DOCUMENT RESUME

ED 110 507 TM 004 780

AUTHOR Rubin, Stanford E., Ed.
TITLE Issues in Program Evaluation Research in State Rehabilitation Agencies: A Symposium Presented at the American Personnel and Guidance Association Convention (New Orleans, Louisiana, April 11, 1974).
INSTITUTION Arkansas Univ., Fayetteville. Arkansas Rehabilitation Research and Training Center.
SPONS AGENCY Department of Health, Education, and Welfare, Washington, D.C.
PUB DATE [Apr 74]
NOTE 54p.
AVAILABLE FROM Publications Section, Arkansas Rehabilitation and Training Center, Hot Springs, Arkansas 71901

EDRS PRICE MF-$0.76 HC-$3.32 PLUS POSTAGE
DESCRIPTORS Accountability; Models; Program Design; Program Development; Program Effectiveness; *Program Evaluation; Program Planning; *Rehabilitation Centers; Rehabilitation Programs; *Research; *State Agencies; *Vocational Rehabilitation

ABSTRACT

The five papers in this symposium address several key issues included in evaluation research. In "Designing State Vocational Rehabilitation Agency Evaluation Research", Stanford E. Rubin discusses the importance of practitioner involvement in the development of research questions and plans and introduces three basic models of evaluative research. Brian Bolton's paper, "Problems in Measuring Outcomes in Rehabilitation Research Projects", comments on such issues as the criterion problem in rehabilitation counseling, measures of client outcome, and procedures for quantifying psychometric change. Problems encountered in conducting large scale evaluation studies are discussed in the context of a Services Coordination Project in Richard Roessler's presentation, "Issues in Initiating Large Scale Rehabilitation Research." Reed Greenwood reviews the need for expanding the parameters of evaluative research to a system approved at both mezzo and macrosystem levels in his paper, "Assessment of External Effects in Rahabilitation Research." Finally, Jerold D. Bozarth, in "Reactions on Evaluation", discusses the differences between research and evaluation. (RC)

Research Abscract

Arkansas Rehabilitation
Research and Training
Center
University of Arkansas
Arkansas Rehabilitation
Services

Vol.___ No.___

July 1, 1974

Issues in Program Evaluation Research in State Rehabilitation Agencies

Evaluation research is becoming increasingly significant to the vocational rehabilitation field with the current emphasis on program evaluation and accountability. The series of papers abstracted here were presented as a Symposium at the American Personnel and Guidance Association Convention in New Orleans, Louisiana April 11, 1974, and review several key issues included in evaluation research.

ABSTRACT

Designing State Vocational Rehabilitation Agency Evaluation Research
Stanford E. Rubin

A large proportion of the research conducted within State Vocational Rehabilitation Agencies during the next few years will probably be evaluative in nature and oriented toward an immediate utilization objective. With this research direction in mind three basic models of evaluative research are introduced: Model 1 - measures the effect of an entire program; Model 2 - compares the specific variables within a program; and Model 3 - measures the differential effects of specific variables within a program. Model 3 is considered to have the greatest utility for providing data on which state rehabilitation agency administrators can make effective program decisions. The reasons for such a conclusion are put forth. Model 3 is then comprehensively described. The importance of practitioner involvement in the development of research questions and plans is discussed.

ABSTRACT

Problems in Measuring Outcomes in
Rehabilitation Research Projects
Brian Bolton

This paper consists of six parts. The first presents a brief introduction to the criterion problem in rehabilitation counseling. The second outlines four conceptual issues relevant to the measurement of outcomes (economic vs. noneconomic, unidimensional vs. multidimensional, absolute vs. relative, and self-report vs. observer ratings). The third reviews seven measures of client outcome which have been developed to address the criterion problem. The fourth briefly discusses two standard procedures for quantifying psychometric change (raw score and residual score methods). The fifth reviews three studies which support the conclusion that rehabilitation outcome is a multidimensional construct. The final section discusses the role of factor analysis in the development of models of the rehabilitation process and outlines two standard models.

ABSTRACT

Assessment of External Effects
in Rehabilitation Research
Reed Greenwood

Evaluative research in rehabilitation calls for approaches which recognize that rehabilitation outcome is determined by a complex group of factors, both internal and external to the rehabilitation system. Focusing on external or environmental influences affecting rehabilitation outcome, this presentation reviews the need for expanding the parameters of evaluative research to a system approved at both mezzo (intermediate) and macrosystem levels. Specific external factors at mezzo and macro levels include social climate, geography, social mores and folkways, economic conditions and institutional policies and practices.

ABSTRACT

Issues in Initiating Large
Scale Rehabilitation Research
Richard Roessler

Problems encountered in conducting large scale evaluation studies in rehabilitation settings are discussed

in the context of a Services Coordination Project (Arkansas Services Center, Jonesboro, Arkansas). In part, the paper discusses differences in perspectives held by those involved in different phases of the project; e.g., those responsible for initiating and managing the project and those responsible for evaluating its effectiveness. Other issues such as setting objectives, defining the experimental treatment, etc. are included. Finally, an effort is made to present resolutions to problems identified in the paper.

ABSTRACT

Reactions on Evaluation
Jerold D. Bozarth

Research and evaluation are often assumed to be different types of processes. Essentially, good evaluation is the same as good research since both processes require characteristics such as internal and external validity. However, most evaluation is at the lower end of the qualitative continuum and may be quite harmful if results from such evaluations are used to determine policy, or to modify programs. Evaluation research should be relevant to questions posed by groups such as legislators and boards of supervisors, and should be directed toward program modification. These questions must often be restated in ways which allow measurement. Other suggestions offered to strengthen evaluation research are to utilize a "helping" set rather than the more threatening "accountability" approach and to incorporate basic research into program evaluation efforts.

A copy of this publication may be obtained by writing to:
Publications Section
Arkansas Rehabilitation and Training Center
Hot Springs, Arkansas 71901

Issues in Program Evaluation Research
in State Rehabilitation Agencies

Stanford E. Rubin, Ed.

Arkansas Rehabilitation Research and Training Center
University of Arkansas/ Arkansas Rehabilitation Service

ISSUES IN PROGRAM EVALUATION RESEARCH IN STATE REHABILITATION AGENCIES

Stanford E. Rubin, Editor

A SYMPOSIUM PRESENTED AT THE

AMERICAN PERSONNEL AND GUIDANCE ASSOCIATION CONVENTION

New Orleans, Louisiana
April 11, 1974

Stanford E. Rubin, Ed.D., Chairman and Presenter

Brian Bolton, Ph.D., Presenter

Reed Greenwood, Ed.D., Presenter

Richard Roessler, Ph.D., Presenter

Jerold Bozarth, Ph.D., Presenter

Arkansas Rehabilitation Research and Training Center

University of Arkansas
College of Education

Arkansas Rehabilitation
Services

This project was supported in part by a grant (16-P-56812, RT-13) from the Department of Health, Education, and Welfare, Washington, D. C.

TABLE OF CONTENTS

INTRODUCTION

The Arkansas Rehabilitation Research and Training Center publishes papers which relate to topics of current significance to the field of vocational rehabilitation. Recognizing the importance of evaluative research in current program evaluation efforts being conducted in public vocational rehabilitation agencies, the authors prepared the papers included in this publication. Although not exhaustive in content, the papers should give the reader a feel for the complexities involved in evaluative research, by explicating design and measurement considerations, and additionally identifying environmental constraints and practical issues involved in implementing such research. The authors invite feedback and reactions to the presentations.

Reed Greenwood, Ed.D.
Director of Research
Arkansas Rehabilitation
Research and Training Center

DESIGNING STATE VOCATIONAL REHABILITATION AGENCY EVALUATION RESEARCH*

Stanford E. Rubin, Ed.D.
Arkansas Rehabilitation Research and Training Center
University of Arkansas

Since we are living through the decade of accountability, a large proportion of the research conducted within State Vocational Rehabilitation Agencies during the next few years will probably be evaluative in nature. Such research will be aimed at providing data on which program or sub-program termination, selective utilization, modification, or expansion decisions can be based. Therefore, such research will obviously be oriented toward an immediate utilization objective. Suchman (1967) succinctly reinforces this point with the following statement:

> "Unlike the basic researcher, the applied researcher must be constantly aware of the potential utility of his findings. Only rarely can he take consolation in the fact that the operation was a success but the patient died." (p. 21)

In spite of its applied nature, though, evaluative research must still adhere as closely as possible to the same rules governing basic research if any faith is to be placed in its findings. For example, evaluative research

*This paper is based on a presentation at the American Personnel and Guidance Association Convention in New Orleans, April 11, 1974.

must utilize measures of program activity and program effect that have demonstrated reliability and validity. Also as Suchman (1967) points out, like basic research, evaluative research requires a theoretical foundation. He goes on to stress that " . . . a test of Does it work? presupposes some theory as to why one might expect it to work." (p. 86)

There are at least three basic models of evaluative research which readily apply to state rehabilitation agencies:

Model 1 - Measures the effect of an entire program.

Example Research Question - What was the total gain in earnings of all clients rehabilitated by our agency?

Model 2 - Compares the effect of specific variables within a program.

Example Research Question - Do the clients of counselors with strong interpersonal skills have greater gains in earnings than the clients of counselors with weak interpersonal skills?

Model 3 - Measures the differential effects of specific variables within a program.

Example Research Question - Does the effect of counselor interpersonal skill level on clients depend on the type of client served?

Although all three models could be used for evaluation purposes, Model 3* is seen as having the greatest utility for providing data on which state rehabilitation agency administrators can make effective program decisions. This is the result of two factors. First, there are many different types of clients (old, young; physically disabled, emotionally disabled, intellectually impaired; etc.) with differing needs seeking rehabilitation services at state rehabilitation agencies. Second, there are many different types of service programs available. Consequently it could be hypothesized that Model 3 effectively and comprehensively implemented could throw light on the differential effectiveness of service programs for rehabilitating client types with different needs. It can easily be assumed that such data would be invaluable to agency decision makers.

At this point a fictitious example of Model 3 in operation will be utilized to demonstrate its application. The research question is:

> Are comprehensive rehabilitation facilities differentially effective in their attempt to rehabilitate physically disabled and psychiatrically disabled clients?

---

*For a more elaborate discussion of this model see Bolton and Rubin (1974).

Table 1 lists the variables (independent, moderator, and dependent). Table 2 illustrates Model 3 in a 2 X 2 two-way

TABLE 1

Project Variables

Independent Facility Variables	Moderator Client Variables	Dependent Client Outcome Variables
Facility A	Physically Disabled	Change in earnings from intake to closure
Facility B	Psychiatrically Disabled	

TABLE 2

Two-Way ANOVA Design for Studying Facility-Client Interaction

Independent Variables	Client Moderator Variables	
	Physically Disabled	Psychiatrically Disabled
Facility A	(Cell 1) Client change in earnings	(Cell 2) Client change in earnings
Facility B	(Cell 3) Client change in earnings	(Cell 4) Client change in earnings

analysis of variance. The six possible outcomes which can result from this design are listed in Table 3.

TABLE 3

Possible Outcomes

1. No significant difference	4. Facility and client main effect without interaction effects
2. Facility main effect	5. Ordinal interaction effect
3. Client main effect	6. Disordinal interaction effect

Only two of these outcomes, ordinal interaction effect and disordinal interaction effect, require consideration of the two sets of variables simultaneously.

The ordinal interaction effect is diagrammed in Table 4. An ordinal interaction effect indicates that working with one type of client, one facility produces results superior to another, whereas, the extent of change for a second type of client is much less related to the facility entered. For example, Facility A may be significantly more effective than Facility B with psychiatrically disabled clients, although they may show little difference in effectiveness when working with physically

disabled clients. If such a finding were real it would obviously suggest that most psychiatrically disabled clients should be sent to Facility A.

TABLE 4

The Effect of Two Facilities on Types of Clients

Facility Variable	Client Moderator		Client Outcome
Facility A working with:	Psychiatrically Disabled	produces	$60 average increase in earnings
	Physically Disabled	produces	$70 average increase in earnings
Facility B working with:	Psychiatrically Disabled	produces	$40 average increase in earnings
	Physically Disabled	produces	$70 average increase in earnings

A disordinal interaction effect indicates that one facility is more effective with one type of client while another facility is more effective with another type of client. The disordinal interaction effect is diagrammed

in Table 5. Obviously, if such a finding were real it would suggest that psychiatrically disabled clients should be sent to Facility A while physically disabled clients should be sent to Facility B.

TABLE 5

The Effect of Two Facilities on Two Types of Clients

Facility Variable	Client Moderator		Client Outcome
Facility A working with:	Psychiatrically Disabled	produces	$70 average increase in earnings
	Physically Disabled	produces	$60 average increase in earnings
Facility B working with:	Psychiatrically Disabled	produces	$30 average increase in earnings
	Physically Disabled	produces	$80 average increase in earnings

Assuming research competence, the effective utilization of Model 3 or any other complex research design is dependent on accessibility to data on program activity,

client characteristics, and client outcome. This requires that the evaluation researcher have the cooperation of state rehabilitation agency administrative and service personnel. Such cooperation cannot automatically be assumed because of at least two factors:

1. Fear that the research results will have a negative effect on program and personnel as opposed to a positive effect.
2. Resistance to the additional work that will be required to provide data to researchers.

Such impediments to effective evaluation research can probably be greatly overcome by providing many opportunities to state rehabilitation agency administrative and service personnel to participate in the development of the research questions as well as the research plan. Such participation should also greatly increase the likelihood that the research results will lend themselves to utilization as well as be utilized.

REFERENCES

Bolton, B. and Rubin, S. E. A research model for rehabilitation counselor performance. Rehabilitation Counseling Bulletin, 1974, 141-148.

Suchman, E. A. Evaluative Research. New York: Russell Sage Foundation, 1967.

PROBLEMS IN MEASURING OUTCOMES IN REHABILITATION RESEARCH PROJECTS*

Brian Bolton
Arkansas Rehabilitation Research and Training Center
University of Arkansas

Introduction: The Criterion Problem

Measures of client outcome have traditionally been utilized for two purposes: (1) as a basis for more fairly assessing counselor performance through the determination of caseload difficulty, and (2) as a basis for assessing agency effectiveness through the determination of client benefit. Underlying the rehabilitation philosophy is the assumption that successful vocational adjustment is fundamental to satisfactory adjustment in other areas of life. Therefore, client benefit has traditionally been synonymous with vocational benefit. In the last few years some rehabilitators have questioned the applicability of this assumption (e.g., Athlestan, et al., 1973). Thus, emphasis has been placed on the assessment of personal-social adjustment as a worthwhile goal independent of vocational success.

Issues in the Measurement of Outcomes

Four conceptual issues which are relevant to the measurement of client outcomes have been delineated by the author in the form of dichotomies (Bolton, 1974a):

---

*This paper is based on a presentation at the American Personnel and Guidance Association Convention in New Orleans, April 11, 1974.

1. Economic (vocational) versus non-economic (personal adjustment) measures of outcome.

2. Unidimensional (single score) versus multidimensional (multi-score) measures.

3. Absolute measures (assessment of final status irrespective of initial status) versus relative measures.

4. Self-report versus observer rating measures (see Fairweather, et al. study presented later).

As can readily be seen the above four dichotomies result in 16 possible combinations of criterion characteristics. For example, the combination of "economic-absolute-unidimensional-objective" is illustrated by salary or Vocational Adequacy at Closure (see below) and the "noneconomic-relative-multidimensional-self report" combination is illustrated by personal adjustment change measures.

The standard procedure followed in most rehabilitation research projects has been to enumerate a variety of relevant criteria and conduct separate analyses for each one (the "shotgun" approach). This multiple criteria approach is distinguished from a multidimensional measure of rehabilitation outcome in that the latter involves a systematic design and statistical procedure.

Selected Measures of Client Outcome

Several measures of client outcome have been developed in recent years in response to the criterion problem. They can be grouped into three unidimensional and four multidimensional instruments. The unidimensional instruments

are reviewed below.

1. Successful versus Unsuccessful closure (R-300): the dominant measure of rehabilitation effectiveness (inadequate).

2. Vocational Adequacy at Closure: a weighted composite score developed by Eber (1966) which includes work status, job code, weekly earnings, and closure status (economic).

3. Rehabilitation Gain: a composite score developed by Reagles, et al. (1970) which is based on 20 items which reflect vocational success and personal-social adjustment.

The above three unidimensional measures of client outcome possess the advantages of comprehensiveness and adequate reliability. Their weakness is that they may be averaging several meaningful independent criteria into a combination score which has no specific interpretation (i.e., Can vocational success and personal-social adjustment be meaningfully added together and labelled gain? Is it possible for clients to succeed vocationally and adjust poorly in other areas, and vice versa? If so, what does their gain score represent?). The four multidimensional measures listed below were developed to assess several relatively independent dimensions of client outcome.

4. Human Service Scale: a self-report questionnaire developed by Reagles, et al. (1973) which results in scores on seven "need" factors, e.g., physiological needs, emotional-security needs, etc.

5. Rehabilitation Services Outcome Measure: a counselor rating instrument developed by Westerheide and Lenhart (1973) which emphasizes the client's functional ability in relation to employment; six

subscale scores result: difficulty, educational status, economic/vocational status, etc.

6. DVR Project Personal Adjustment Change: a procedure which scores the pre/post responses of clients to the TSC, 16 PF, and Mini-Mult on six dimensions of change: self-concept, anxiety, extroversion, etc. (Rubin, et al., 1974).

7. DVR Project Vocational Change: a procedure which scores the pre/post R-300 variables of salary and work status on three (highly correlated) measures of change: salary change, work status change, and vocational success (Krauft, et al., 1974).

Measurement of Change

The quantification of change continues to constitute a major problem for psychometricians and applied researchers alike. Two major procedures, or some variants thereof, are in standard use. The first, the raw change score, is simply the difference between posttest and pretest performance. The raw change score procedure is inappropriate when ceiling effects or regression effects are serious problems. The residual change score may be a useful substitute under the conditions present wnen the raw score is inappropriate. Essentially, the residual change score compares individuals who have the same pretest status. O'Conner (1972) has provided the most up-to-date summary of the issues in measuring psychometric change.

Dimensionality of the Criterion

As suggested by the multidimensional measures outlined above, rehabilitation outcome is probably a multidimensional construct. Three research studies which further support

this conclusion are summarized below:

1. Fairweather, et al. (1960) intercorrelated 13 criterion measures collected in conjunction with a research project studying hospitalized psychiatric patients. Two relatively independent clusters appeared: (a) subjective self-evaluation via self-report inventories, and (b) objective interpersonal evaluation reflected in follow-up items and group therapy ratings.

2. Rosen, et al. (1970) factor analyzed 22 criteria of community adjustment for a sample of mentally retarded clients. Five factors resulted: coping behavior, job satisfaction (A), job stability, job satisfaction (B), and general adjustment.

3. Bolton (1974b) factor analyzed 12 measures of change in psychological adjustment (self-report) and three measures of change in vocational adjustment (R-300). Three factors resulted: I. Improvement in work status and salary, II. Improvement in self-concept (or Reduction in neurosis), and III. Reduction in psychosis. The first factor was uncorrelated with the other two factors. Thus, improved vocational functioning may be independent of improved psychological adjustment.

The Role of Factor Analysis in the Development of Rehabilitation Treatment Models

The development of quantitative models of the rehabilitation process constitutes an important step toward understanding how client characteristics interact with treatment parameters to produce successful outcomes. Such models can be effectively utilized to sort multiple variables into meaningful categories under the client and outcome components of any rehabilitation treatment model. They are less appropriate for treatment variables which are usually pre-determined by practitioners who must operate under restrictions imposed by realistic

considerations such as cost and client-staff ratio. Generally speaking, the treatment variables are not reducible or otherwise subject to statistical re-combination. The three major components of any rehabilitation treatment model may be further identified as follows:

A. Client variables
   1. Attributes (relatively fixed measures such as sex, age, education, etc.)
   2. Characteristics (relatively modifiable or changeable dimensions such as motivation, self-concept, vocational skills, etc.)

B. Treatment variables
   1. Counselor interpersonal skills and counseling techniques (friendliness, advising, occupational information, etc.)
   2. Services (physical restoration, psychotherapy, personal adjustment training, etc.)

C. Outcome variables
   1. Measured change on psychological characteristics (by self-report or observer rating techniques).
   2. Global indices of success (program completion, employment, self-sufficiency, etc.).

Some outcome variables may be reduced to composite dimensions while others should not be. In order to report the results of rehabilitation treatment efforts, administrators are required to cite statistics which reflect directly upon economic issues, e.g., employment, welfare recipients, etc. Thus, theoretical models of the rehabilitation process should be designed to incorporate directly the measures of treatment effort and outcome.

The situation is entirely different in the domain of client variables. Except for the need to record basic

attributes in order to assign clients to optimal treatments, client variables should be reduced to more fundamental dimensions whenever possible. Factor analysis is used to achieve parsimony in the client domain, which usually contains the largest number of variables. The factors then represent the major theoretical constructs in the psychological model of the rehabilitation process. The model gains generalizability by using the stable dimensions of client functioning underlying the treatment process.

Treatment Models

In the final section of this paper two basic rehabilitation treatment models are outlined:

Model I. This model assumes that the client's initial status regarding psychological characteristics in the domain of abilities and skills and intensity of services interact to determine the rehabilitation outcome. Factor analysis is employed to reduce the client initial status characteristics (which are assumed to be relatively stable) to fundamental dimensions. A moderated multiple regression analysis or computer pattern analysis (AID) is used to assemble the client, treatment, and outcome variables.

Model II. This model is appropriate for studying the counselor as the primary effector of change in the client. The assumption is made that various combinations of

counselor characteristics and skills interact with client characteristics in producing client change and rehabilitation success. Or stated in a slightly different manner, the effect of the counselor is moderated by client characteristics. The dependent variables in this ANOVA design can be psychometric change scores or global indices of rehabilitation success. Factor analysis is employed to isolate the primary dimensions of client change.

References

Athlestan, G., et al. Legitimizing non-vocational goals in rehabilitation counseling. Paper presented to the APA Convention, Montreal, 1973.

Bolton, B., et al. Clinical versus statistical prediction of client feasibility. Madison: UW-RRRI, 1968.

Bolton, B. The prediction of rehabilitation outcomes. Journal of Applied Rehabilitation Counseling, 1972, 3, 16-24. (a)

Bolton, B. Factor analysis in rehabilitation research. Rehabilitation Counseling Bulletin, 1972, 16, 94-105. (b)

Bolton, B. Assessment of Outcomes (Chapter 7). Introduction to rehabilitation research. Springfield, Illinois: Charles C. Thomas, 1974. (a)

Bolton, B. Dimensions of rehabilitation outcome: A factor analysis of personal adjustment and vocational measures of client change. Research Report, 1974. (b)

Fairweather, G., et al. Relative effectiveness of psychotherapeutic programs: A multicriteria comparison of four programs for three different patient groups. Psychological Monographs, 1960, 74 (Whole No. 492).

Krauft, C., Bolton, B., Richardson, B., & Rubin, S. Rehabilitation counselor behavior and client case service outcome. (Monograph XIII). Fayetteville: ARR&TC, 1974.

Muthard, J. E., & Miller, L. A. The criteria problems in rehabilitation counseling, 1966.

O'Conner, E. F. Extending classical test theory to the measurement of change. Review of Educational Research, 1972, 42, 73-97.

Reagles, K., Wright, G., & Butler, A. J. A scale of rehabilitation gain for clients of an expanded vocational rehabilitation program. (Monograph XIII), Madison, Wisconsin, RRRI, 1970.

Reagles, K., Wright, G., & Butler, A. J. The Human Service Scale. Madison, Wisconsin: Human Service System, 1973.

Rosen, M. et al. Prediction of postinstitutional adjustment of mentally retarded adults. American Journal of Mental Deficiency, 1970, 74, 726-733.

Rubin, S., Bolton, B., Krauft, C., Richardson, B., & Bozarth, J. Rehabilitation counselor behavior and client intermediate-term personality change. (Monograph XII), Fayetteville: ARR&TC, 1974.

Tucker, L. R. Elements of a central prediction system. Psychometric monograph No. 11, 1956.

Westerheide, J., & Lenhart, L. Development and reliability of a pretest-posttest rehabilitation services outcome measure. Rehabilitation Research and Practice Review, 1973, 4, (No. 2), 15-24.

ASSESSMENT OF EXTERNAL EFFECTS IN REHABILITATION RESEARCH*

Reed Greenwood, Ed.D.
Arkansas Rehabilitation Research and Training Center
University of Arkansas

Research, particularly of an evaluative type, has recently become a matter of central focus in the field of rehabilitation. The emphasis from the present executive and congressional leadership is for program accountability and demonstrated program effectiveness. In the past, vocational rehabilitation programs have often been supported and funded because of noteworthy intent and a deserving client population, in addition to demonstrated and assumed effectiveness. However, the era of reliance upon input measures of effectiveness and gross, cumulative output measures (rehabilitation closures) is bygone. Present policies not only encourage, but mandate more sophisticated evaluations of public vocational rehabilitation programs. This mandate places a heavy responsibility upon program administrators who in turn look to the expertise of rehabilitation researchers for guidance in developing program evaluation techniques.

---

*This paper is based on a presentation at the American Personnel and Guidance Association Convention in New Orleans, April 11, 1974.

Researchers do have an arsenal of techniques and evaluation methods available for use in assessing the effectiveness of rehabilitation efforts. However, these techniques are more often appropriate for controlled experimental conditions rather than for applied research where tightly controlled designs may not be possible. The challenge of identifying evaluation methods suitable for utilization with ongoing rehabilitation programs is considerable. The purpose of the following discussion is to focus on a somewhat neglected area of concern in which the need for careful planning and conceptualization for evaluation is essential; namely, the influence of the environment external to the rehabilitation process. Although external factors are difficult to assess in terms of strength of effect, they may contribute considerably to rehabilitation outcome.

Because the primary focus of the rehabilitation process has been the rehabilitation client, heavy emphasis in rehabilitation research has been placed on the individual and his effectiveness in coping with the environment. Efforts to predict rehabilitation outcome using client characteristics have resulted in findings which could account for only a portion of the variability contributing to outcome. After reviewing 40 such prediction

studies, Bolton (1972) found few which accounted for more than 25 percent of the variance in outcome. The remaining variance is determined by such factors as treatment parameters, client-counselor interaction, and other factors external to the rehabilitation process. These external factors include, but are not limited to, economic conditions, institutional barriers, and societal attitudes which may differentially affect the disabled.

Before looking more closely at these specific external factors, a framework for model building and analysis is needed. Mullen and Dumpson (1972) offer just such a framework for evaluative research in programs like vocational rehabilitation. Drawing upon general systems theory, three levels of systems are proposed: (1) microsystems, in which problems are viewed in terms of the determining focus being within the boundaries of either an individual or a small group directly experiencing the problem, (2) mezzosystems, in which the determining forces are seen as directly involving not only human systems but social systems on the scale of neighborhood and communities, and (3) macrosystems, in which the critical determinants extend beyond given individuals, groups, or localities, and occur on the broadest level of social organization involving large geographically scattered populations. Having emphasized microsystem levels in the past, rehabilitation researchers are now

required to move to mezzo-and macrosystem level analysis for a more adequate understanding of the multiple variables affecting rehabilitation outcome. Evaluation research models by necessity must be developed for these various levels if additional understanding and knowledge are to be obtained.

Fairweather (1967) offers some guidance regarding the specific external factors which should be incorporated into a more comprehensive model. The more potent of his suggestions are social climate, geographical location, social mores and folkways, and economic conditions. For disabled populations it would also be necessary to include institutional policies and practices.

The identification of various external factors influencing the evaluation research process and the measurement of their effect on outcome is not an easy task. However, examining the mezzo- and macrosystem factors in more detail may result in some useful, if crude, measures.

Social Climate: Multiple factors are involved in assessing the social climate in which a social service system such as vocational rehabilitation functions. However, as an example of assessment of one aspect of social climate, the degree of acceptance of the rehabilitation process and system can be assessed at both mezzo- and macrosystem levels. Through carefully conducted surveys, levels of acceptance by

various groups can be assessed. This would provide additional insight into the problems and obstacles to be encountered in communities and larger geographical areas. If a generally supportive social climate exists, the realization of positive rehabilitation outcomes is obviously enhanced. Other areas for further research and study include attitudes toward the disabled (particularly attitudes relating to employability), attitudes toward roles such as husband/wife, cultural and religious biases, and the sympathy-apathy-charity public response triad.

Geographical Location: Obviously attention must be given to the similarities and differences which exist among and between rural-urban, county, state, city, or national areas. The effect of geography including climate, topography, and other variables has been recognized in cultural, social, and economic development. It is also an external factor which should be identified and assessed in terms of its effect upon rehabilitation outcome.

Social Mores and Folkways: The variability among subsystems within a larger geographical area in terms of social mores and folkways is considerable and could account for some outcome variability. For example, the mores of simple hill folk in regard to

the disabled may vary greatly from those of more sophisticated urban dwellers. This variability has often been lightly dismissed because of the assumed randomization of such factors across the population. However, it is an external variable which requires greater attention, as well as a more comprehensive assessment of the effect such mores have on rehabilitation outcome.

Economic Conditions: This is one of the more potent external factors, and one for which a considerable amount of data exist at both mezzo- and macrosystem levels. Attention should be directed toward assessing the effects of employment (or unemployment) rates, technological change, the rate of economic development (or underdevelopment), and the production/profit vs. value of the individual attitudes of employers. One rather obvious problem is that the differential effect of unemployment on minority groups such as Blacks would likely apply to the disabled, thus creating placement difficulties. Close scrutiny of economic indicators and the effect of such indicators on rehabilitation outcome is crucial to the evaluation process.

Institutional Policies and Practices: The complexities of existing private and public institutions have been

recognized as influencing the success of the rehabilitation process. However, the effect of such "competing" public policies as social security disability insurance and vocational rehabilitation programs has not been clearly documented. Although complementary in many ways, vocational rehabilitation and social security programs offer competing opportunities (vocational rehabilitation leading to independence on one hand versus social security emphasizing financial dependency on the other). These policies have a widespread macrosystem impact upon rehabilitation outcome since they exist nationwide.

In summary, several issues have been reviewed for consideration in conducting program evaluation research. These were:

(1) Rehabilitation outcome is determined by a complex group of factors including participating clients, treatment staff and services (the rehabilitation system), and factors external to the system.

(2) Microsystem research on the client or small groups has resulted in findings which account for only a limited portion of the variability in rehabilitation outcome.

(3) A systems approach will provide a better model for conceptualizing the boundaries of program evaluation research and for expanding those boundaries to

mezzo-and macrosystem levels.

(4) At mezzo- and macrosystem levels the identification of factors external to the rehabilitation system and the measurement of such factors should and can be realized.

(5) Specific external factors were identified and discussed. These were limited to social climate, geographical location, social mores and folkways, economic conditions, and institutional policies and practices.

References

Bolton, B. The prediction of rehabilitation outcomes. Journal of Applied Rehabilitation Counseling, 1972, 3, 16-24.

Fairweather, G. W. Methods for experimental social innovation. New York: John Wiley and Sons, Inc., 1967.

Mullen, E.J., & Dumpson, J.R. Evaluation of social intervention. Jossey-Bass Inc., 1972.

ISSUES IN INITIATING LARGE SCALE REHABILITATION RESEARCH*

Richard Roessler, Ph.D.
Arkansas Rehabilitation Research and Training Center
University of Arkansas

Introduction

The purpose of this paper is to describe problems encountered in conducting large scale rehabilitation research studies. In particular, comments in the paper refer to the Services Coordination Project currently underway at the Arkansas Services Center in Jonesboro, Arkansas. The Arkansas Rehabilitation Research and Training Center has contracted with Social and Rehabilitative Services of Arkansas to evaluate the coordination project.

In developing and administering the evaluation of the project, research staff have identified a number of social and political forces that affect the conduct of the research. The remainder of the paper attempts to describe these areas of conflict and to recommend methods of resolution.

Specifying Objectives

The first place that evaluation research interacts with the real world is in the setting of objectives and goals for the social impact program.

---

*This paper is based on a presentation at the American Personnel and Guidance Association Convention in New Orleans, April 11, 1974.

As Weiss (1973) points out, projects often have difficulty translating Federal legislation into program activities because legislation specifies neither the operations involved in reaching goals nor the evaluative criteria for determining whether goals have been reached. The legislative base for the Services Coordination Project, the Allied Services Act[1], itself does not clarify such terms as coordination, independence, dignity, and, to a lesser degree, self-sufficiency.

Compared to What?

Unclear objectives in legislation create an interesting projective test for State, project and evaluation personnel. In terms of Services Coordination, State and project personnel saw the charge of the Allied Services Act as one of defining and implementing a management system to coordinate social and rehabilitative services.

The evaluation staff, on the other hand, was concerned with questions regarding effectiveness and efficiency, e.g., is the experimental service program superior to the traditional one? In other words, to determine if services coordination is an improvement, one must ask, "Compared to what?"; obviously, "Compared to nonintegrated service delivery."

---

[1]Presently under consideration in the Senate.

Specifying ways in which services coordination should be an improvement over nonintegrated services involves defining social change criteria (Fairweather, 1971). According to Fairweather, social change criteria represent a consensus among interested parties regarding what problems the program should solve.

Since such consensus rarely exists, evaluators end up initiating a process that should have started with Federal legislation, that is, developing appropriate comparisons in terms of relevant social change criteria.

Asking such questions as "Compared to what?" does not endear researchers to those who have the responsibility for initiating and maintaining a system. Indeed, for project maintenance purposes, the only real need for evaluation is to answer formative or management oriented questions. For example, in the Services Coordination project, State and project staff wanted the evaluation team to compare alternative management procedures for carrying out coordinated service delivery.

Evaluation personnel, however, felt that the first priority was to demonstrate a need for change in current practices. Establishing the need involved those effectiveness and efficiency questions, which require system comparisons in terms of client outcome and case process.

Due to differing perceptions of evaluation, discussion went back and forth within the project as to the

best way to proceed with the evaluation. Resolution of the question of management versus outcome evaluation came during a joint meeting with Federal, State, project, and evaluation personnel. The balance was tipped when Federal project officers emphasized the importance of the effectiveness and efficiency issues; Federal representatives were less interested in "fine tuning" and more interested in comparative outcome data.

At first, the time invested in the dialogue between the evaluation staff and project and State personnel seemed wasted effort. But, as a result of the dialogue, the evaluation team recognized the need for a balanced evaluation study. The final evaluation design for Services Coordination, as a result, included studying client outcome in coordinated and noncoordinated service delivery and studying the differential effects of alternative management procedures to carry out case handling in services coordination.

Relationship of Evaluation Process to Project

Differing perceptions of evaluation illustrate the need for greater communication and coordination among Federal, State, project, and evaluation representatives. However, such communication raises a question regarding the relationship of evaluation to the overall project. In the past, evaluation literature has stressed the need

for objectivity; the evaluation team should avoid contacts that might contaminate its objectivity. If necessary, the team should speculate regarding program objectives rather than communicate directly with project staff (Freeman and Sherwood, 1965).

Though objectivity is not overlooked, current trends emphasize impact and utilization of evaluation, which necessitates a relationship between evaluation and project efforts. Indeed, Freeman and Sherwood (1965) contend that the evaluator is socially responsible for working with social impact progra... - the development of measurable objectives and appropriate program practices.

Experimental Treatment: Definition and Acceptance

Involvement of evaluation personnel or appropriate consultants in the development of project objectives and activities helps overcome another problem often cited in evaluation research, that of a constantly changing experimental program. If the program is constantly changing, it becomes increasingly difficult to determine the nature of the experimental treatment; i.e., to identify program features that might have had a positive effect.

Initial planning for the Services Coordination Project involved extensive use of a management consultant to develop an operational management system. The consultant did his job well and, from the viewpoint of the evaluation

team, developed a system, that could, in fact, coordinate tne delivery of social services. A clear experimental treatment had been designed.

The irony of the project was that the operational management system was slow to move into full operation. The delay seemed to result from (a) the perceived threat of coordinated services, (b) the lack of perceived need for services coordination, (c) the lack of involvement of service delivery people in the planning of the management system, and (d) the lack of authority of the project over service delivery personnel.

The time spent in defining project activities and in implementing those activities has a tremendous impact on evaluation. For research on Services Coordination, the chief loss was in terms of data collection time for pre- and post-testing.

With better planning, the time lag might have had some value to the project. For example, the period between program development and program implementation could have been used to gather baseline data regarding current client outcomes in noncoordinated services.

However, eagerness to start the project resulted in tentative coordination efforts that contaminated the setting. Service providers realized that traditional service delivery was no longer in existence, although they were not always sure what had replaced it.

Eagerness to begin is certainly a positive attribute of any project. But, it, too, can be affected by some of the same political and social forces that affect evaluation. For example, the Services Coordination project has been hampered by perceived shifts in commitment on the Federal level to coordination of service delivery. Such shifts are manifested in the on again/off again funding rumors that have a definite effect on project morale.

Summary

In summary, numerous factors, beyond those technical ones of research methodology, impinge on evaluation research. In terms of the Services Coordination project, these factors manifested themselves in the following problems:

1. The problem of starting with objectives stated in political or, at least, nonoperational terms.
2. The problem of multiple interpretation of ambiguous objectives.
3. The lack of clearly defined social change criteria.
4. The conflicting expectations of evaluation for either management (formative) or outcome (summative) studies.
5. The problem of evaluation meeting its social responsibility while at the same time maintaining its objectivity.

6. The problem of a constantly changing treatment effect that makes it difficult to identify causes of program outcomes.
7. The problem of program acceptance by service providers.
8. The problem of project eagerness as it effects the collection of baseline data.
9. The effect of changing political trends on program support and, consequently, on morale of project personnel.

Responding to problems of research in an applied setting requires reviewing aspects ranging from Federal legislation to project enactment. Some recommendations for evaluation in social/political settings are:

1. Evaluation specialists should participate in drafting Federal legislation to accomplish the following:
   a. Specify objectives in operational terms.
   b. Identify relevant social change criteria.
   c. Establish procedures to ensure that all proposals include appropriate objectives and criteria.
   d. Develop review methods for projects so that long-term funding can be committed contingent on specific project activities and behaviors.

2. Joint objectives writing sessions involving grantees (State and local) and intended or outside evaluation personnel would help individual programs establish their goals in measurable terms.

3. In establishing an evaluation program mutual discussion among Federal, State, project, and evaluation personnel is needed regarding:

   a. Evaluation's role and responsibility in the project.

   b. The nature of social change criteria for the study and types of acceptable evidence for project success.

4. Service providers need to be involved in the development of the project in terms of stating objectives and developing program practices to meet those objectives. Involvement will build understanding of the need for the program and commitment for carrying out program policies.

5. Evaluators need to develop balanced evaluation studies including both outcome and management data, for example, comparing an innovation with present practice and experimenting with additional ways to implement innovative practices.

6. Baseline data should be collected before instituting innovative practices. Outcomes of ongoing programs can then be compared with those resulting from new programs.

Selected References

Fairweather, G. Methods for experimental social innovation. New York: J. Wiley and Sons, 1967.

Freeman, H. & Sherwood, G. Research in large-scale intervention programs. In Readings in evaluation research, Caro, F. (Ed.). New York: Russell Sage Foundation, 1971, 262-275.

Suchman, E. Action for what? A critique of evaluative research. In The organization, management, and tactics of social research, O'Toole, R. (Ed.). Cambridge, Massachusetts: Schenkman Publishing Company, 1971, 97-130.

Weiss, C. Between the cup and lip. Evaluation, 1973, 1 (2), 49-55.

REACTIONS ON EVALUATION*

Jerold D. Bozarth, Ph.D.
Iowa Mental Health Authority
Department of Psychiatry
University of Iowa

One thing which evaluators do very often is to assume that everyone knows the difference between evaluation and research. The difference between definitions of "evaluation" and "research" often cited are a) research is directed toward adding new knowledge while evaluation is directed toward noting information associated with a particular program; b) research utilizes high levels of "good" design, including randomization, control groups, etc., while evaluation is usually unable to establish such designs; and c) research is longitudinal and evaluation is short-term. Many of these differences do not, in fact, need to exist. In an attempt to throw light on both the state of art in evaluation research as well as examine the relationship between "research" and "evaluation", a number of points seem pertinent. These are:

1) "Good" evaluation is the same as "good" research. That is, the evaluation research design should have good

*This paper is based on a presentation at the American Personnel and Guidance Association Convention in New Orleans, April 11, 1974.

internal validity allowing confidence in the results and good external validity allowing for the practical application of the results within specified populations. Unfortunately, most evaluation is at the lower end of the qualitative continuum in respect to validity considerations such as statistical design, control of variables, statistical power and availability of cross-validation. Yet, it is quickly adapted for administrative actions; especially as a substantiation of administrative designs. Hence, evaluation is probably more harmful than helpful the way it is usually practiced. For example, one evaluation report that intelligence quotients could be improved by 100 points by altering environment was cited in a federal report--slightly misinterpreted-- and latched onto by national and international media, producing strong incentives for legislators to act in certain ways. The lack of validity of the findings, as well as the misinterpretation by some myopic bureaucrats had to be pointed out before there was even an awareness of great perpetuation of error.

2) Evaluation research should be relevant. If evaluation research is to be relevent, three question areas should be considered. These questions are:

a) Questions which are asked by legislators, boards of supervisors, the general public, etc. These

questions must often be reworded by the questioner, with the help of the evaluator.

b) Questions which can assist particular programs to modify and revamp their effort towards more effectiveness.

c) Questions which bear on both a and b above and also on basic research to add new knowledge to questions of rehabilitation counselor effectiveness. This is basically information which can be applied to the broad question of "What kinds of counselors providing what kinds of services to what kinds of clients are being effective in what kinds of ways?"

3) Evaluation generally has an accountability set which is, in the opinion of the present writer, a set which produces only threat to, evasion of, game-playing-by, and resistance by the project personnel being evaluated.

4) The data gathered via evaluation are seldom used, and often cannot answer the original questions. The questions which usually initiate evaluation components of programs are generally value-type questions, rather than measurable-type questions. In addition, evaluation research efforts, by-and-large, lead to a collection of invalid, superfluous data which are seldom related to the original questions anyway. A common response of at least a number of federally funded mental health centers which are required to turn

in numerous accountability data is, first, they don't know how the information is used; second, the reports which are produced are of no reasonable value; and third, information about essential questions is nearly impossible to retrieve.

This should not be interpreted as a plea to eliminate evaluation. It is a plea for a focus on certain considerations. These considerations are: 1) that "basic" research must be included in the evaluation of programs. It is the long-term cross-validation and replication which is the crux of advances in rehabilitation programs; e.g., research in the pecking of pigeons (operant conditioning) eventually led to behavior modification approaches to help people. Likewise, extensive research on variables related to effective psychotherapy has led to training programs for rehabilitation personnel, mental health center personnel, nurses, etc.; 2) that questions originally posed by legislators, administrators, etc., need to be altered so that they are posed in ways which lend themselves to m[redacted]rement. Like doctoral dissertations, one of the majo[redacted]blems is finding and stating the questions. The remainder of activity to follow an adequately stated question is simply a complex technological footnote; and 3) that evaluation ought to

be a "helping" set, rather than determination of the worth (funding?) of a program. The set ought to be one which assists programs to identify their weaknesses and strengths and modify accordingly. An example of this approach occurred in a small, unpublished study which took place in a mental health treatment facility. Some measures of client goal accomplishment were developed, and implemented as one criterion. Over the span of a year, it was found that better than sixty percent of the clients served depicted zero improvement on the criterion measure. With an accountability set, this finding would have been quite threatening (and resisted on the basis of being a gross study, having only one criterion, using poor evaluation procedures, etc.). However, part of the evaluation included a more thorough examination of preliminary findings and was geared toward helping the center improve services. The further examination indicated that, on the one hand, a high percentage (90%) of clients of some of the therapists were improving, while a small percent (like zero) of the clients of other therapists were improving. On the other hand, it was found that clients receiving certain services were not improving while other clients receiving other services were split in their improvement, depending upon the therapist with whom they had contact. This was an interesting finding, in that it corresponded closely to extensive reports of differential

effectiveness in psychotherapy literature. It was also a useful administrative bit of information, largely because of a "helping" set rather than an "accountability" set. The administration of the center was able to take the findings of clients' lack of improvement to the individual therapists confidentially, with the idea of looking further at this phenomenon with the therapists. It was realized that this was only one criterion, but still worth consideration by therapists, since clients of some of the other therapists were improving on that same criterion. Such determinations were eventually established as that of one therapist changing assignments and feeling relieved about the change. Another therapist was able to change clientele and begin working with a less threatening clientele. The situation had cast the therapist into a necessary work load of clientele which posed problems for that therapist. The therapists and the center eventually benefitted from the initially threatening (in an accountability set) evaluation finding.

The administration also identified, at least, one of the services as being inadequately structured. This was the service of group therapy, which the evaluation had revealed had few clients improving. The apparent reason was that there was no core group in the therapy sessions; that the clients might come for one or two weeks and never

return, and then a complete cycle of clients participated in the groups in this way. Alterations to include more core structure to the groups was thought to be a way to alter the apparent ineffectiveness of the group therapy program.

All in all, evaluation is viewed by the present writer as being functionally weak and probably perpetuating error. Evaluation, however, can be strong. It can be helpful. It can aid in the strengthening of programs.