Vocational Rehabilitation
1 West Wilson St., Room 720
Madison, Wisconsin 53702

John F. Schlicting
Div. of Vocational Rehabilitation
Pouch F State Office Building
Juneau, Alaska 99803

Mr. Schlothauer
Commission for the Blind
Austin, Texas 78735

Jack C. Scott
Services for the Blind
P.O. Box 2658
Raleigh, North Carolina 27611

Julien C. Seibert
Mass. Rehabilitation Commission
296 Boylston Street
Boston, Massachusetts 02116

Duane T. Sermon
Div. of Vocational Rehabilitation
Third Floor Space Center Bldg.
444 Lafayette Road
St. Paul, Minnesota 55109

Roberta S. Shaw
Developmental Disabilities
1418 Luisa Street
Santa Fe, New Mexico 87501

Catherine Sheridan
Office of Vocational Rehabilitation
99 Washington Avenue
Albany, New York 12210

Neil R. Sherwood
Vocational Rehabilitation Div.
2045 Silverton Road, NE
Salem, Oregon 97310

Asher R. Soloff
Jewish Vocational Service
1 S. Franklin St.
Chicago, Illinois 60606

David O. Songer
Bureau of Rehab. Services
122 "C" Street, NW
Washington, DC 20202

Dean G. Stanzel
DHEW/OHDS/ORS
610 East 12th, Room 384
Kansas City, Missouri 64106

Jim Stelling
Vocational Rehabilitation
1808 West End Bldg., Room 1400
Nashville, Tennessee 37203

Howard J. Stewart
Office for the Visually Handicapped
Room 330 CAB Building
Harrisburg, Pennsylvania k6k29

Darlene Taylor
Div. of Vocational Rehabilitation
603 State Office Building
47 Trinity Avenue
Atlanta, Georgia 30334

Joseph G. Taylor
Vocational Rehabilitation
1309 Winewood Blvd.
Tallahassee, Florida 32301

Edwin C. Thomas
Commission for the Blind
1430 Confederate Avenue
Columbia, South Carolina 29210

Robert T. Tipton
Bureau of Rehabilitation Services
Capitol Plaza Tower
Frankfort, Kentucky 40601

Francisco Vallejo
Vocational Rehabilitation Program
Box 1118
Hato Rey, Puerto Rico 00919

Guy Veach
Div. of Vocational Rehabilitation
1500 Shallcross Avenue
Wilmington, Delaware 19899

Robert Wares
Vocational Rehabilitation
3523 North Ten Mile Drive
Jefferson City, Missouri 65101

Carol J Whitcraft
Texas Rehabilitation Commission
118 East Riverside Drive
Austin, Texas 78704

Barbara Witten
Syracuse Universtiy
Div. of Special Ed. & Rehabilitation
805 S. Crouse Avenue
Syracuse, New York 13210

Everett L. Wright
Vocational Rehabilitation
P.O. Box 573
Baton Rouge, Louisiana 70821

Syracuse University student participants: John Goeke, Naomi Meyer, Donna Rubin