THE USES OF PERFORMANCE MEASURES IN REHABILITATION PROGRAMS

Jeff Harris

August 1973

Working Paper No. 215/RS022

Frederick C. Collignon, Project Director - Michael B. Teitz, Principal Investigator
Project for Cost Benefit Analysis and Evaluation of Rehabilitation Services

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

PREFACE

As the demand by taxpayers, client groups, and legislators for program evaluation continues to grow, the importance of developing better measures of program impact and performance and of creating incentives for organizations to increase their effectiveness and efficiency is becoming more widely recognized. In the vocational rehabilitation field, the indicators of program performance and impact have been primarily drawn from the information system in the program which monitors the client's movement through the program and ultimately into a job or homemaking. The movement of the client through the system and the final characterization of services as successful or unsuccessful draw heavily upon the arbitrary judgments of professional counselors. Calls for reform of the indicators have been heard for many years in the rehabilitation field, but few specific suggestions have come forth. More recently, policymakers, backed by legislation, have begun exploring the use of weighted case closure and benefit-cost indicators for program performance. The underlying social values behind such measures have seldom been systematically considered, however. Nor have the impacts of the use of such indicators upon the behavior of counselors in the rehabilitation system been considered.

In this paper, Mr. Jeffrey Harris analyzes in depth the use of indicators generally in social program evaluation and specifically in the rehabilitation field. He evaluates both the current indicators being used in vocational rehabilitation and the various alternative

indicators which have been proposed. Finally, he examines the organizational problems in introducing new indicators and recommends strategies for such organizational innovation.

We believe that this paper will prove useful not only to those concerned with program management and evaluation in rehabilitation, but also to those generally concerned with evaluation and the design and use of indicators. This paper draws from the thesis submitted by Mr. Harris for his Masters degree in City and Regional Planning. The paper also reflects Mr. Harris' program experience as Assistant Director of the nationally acclaimed Model Cities Program in Alma, Georgia, and his consulting and field work with the State of California Department of Rehabilitation.

Frederick C. Collignon
Michael B. Teitz

ACKNOWLEDGEMENTS

Only some of the intellectual debts I have incurred in preparing this paper can be reflected in formal citations.

Appreciation is due to Fred Collignon not only for his patient and perceptive criticisms, but also for providing me with a view of the world and a role model that begin to traverse the gap between analysis and involvement. Michael Teitz's thoughtful questions and observations have helped me on several occasions to create sense and synthesis where both were lacking.

Virtually all of the graduate students participating with me in the "Project for Cost-Benefit Analysis and Evaluation of Rehabilitation Services" at U.C. Berkeley have contributed in some way to this paper -- if only by a casual comment that has since generated a long line of thoughts.

Over the past few months I have interviewed a number of individuals, many of them employed by the California State Department of Rehabilitation. The list of their names at the end of this paper is inadequate compensation for their time and for the valuable insights that each one managed to convey to a non-practitioner.

Partial support for the research reported here was provided by a training grant from the Rehabilitation Services Administration of the Social and Rehabilitation Services, U.S. Department of Health, Education, and Welfare.

An earlier version of this paper was submitted as a Master's Thesis to the Department of City and Regional Planning, U.C. Berkeley, in June, 1973.

TABLE OF CONTENTS

ABSTRACT

The Uses of Performance Measures in Rehabilitation Programs

The growing interest in the evaluation of social programs in general has also been evident in the Federal-State vocational rehabilitation program. Here, though, the basic tasks of measuring program output and evaluating performance are already well underway. Current problems are concerned more with refining the existing measure, the "26-closure," to deal with problems of accountability, choice, and control that have accompanied major expansions of the program's scale and scope over the past decade.

Performance measures, along with other modes of monitoring and evaluation, can be used in rehabilitation programs to address the issues of: who is served, how well, and with what final results (that are attributable to the program).

There are four major roles that performance measures can potentially play in a rehab program:

(1) providing information for program improvement (the traditional role);
(2) providing information for outside justification and "program defense";
(3) serving as a means of direct administrative intervention (through the creation of incentives); and
(4) helping to raise new issues, educate participants, and constructively channel conflicts.

The existing measure, the 26-closure, gives the same credit for all "successful" rehabilitations -- regardless of the severity of the disability, the amount of effort and resources expended, or the actual degree of benefit to the client. As a result, many have objected to the 26-closure as a performance measure because it provides inadequate information and creates dis-incentives for serving cases that are relatively more difficult, time-consuming, or less certain of success.

Chapter 2 provides a brief overview of current practices in monitoring and evaluation, primarily at the State level of the rehabilitation program. This is followed in Chapter 3 by a literature review of proposals for modifying or replacing the 26-closure as the program's basic performance measure, including several suggestions for "weighting" 26-closures and constructing aggregate indices. These proposals are critiqued in terms of their practical consequences if introduced in an

agency, as well as on the grounds of their technical design. The general conclusion is that none of the proposals advanced thus far is likely to prove effective, on its own, either as a means of providing more accurate information or as a way of deliberately altering existing incentives in the intended directions.

Some understanding of the structure and control mechanisms of a rehab agency is necessary, in order to evaluate the possibilities and limitations of using performance measures for control and positive change. Toward this end, Chapter 4 provides a brief review of some relevant literature on organization theory. It then summarizes the limited empirical research on power and decision-making in rehab agencies. The chapter concludes with a review of potential means of intervention in the program, at both the policy and the administrative level.

The final chapter attempts to synthesize the preceding discussions of: (a) monitoring evaluation activities and (b) intervention and control functions. It strongly suggests that for complex social programs these two functions be integrated into a "guidance process."

In the case of rehabilitation, key features of such a guidance process are that it be decentralized and pluralistic -- that it include a wide range of participants in the processes of analysis and response, along with multiple indicators of the program's performance. At the same time, the new guidance processes should look primarily to the field (including both counselors and clients) for sources of innovation. A greater degree of risk-taking and diversity should be rewarded, and learning and adaptation encouraged at all levels of the agency -- rather than strict adherence to pre-planned, centrally defined policies.

The paper concludes with a brief look at one of the possible futures for the rehabilitation program: increasingly close ties, and perhaps consolidation, with other manpower and social service programs at both State and Federal levels. Further developments in this direction could take a variety of forms, bringing with them new opportunities as well as constraints. A process of program evaluation that is strengthened along the lines suggested here may be of significant help in resisting the constraints and seizing the opportunities.

CHAPTER I

INTRODUCTION

THE PROBLEM OF MEASURING PERFORMANCE

Context: the rise of program evaluation. Trends in rehabilitation programs. Alternate views of performance measures; definition; sources of discontent. Potential roles for performance measures: program improvement; program defense; administrative intervention; issue-raising and education. Scope of the paper: biases and limitations; format.

Context: The Rise of program Evaluation

> In reaction to the "exuberance of the 1960's," with its profusion of overlapping and sometimes contradictory efforts to create a more equitable and productive society, the pattern of social programs in the 1970's must reflect the cautious, deliberate development and testing of new strategies to determine their effectiveness, before they are launched full-scale. At the same time, existing programs should be subject to a thorough re-assessment of their actual performance -- not just their promises or potential. This will mean a rigorous questioning of familiar assumptions and an impartial scrutiny of program effectiveness, measured objectively against the resources society has invested. . .

Or so runs the sharp-pencilled litany of the proverbial policy analyst these days. Given that, in practice, the "impartial analyses" have often served simply to buoy sinking programs or (more recently) to torpedo them, rhetoric such as the above may seem somewhat naive. Nevertheless, there is little doubt that the habit of "systematic thinking" (Rivlin, 1971; Schultze, 1968) is continuing to spread throughout the public sector -- though haltingly and with considerable trauma in some quarters.

This broader trend toward the explicit analysis of program performance is part of an older tradition, however, in the Federal-State vocational rehabilitation program. Throughout the program's fifty-year history the separate agencies in each State, funded and rather loosely guided at the Federal level by the Rehabilitation Services Administration (RSA),* have progressed further than most other manpower and social service agencies toward determining their own goals and mandate, developing an array of service technologies that on the whole seem to "work," and defining at least some basic measures of program output that can be used to judge and compare performance among agencies or over time.

The current thrust towards improving program evaluation, then, takes a different form in rehabilitation than in most other social programs. In rehab programs, the basic problem is not how to begin measuring performance, but rather how to refine and broaden the scope of the performance measures that are already well-established. The questions raised are not only technical ones, but organizational and political as well:

> How are current performance measures actually used (or misused) in State rehabilitation agencies?
>
> What changes are needed, in the performance measures or in the evaluation and management processes in which they are inevitably embedded?
>
> What might be some of the unintended consequences of these changes?
>
> What <u>should</u> be the roles of performance measures in the ongoing operation of rehab programs?

Each of these questions will receive some attention in the pages to follow.

---

*Formerly the Vocational Rehabilitation Administration.

The purpose of this paper, then, is to consider the relationship of performance measures and evaluation to the larger tasks of effectively managing and changing a rehabilitation program. Problems with the current measures of program performance will be reviewed, but recent proposals to replace or supplement them will also be critically examined. The focus will be less on the technical features of such proposals than on their possible consequences if actually introduced in an agency. Finally, some suggestions will be offered for the roles that performance measures can play in rehabilitation programs, and some implications of this for their design and use.

Trends in Rehabilitation Programs

Many of the current issues in monitoring and evaluation stem not from the poor performance of rehab programs, but from their widely acknowledged success in serving the needs of the disabled. The program's basic service strategy, the "rehab model," aims at the delivery of flexible, comprehensive, individually-tailored services that focus on a relatively well-defined goal for each client. This general approach has proven effective enough to interest other agencies in adopting at least some of its features.

At the same time, recognition of the success of rehabilitation programs has generated pressures to expand both their scale and scope, by: (a) providing increased funding to serve a greater proportion of those in need, (b) broadening eligibility requirements to allow new client groups (or new segments of existing groups) to be served, (c) expanding the range of services offered, and (d) accepting new types of client benefits as evidence of rehabilitation "success."

Some simple statistics reveal the magnitude of growth that has already occurred in the program (U.S. RSA, 1972). In the early years, the annual total of individuals rehabilitated grew slowly, amounting to only a few thousand nationwide during the 1920's. The total jumped noticeably during World War II, but didn't reach the 100,000+ mark until 1962 -- forty years after the program began. Over the next decade, Federal legislation (most notably the 1965 amendments) increased the overall level of program funding, provided more generous matching provisions for the States, and broadened the program's mandate for whom to serve and what services to provide. It took only six years for the annual number of rehabilitants to increase another one hundred thousand (to 207,900 by 1968), and then slightly over three years to reach the third hundred-thousand (326,100 in 1972).

The growth in program expenditures has been even more dramatic. From an initial level of only $0.3 million in 1921, the program's expenditures nationwide now total over $600 million annually (U.S. RSA, 1970). Once again, the most rapid growth occurred during the 1960's, with a doubling in expenditures from 1960 ($78.7 million) to 1965 ($154.1 million), and then a further three-fold increase by 1969 ($455.9 million).

At the same time, the mix of clients served and of those actually rehabilitated by the program was changing in response to the new opportunities and pressures mentioned above. For example, in 1958 only 2.1% of all those rehabilitated nationwide were identified as mentally retarded; ten years later this had increased to 10.7% of the (much larger) total. Similarly, the mentally ill represented only 5.0% of the rehabilitated cases in 1958. Their share of the growing number of

rehabilitants had nearly doubled (9.8%) by 1963, and then redoubled in another five years (19.6% in 1968 -- U.S. RSA, 1970).

This expansion and broadening of the program, however, have raised new problems. These include the usual difficulties of managing any major program under conditions of growth and change: the need to locate or train enough administrative and professional manpower,* and to maintain an acceptable level of service quality and personal attention. But in rehabilitation there have been other problems, as well, such as the inadequacy of existing service strategies for dealing with some of the newer categories of disabled clients (for example, the mentally retarded, those with "behavioral disorders," and individuals with severe or multiple handicaps). For many of these individuals, the traditional pattern of services (medical restoration, preparation for employment, job placement, and closure) appears to be inappropriate. There may be several reasons for this: either the goal of competitive employment is not a realistic one, the services provided by the agency fall short of meeting the client's needs, or in some cases the basic assumption that a client's problems can be "solved" and then services terminated is simply not meaningful.

A further problem has been that as the rehab program expanded, it inevitably (and partly by intent) began to serve more cases that were relatively difficult, time-consuming, and expensive. It seems reasonable to expect that rehab, like most social programs, would experience declining cost-effectiveness as it expanded to cover more of the target population -- and thus harder cases (Rossi, 1969). In fact, though, the

---

*Sexist terms such as this will be used throughout this paper only for lack of alternatives.

expectations of legislators seem to have run in very different directions. They have generally expected (numerical) gains in program output to be at least proportional to budget increases.

In the case of rehabilitation, the fact that these expectations may have been created, or at least reinforced, by the enthusiasm of the program's own supporters has proven of little comfort. In most States the annual number of rehabilitants has not risen in proportion to budget increases, nor have outside critics been sufficiently educated to the changing nature of the program's goals and clientele -- or to new ways of measuring these changes. The pressures for increased "production" remain.

In spite of the growth in the budget and caseloads of rehabilitation programs nationwide, the estimated backlog of those considered eligible and in need of services has grown even more rapidly -- due to an aging population, a broader recognition of need and understanding of the nature of rehabilitation, and an easing of the eligibility requirements themselves. Using one estimate of the "pool of need" (U.S. RSA, 1972), vocational rehabilitation programs in all States together served fewer than 10% of the eligible disabled population in 1972, and actually rehabilitated fewer than 3%. Using other, larger estimates of potential need, these figures are even less "optimistic."

Even if more funding were to become available, immediate expansion of the program would be subject to other constraints: the shortage of trained counselors; limited capacity of the rehab facilities and private vendors that actually provide, under contract, many services to clients;* and the shortage of competitive or sheltered job

*See Markowitz (1972), and Nat. Rehab. Assoc. (1964), ch. VI.

placements for clients once they have completed their program of rehabilitation.

In general, expansion in the scale and scope of the rehabilitation program has generated strong pressures for greater accountability, choice, and control. The program not only has to demonstrate increased results, but also to make and enforce a greater number of choices of whom it serves, in what ways, and towards what goals or criteria of success. These choices are sledom based on explicit, publicly-stated policies. Instead, they are the unplanned consequences of diffused and decentralized patterns of program administration and service-delivery, that allow considerable autonomy to individual counselors to make what amount to basic policy decisions. Except in the broadest terms, there is little consensus on the program's goals and priorities for service -- or on who should have a say in setting these priorities.

While the program's dispersed pattern of administration and its strong tradition of counselor professional autonomy have definite advantages, they also tend to weaken and fragment any attempt at centralized policy-making or hierarchical control of the program. Rather than suggesting that these tendencies be fought in conventional terms, by seeking more effective means of centralized control toward pre-determined outcomes, an alternative strategy will be developed later in this paper. It will concentrate on finding ways to disperse the responsibilities for information, learning, and program guidance among all levels of the rehab agency; to raise new policy issues; and to involve new participants in their solution.

A final point to be emphasized is that rehabilitation, along with many other social programs, will be operating in the years ahead in an

increasingly uncertain environment. Current signs of this can be found in the efforts to decentralize major Federal-State programs, the accompanying (and perhaps conflicting) trends toward agency consolidation, the recent cutbacks and ceilings imposed on Federal expenditures for social programs, and in what appears to be a spreading philosophy of re-privatization for many forms of social intervention that are themselves of recent origin. All of these suggest that future planning for rehabilitation programs should be flexible and "contingent," avoiding heavy commitments to any single strategy for evaluation and program guidance that may prove inappropriate under changing conditions.

Alternate Views of Performance Measures

Definition. While this paper will touch on a range of issues involved in evaluating and managing rehab programs, the emphasis remains on one particular "mode" of analysis, performance measures. As the term is used here, performance measures differ from other sources of evaluative information in that they tend to be: expressed in quantitative terms, amenable to "external" (non-participant) observation, and relatively efficient for mass collection and analysis. A further characteristic of performance measures as defined here is that they generally are not collected for a single study or to answer one specific question, but are more commonly gleaned from the day-to-day operating records of an agency. While this may make them less "obtrusive" as measures of performance (if, indeed, this is desired -- see below), it also means that they will seldom provide the ideal data for responding to any given policy question.

Performance measures may focus on program processes, program outputs, or the actual impact of the program on its target group or its initial goals. A distinction is made between outputs and impact because the rehab program is seldom the only force acting on its clients. Other "environmental factors" may limit or undo the positive effects of the program -- or in some cases reinforce or overshadow them. Either way, measures that distinguish between program output and impact can begin to raise a whole new set of issues concerning the attribution of observed change to the program, and the possibility of shifting resources outside the formal rehab program to achieve the same goals.

Any comprehensive evaluation should deal with at least two broad sets of issues in rehabilitation programs:

(1) Who is served by the program? What proportions of the various sub-groups of clients (defined along a number of dimensions, not just disability type) end up as successfully rehabilitated? What are the actual selection processes at work, whether deliberate or unintended, internal or outside the agency?

(2) How are clients served, how adequately, and with what results? Are services comprehensive, thorough, and appropriate to the client's needs? Are they reasonably efficient -- considering not only direct program costs but also indirect costs, to the client and to society in general? To what extent are observed gains to the client, his family, and the community actually attributable to the program? How stable are these initial gains over time?

The collection and analysis of performance measures can be distinguished from other modes of program monitoring and evaluation, each of which may be appropriate under certain circumstances, and each producing somewhat different types of information. Examples of these other modes include:

--case-review and "management audits" - an examination of program records for conformance to legal requirements or professional standards of service-delivery.

--field observation, site visits, informal interviews - to gain a "feel" for the processes of service-delivery and program management.

--review panels, conferences, and other forms of direct participation - to exchange information and points of view, working towards informal consensus or formal action.

--experimental research - controlled studies to determine cause-effect relationships.

--full-scale program evaluation studies - to develop (for example) system models of program structure or process, overviews of client or staff attitudes, or definitive statements of program impact and cost-effectiveness (for examples, see Abt Assoc., 1972; and I.I.S., 1971).

--intuition, judgment, experience - preferably part of each of the above modes, but also used to link and interpret them and as the most common means of responding in day-to-day decision situations.

All of the monitoring and evaluation modes listed above can be distinguished from performance measures in being neither routine nor generalized -- each is a specific activity undertaken in response to a (presumably) well-defined problem. In one important sense, these other modes can complement the ongoing collection and analysis of statistical performance measures. The performance measures can uncover a potential problem, but then the other, more intensive modes of analysis are needed to specify its exact nature and dimensions, its probable causes, and the best means of responding to it.

One reason why this paper concentrates on performance measures rather than other modes of program evaluation is that performance measures offer a relatively efficient means of "scanning" a complex program, and thus represent a useful first step in many program analyses.

In addition, the subject is one that can be discussed usefully in analytical terms, while the other evaluation modes are perhaps best dealt with in the same way they are most often learned: through intuition and direct experience.

Sources of discontent. Currently the basic measure of performance in rehab programs is the total number of cases that are closed each year as "successfully rehabilitated" (termed "status 26" in the standard case-recording system). As will be described in further detail in Chapter 3, many people are dissatisfied with the current measure. These concerns usually fall under one of two broad categories, poor information or wrong incentives.

The information produced by the 26-closure measure is often seen as inadequate, misleading, or unfair. Since it counts all cases the same, the 26-closure by itself cannot distinguish between "difficult" cases and "easy" ones; between services that are comprehensive and high-quality or nominal and inadequate; between clients whose needs have been effectively met at closure and those who are still far from realizing their full vocational or personal potential. Because the 26-closure does not reflect differences in service quality or caseload difficulty, administrators cannot rely on the total number of closures to realistically compare the performance of field units or individual counselors. Finally, as long as outside observers of the program focus their attention on the total number of closures, they tend to overlook significant long-term changes (or the failure to change) in the quality of services, the difficulty of cases, or the mix of clients being served by the program.

A second source of concern is that the 26-closure, in conjunction with numerical "quotas," tends to produce undesirable incentives among

rehabilitation counselors and others in the agency. In order to achieve his annual quota of 26-closures, a counselor may be encouraged: (a) to accept a larger proportion of "easy" cases (or those which appear to be more certain of success); (b) to emphasize services of a short-term nature; or (c) to close a case at too low a level of (vocational) attainment for that client -- or else too quickly, before it is clear that employment and other gains can actually be maintained.

While these objections have been directed at the 26-closure itself, the larger question of how to measure program performance has been approached in a number of ways: as a technical problem, a systemic one, or a problem in organizational and political terms. Each of these points of view leads toward a somewhat different set of proposed solutions.

On a technical level, the problem of performance measurement is seen as either one of faulty design of the measure itself, questionable reliability or completeness of data provided from the field, or inadequate analysis. In any case, the solution is assumed to lie in better technique.

A systemic view of the problem, on the other hand, would see inadequacies in the data or its analysis as resulting not primarily from poor instruments, techniques, or concepts, but from: a lack of time or incentives to provide data from the field, incentives to "fudge" the data, distortion arising from the communication process itself, or the different perceptions that analysts and practitioners have of the system and of what information is relevant.

The organizational-political point of view might raise many of these same points, but would also introduce a concern for what happens after information is collected and analyzed. Ineffectiveness or

inaction at this stage would also be considered part of "the problem" of performance measurement.

This third view also focuses attention on how performance measures affect ongoing relationships and the balance of power -- both within the agency and between it and the surrounding environment. For example, as long as performance measures are taken seriously by higher administrators or outside observers, they will always have the potential -- with or without justification -- to make individuals or the organization as a whole look bad. The best response to this would be to develop more complex and sensitive measures; the easiest response (in the short-run) is simply to use more flattering ones, or to invent strategies that might discredit or offset potentially damaging findings. The first approach is the one that will be pursued in this paper.

Potential Roles for Performance Measures

A different framework can be used to discuss performance measures in rehabilitation, based on the roles they might play in the program, rather than on their technical characteristics or the substantive ground they cover. Four basic roles for performance measures will be briefly described:

- -- sources of information for program improvement.
- -- sources of information for justifying and "defending" the program (or agency).
- -- a means of administrative intervention.
- -- a way of raising issues, educating participants, and generating then guiding conflict.

Even organizing the discussion in terms of roles makes it easier to consider certain issues. For example, one conclusion to be

developed later in this paper is that different types of performance measures may be required for at least some of the different roles.

Program improvement. Traditionally, the preferred role of performance measures and other forms of evaluation is to provide information that can be useful in improving a program (whether this information is in fact used is a further problem). Performance measures may address relatively short-term, incremental "management" decisions, or they may be directed toward more fundamental shifts in program strategy. Often, program improvement is seen as the only legitimate role for evaluation. Other uses would be considered subversive, such as: defending a "bad" program or covering up its failures, justifying the pre-planned termination of a program, maintaining a false front of scientific objectivity, or simply delaying action (Suchman, 1972). This cynical list, however, ignores some very functional uses of performance measures other than as sources of information for program improvement. These other uses will be introduced in the following pages, and further developed later in the paper.

Program defense. It is heresy in most quarters to openly acknowledge the use of evaluative information for "program defense" as a legitimate core. Nevertheless, some data is in fact used in almost every agency to explain the program and its accomplishments to outside audiences. Practically speaking, rehab agencies will (and perhaps should) continue to emphasize that portion of the available data which shows them in at least reasonably good light. Rather than try to fight this tendency on its home ground (within the agency), it may be more sensible: (a) to attempt to contain it within reasonable bounds,

(b) to balance agency reports by providing for external sources of program analysis (funding for client-led evaluations, etc.), and (c) to make sure that at the very least program administrators are not misled by their own published data, or tempted to confuse it with the information they need for internal management.

Administrative intervention. A relatively recent trend in rehab programs is the widespread interest in using performance measures as direct means of intervention and control -- a way of deliberately modifying the incentives faced by counselors and others, and thus altering their behavior in the intended directions. Rather than using performance measures simply as a source of information to supervisors and administrators who in turn take appropriate action, this approach tries to "internalize" the feedback process -- with performance measures providing an explicit standard of judgment and an "automatic" means for counselors to know how well they are doing. The further assumption is, of course, that counselors would in fact be motivated to improve their ratings, and also know how (and be able) to do so.

For reasons to be elaborated below, I question the degree to which performance measures in rehab programs can actually serve in this role, as direct means of intervention and control. It may be noted that, to date, most of the enthusiasm for using measures in this way has been expressed in journal articles and papers delivered at professional conferences (Lawlis and Bozarth, 1968; Miller and Barillas, 1967; NCACVR, 1968), rather than through a ground-swell of demand from practitioners.

Issue-raising and education. The potential role of performance measures in raising new issues, educating program participants, and channeling basic value conflicts toward some type of resolution may be

one of their most important contributions over the long run -- yet it is one that is complex, subtle, and seldom mentioned. Even the process of trying to define performance measures in operational terms may help to clarify goals and priorities. Widespread participation in this process can help to resolve conflicting points of view, provide a common language and focus for discussion, and clarify some of the obstacles to program change (Collignon, 1973). In order for performance measures to effectively play this role of stimulating and guiding debate, however, they should be:

-- linked to a set of issues that are widely perceived as important within the program.

-- based on concepts that have some intuitively clear meaning.

-- tied to acceptable theory and to data that is valid and readily available.

-- flexible enough to be changed and refined over time.

-- relevant to the policy decisions that have to be made, and to the tools available for implementing them (DeNeufville, 1972).

The "issue-raising" role of performance measures will not receive as much attention in this paper as the other roles. This is not intended as a reflection on its relative importance, but may simply indicate how under-used this role currently is, and how difficult to talk about outside the context of a specific program.

Scope of the Paper

Biases and limitations. At this point, a few of the key assumptions in this paper should be made more explicit. First, it is assumed that the basic strategy of rehabilitation works fairly well, but that there also remains room for improvements -- for example, the introduction of

new service-delivery techniques or a shift of emphasis toward serving new client groups. It is less important to add to the array of services already available than to preserve, in the face of continued pressures for standardization, the tradition of services being tailored to each individual client, with most decisions made jointly by clients and counselors rather than administratively imposed.

Second, I will be concerned with the use of performance measures only in the public sector of what is, in reality, a much broader rehabilitation system. This larger system includes additional sources of services and support to disabled individuals, resources whose impact is significant but perhaps even more difficult to assess or control. These resources are provided by: privately operated facilities and service-vendors; other social service and income maintenance agencies; voluntary associations; and the entire network of independent professionals, family, and friends available -- or sometimes not available -- in the community.

In a similar fashion, while this paper will deal mainly with explicit, formal performance measures and evaluation processes, this is not to deny the practical significance of informal channels of information. Both have major roles to play in understanding and managing any large-scale organization. Ideally, the formal system should support and complement the informal one -- or at least not undermine it.

An additional limitation in the scope of the paper is that it emphasizes policies and practices as seen at the State level and below, rather than primarily from the Federal agency perspective. One reason for this is that the State agencies have more of a direct responsibility for program operations, and thus a greater need for performance measures

that are effective in the "program improvement" role. As for explaining and "justifying" the program to outside observers, performance measures that are useful in this role to State agencies can also be readily adapted for use at the Federal level.

Perhaps the central reason, though, for concentrating on performance issues as seen from the State level and below is my own assumption that positive change in the program will not come predominantly from the higher levels of the system or from the outside, but from within and below. However, this is not to say that basic changes in the Federal legislation (such as the proposed "Rehabilitation Act of 1972," see: U.S. Senate, 1972; and U.S. House, 1972) cannot still play important roles in: (a) helping to ease existing administrative constraints; (b) providing budgetary resources and legitimacy for innovations; and (c) changing the framework and rules for decision-making, to encourage new issues to be raised and to involve new participants (clients) in resolving them.

Two of the problems that will receive little attention here are (1) how to relay performance information from the States to the Federal agency, and (2) in turn how to convey Federal policies back to the States in ways that assure their enforcement. I see these tasks as less urgent than others, and in a sense, as problems that should remain difficult.

There is a similar bias in the paper toward decentralized and multi-centered strategies of information and management within each State agency, to be accompanied by the "opening-up" processes mentioned above. These views will be further developed in the final chapter.

Format. The following chapters begin (Chapter 2) by summarizing the available information on current monitoring and evaluation practices in State agencies, as a point of departure for considering alternative approaches. Several such alternatives found in the literature will be reviewed and critiqued in Chapter 3, and a number of general issues raised for consideration. The fourth chapter shifts from a direct concern with program evaluation (determining what should be done) to the problems of intervention and control (how to create the intended changes in the system). These two points of view converge again in the final chapter, where an effort is made to outline some key features and criteria for an effective "guidance process": one that can help to raise, answer, and ultimately act on some of the basic issues facing rehabilitation programs.

CHAPTER 2

CURRENT PRACTICES IN STATE REHABILITATION AGENCIES

Overview. Levels of monitoring and evaluation: program management; "Lower-order" (managerial) efficiency; program effectiveness; client and community impact; "Higher-order" (allocative) efficiency.

Overview

Current monitoring and evaluation practices vary greatly among State agencies -- in both their level of effort and degree of sophistication. In a recent survey of rehabilitation agencies (Ridge, 1973) three-fourths of the "general" agencies (excluding those that serve blind clients only) reported that they had a "program evaluation division." Of these, however, one agency out of three said they spent less than one man-year of staff time on evaluation. A similar number devoted fewer than three man-years to evaluation for the entire agency.

The focus of monitoring and evaluation also differs among States. Some agencies are primarily concerned with tightening up their day-to-day management and avoiding conflict with Federal regulations; others have begun to sponsor major research dealing with program impact, causal relationships in the service-delivery process, and other complex topics.

There is also wide variation among State agencies in terms of their commitment to evaluation and monitoring, and in the extent to which findings influence the actual operation of the program. In some cases, studies are undertaken in response to real program issues, and the

answers (when adequate themselves) are taken seriously. Other agencies, though, seem willing to devote only a small portion of their budget to monitoring and evaluation -- perhaps reluctantly, to avoid being left behind in the popular move toward "rational analysis" in government.

On the Federal level, RSA obtains basic program statistics through the "R-300" reporting system. Regular reports are produced by each field office, compiled at the State level, and then forwarded to the Federal agency. The RSA staff performs some data analysis and then publishes a series of statistical reports, often formatted to allow the basic statistics and some derived ratios to be readily compared among the State agencies. Basic research and program evaluation studies are also sponsored by RSA and its umbrella agency, the Social and Rehabilitation Services (SRS) branch of HEW (for recent examples see Abt Assoc., 1972; I.I.S., 1971; Teitz and Collignon, 1973; and the list of currently-funded research projects in Newman, 1972).

While there remains considerable room for improvement, interest and commitment to program evaluation at both the State and Federal levels have grown rapidly in recent years. Evidence of this is found in the "Rehabilitation Act of 1972" (H.R. 8395 -- reprinted in: U.S. Senate, 1972). This bill was passed by Congress in 1972 (unanimously) and again early in 1973, but vetoed by the Administration both times. A special section of the Act required that <u>all</u> Federally-funded rehabilitation programs and special projects be evaluated for impact and cost-effectiveness, and that all such evaluation findings be made public. In addition, every evaluation study would itself have to show that an effort had been made to include the opinions of "program participants." Finally, the Act required that an annual report be

submitted to the President, including data on services provided by the program and the number of successful closures -- broken down by type of job placements (or other closure status). Considering the wide support for H.R. 8395 among concerned interest groups, it is likely that a similar emphasis on program evaluation will emerge in future versions of the legislation.

Levels of Monitoring and Evaluation

Collignon (1972-B, 1973) offers a useful framework for discussing monitoring and evaluation practices within State agencies. Evaluation efforts can be focused at five different levels, the first two concerned with program processes and the latter three oriented more to the program's outputs or impact:

--program management

--"lower-order" efficiency

--program effectiveness

--client and community impact

--"higher-order" (allocative) efficiency

At each level, different types of questions are asked, different analytical techniques tend to be the most useful, and the results are generally used for somewhat different purposes. Each level of analysis will be described briefly in the sections to follow.

Program management. At this level, performance questions have to do with whether the program is operating legally and following the principles of sound fiscal management. Comparisons are made with standard procedures in other rehabilitation agencies. Attention is focused on conserving resource-inputs (staff time and funds) and on

following agency procedures. Accounting reports and basic management audits monitor the use of program resources; reviews of case records are the usual means of assessing counselor activities. All State agencies perform this type of routine monitoring for purposes of program management.

"Lower-order" (managerial) efficiency. The next step in performance monitoring considers not only program inputs and routine procedures, but also the service-delivery process and the relationship of inputs to the program's immediate "output": successfully closed (status 26) cases. Attention might be directed to various input-output ratios, the analysis of case flows, or to other process indicators such as staff turn-over rates. According to the above-mentioned survey (Ridge, 1973), the indicators commonly monitored at this level by State agencies include: number of new referrals, new plans, and 26-closures per counselor each month; the ratio of new plans to new referrals; the ratio of 26-closures to plans; the number of clients per counselor; average program costs per client (or per 26-closure); and so forth.

The "flow" of a counselor's caseload may also be monitored, based on the time each of his clients takes from referral to plan or from plan to closure. Some field supervisors check further to see that their counselors have a "balanced" caseload, without too many or too few clients at any single stage in the process (which would tend to produce an uneven stream of closures during the program year).

In many State agencies, the Federal R-300 reporting forms have been found to be inadequate, and have been supplemented by additional reporting requirements. At the field level of the State agencies, however,

the common complaint is the opposite: not that the amount of information collected is insufficient, but that much of the data is unnecessary -- or else misinterpreted by higher-level administrators. Supervisors and field-level administrators often complain that the statistical reports distributed from the central office arrive too late to be used for their management purposes. They also complain that data is provided in too complex a format, without adequate explanation or analysis. Finally, they state that much of the data is simply too aggregated to be useful. For example, it is difficult to separate out the data on special programs or on counselors with specialized caseloads for proper comparison with others throughout the State. As might be expected, many field-level managers admit to using other informal sources for their "real" information, or else have developed their own rules of thumb, using one or two reported statistics as lead indicators to spot possible problems.

Some general sense of the extent of client attrition during the rehabilitation process can be gained through comparisons of: the total number of cases referred, the number of new plans written, and the number of actual 26-closures for a given counselor or field unit. At present, though, there appears to be no State agency that analyzes the attrition process in depth, as it affects different groups of clients, both during service-delivery and at the initial stages of referral, intake, and eligibility determination (see Chapter 3).*

A different type of analytic technique that may eventually prove powerful is only now being introduced at the State level, for

---

*Systematic analysis of this sort was called for as early as 1965, during a series of National Institutes for Rehabilitation Research (Dishart, 1965).

example in the Washington State agency (Serot and Collignon, 1973). This approach attempts to identify, through statistical analysis, the rough equivalent of a "production function" for rehabilitation programs. The basic aim is to determine the effect on the program's performance* of each major -- and potentially controllable -- input to the service-delivery process, for example: type of services provided, amount spent by the agency, counseling variables, client characteristics, and so forth. Of course, this assumes that, even in a program as complex and variable as rehab, there are some systematic, discoverable, and potentially controllable relationships between program inputs and outcomes. At the very least, further studies in this area might help to reveal implicit trade-offs among program inputs or across the various dimensions of performance.

Program effectiveness. This third level of analysis deals more directly with the outputs of the program: the quantity and quality of "successful" case closures and the degree to which these reflect the program's objectives and the needs of the disabled population. Quantity of closures is commonly expressed not only in terms of the number of rehabilitations per year, but also as a rate per 100,000 population in each State (or each district). In many cases, a rate of rehabilitations based on the total population may not provide valid comparisons among different geographic areas, since the number of disabled who are eligible for services (i.e., the "pool of need") may represent varying proportions of the population (Ridge, 1972). Statistics on "coverage" (percent of need actually served) are generally not broken down for sub-groups of

*Defined as the probability of closure or as changes in earnings.

clients, due to: (a) the lack of data, (b) problems in precisely defining disability and eligibility for different client groups, and (c) the potential political implications of exposing major differences in the program's treatment of different client groups.

One exception to this was a recent study by the California State agency (1972-B), which provided estimates by broad disability group of the total need, the number served, and the number (expected) to be rehabilitated in fiscal 1972. While the report itself did not furnish percentages (these would have made comparisons among disability groups even more obvious), simple computations revealed that the proportions of estimated need in each sub-group projected to be rehabilitated in 1972 ranged from 1.4% (alcoholics) to 7.1% (mentally ill) -- or in other words, varied by a factor of four.

The "quality" of vocational rehabilitation services delivered to the client (as opposed to the difficulty of the case or the actual degree of benefit to the client) is most often assessed through the "expert judgment" of supervisors during periodic case record reviews. The review process may involve the counselor himself, and occasionally other counselors acting as "peer reviewers."

Criteria to be used for reviewing the quality of case work (or, to be more exact, the quality of case records) are contained in the program manual. These may be supplemented by professional guidelines, State agency criteria, or by additional standards generated within each field office. For example, criteria used for case review in California (Collignon et al., 1972-B) may include:

--client eligibility for services (presence of a disability, demonstration of an "occupational" handicap, expectation that services can alleviate or remove the handicap).

--appropriate diagnosis of client's disability and needs.

--formulation of adequate plan for services, with suitable goal(s).

--plan actually carried out.

--counseling provided to client.

--successful closure and employment (etc.) for 30 days or more.

In one region of the State, a recent sample review of case records (Dieckmann, 1972) included the above criteria plus others dealing with:

--adequacy of documentation and case-recording.

--flexibility in the client's plan, as needed, during the period of service-delivery.

--prudent use of funds.

--sufficient counselor-client contact.

--appropriate "movement" of the case, without undue delay in any status.

The quality and effectiveness of services provided by private vendors and facilities under contract with the State agencies is another major area of concern that has recently begun to receive attention both at the State level and among counselors in the field. The significance of this effort is suggested by the fact that between one-third and one-half of all case service expenditures in an agency may be budgeted for such private contract services. Inadequate services from private vendors, or even delays in starting services to a client, can be major causes of attrition -- over which both the client and the agency have up to now had limited control (Abt Assoc., 1972; Markowitz, 1971 and 1972; Markowitz and Collignon, 1972; and NCACVR, 1968).

Client and community impact. At this level of program evaluation, the issues become more complex and the techniques more time-consuming

and expensive. Few State agencies carry out systematic studies of the ultimate effects of their program on clients, their families, and the community at large.

One critical evaluation question at this level is the extent to which client benefits, both vocational and non-vocational, can actually be attributed to the agency's services. Benefits in this sense might be measured not only by job placements, but also by: changes in client earnings (or those of other family members who are freed to work), the ability of the individual to function in daily living situations, and the equally important (but difficult to quantify) gains in the client's physical and psychological well-being. According to the survey of State agency practices mentioned above, only six agencies reported efforts to regularly assess any of these non-employment benefits to rehabilitated clients. An equally small number reported having ever used a control group in evaluating benefits to clients -- even though this is necessary to determine the amount of observed "client change" that can actually be attributed to the program rather than to some outside factor.

The stability of gains measured at the time of case closure and the degree of client satisfaction with the services are also important criteria for assessing program impact. Information of this sort can best be obtained through periodic follow-up studies, using a sample of 26-closure cases. However, fewer than three out of five of the State agencies responding to the survey reported having ever undertaken a follow-up study of closed cases. Formal surveys or other modes of client participation can also be direct sources of feedback on the program's quality and impact. Only two out of five agencies, though,

reported that clients or their representatives were regularly involved in any monitoring or evaluation of the program.

Another interesting finding emerged from a series of informal discussions with several counselors, supervisors, and field-level administrators in the California State agency. When individuals at each level of the agency were asked to identify their most important source of information on "how well the program was doing," they replied that in general this was obtained neither from routinely reported data nor from the results of special studies, but rather through informal feedback -- positive and negative -- from clients, employers, other agencies, and the community in general. While such "informal" sources of information on program quality and impact are often overlooked by evaluation staff in the agency's central office, they may nonetheless represent a major untapped resource -- perhaps one that is already being used tacitly but effectively at the field level.

In an even broader sense, "community impact" might be judged by the extent to which the program helps to re-educate public attitudes toward the disabled, to reduce the notion of stigma associated with physical or mental disabilities, and to enlarge the community's recognition of need and acceptance of a social responsibility to deal with it. Most State agencies have not directly examined this level of program impact in the past.

"Higher-order" (allocative) efficiency. At this fifth level the problem is to evaluate the overall "worth" of the entire rehabilitation program, in terms of the resources used and the achievement of socially-defined goals. This sort of an assessment is commonly approached

through some form of benefit-cost analysis. While such studies have become considerably more complex and sophisticated over time -- in terms of their assumptions, their methods of measuring (and actually pricing) a wider range of program benefits and costs, and their ability to disaggregate benefit and cost data to suggest some of the distributional effects of a program -- they still share some fundamental shortcomings in both concept and technique. One basic problem comes from assuming that the huge variety of goals and resources valued by different members of "Society" can ultimately all be expressed meaningfully in a common term such as dollars.

When benefit-cost studies necessarily fall short of translating all goals and resources into dollar terms, their failure is not a random one. Instead, a systematic technical bias is introduced -- one that leads inevitably to ethical and political biases, as well: all factors that cannot readily be measured and priced are delegated to a verbal "discussion," while the tangible benefits and costs that can be buttressed with "hard" numbers find their way into much more visible numerical calculations.

In spite of these problems, benefit-cost analysis has had a long history within the rehab program nationally (Conley, 1965 and 1969; U.S. DHEW, 1967; Abt Assoc., 1972). But it has been used more often for proving to Congressional committees that the "investment" of public funds in rehabilitation can yield impressive monetary returns than for actually making "rational" resource-allocation decisions, by comparing the effectiveness of alternative human resource programs.*

---

*For a series of benefit-cost studies presented in a way that deliberately discourages such direct comparisons among similar social programs as "premature," see: U.S. DHEW (1966).

Attempts have also been made at the State level to employ benefit-cost techniques in assessing the "overall worth" of rehabilitation programs as a use of public resources (for some recent examples, see Bellante, 1971; Michigan State DVR, 1970; Wright et al., 1969; and Collignon and Dodson, 1973). Once again, the aim has been primarily to demonstrate the value of the program to outside observers in the State Legislature and the Governor's Office -- rather than to "rationally" guide a State's allocation decisions.

Given the state of the art in benefit-cost analysis, though, it is probably fortunate that the analysts' achievements have been so modest. A serious attempt to rationalize broad-scale allocation decisions using benefit-cost criteria would probably have served only to further disguise the questions of distributive equity that are implicit in all allocation decisions.

With this brief overview of current monitoring and evaluation practices in rehabilitation agencies, the following chapter will turn to a discussion of several major problems with the current performance measures, and to some of the specific proposals that have been made for changing them.

CHAPTER 3

PERFORMANCE MEASURES - A REVIEW AND CRITIQUE OF RECENT PROPOSALS

Introduction. The current measure: advantages; information problems; incentive problems. Weighted case closure: statistical correlation studies; expert ("clinical") judgments; agency performance studies; general critique of weighted closures. Other approaches to performance measurement: refine/disaggregate the 26-closure; replace the 26-closure; analysis of selection/attrition. Introducing new performance measures - some further issues. Re-thinking the need for new measures.

"...the evaluation measures look at rehabilitation as an event -- when in reality it's a process."

--rehab counselor,
Calif. State DR

Introduction

This chapter will discuss some of the problems (and a few advantages) of the 26-closure as a performance measure, and then examine critically several alternatives that have been proposed recently. Particular emphasis will be placed on the advantages and limitations of various schemes for differential "weighting" of 26-closures. In general, I will argue that while each of the proposed alternatives may represent an improvement over the existing system, each one also retains some of its disadvantages (emphasis on a limited range of program objectives, self-judgment by the counselor, incentives

to fudge the data or "game" the reward system) while often adding new ones as well (complexity, expense, new perverse incentives, inability to deal with the outside constraints that also affect counselor decisions).

In a larger sense, perhaps the most serious problem is the tendency of observers and reformers to view each new proposed performance measure in isolation -- either as the "total solution" or as totally useless. In contrast, the approach to be outlined in Chapter 5 calls for the simultaneous use of multiple performance indicators, along with fundamental changes in the process of analyzing data and using the results.

The Current Measure

Routinely collected performance data provide much of the information needed for monitoring and evaluation activities at all five of the levels described in the previous chapter. The basic building-block for a number of other ratios and composite indicators is the number of "successful" closures in status 26. The requirements for a 26-closure (U.S. DHEW, 1969) are that clients:

> ...(1) have been declared eligible, (2) have received appropriate diagnostic and related services, (3) have had a plan for vocational rehabilitation services formulated, (4) have completed the plan insofar as possible, (5) have been provided counseling and one or more other rehabilitation services, and (6) have been determined to be suitably employed for a minimum of 30 days.

In practice, counselors exercise wide latitude in applying these criteria to individual cases.

Advantages. A number of advantages are commonly ascribed to the 26-closure, as a well-established output measure for rehabilitation programs. In comparison to other social programs, at least, the

26-closure represents a relatively well-defined standard, a means of keeping the agency fairly visible and accountable to the public. From the agency's point of view, this translates into a more effective means of demonstrating achievement and justifying annual budget requests.

The 26-closure also provides a focus and an end-point for the service-delivery process. This may be important to the client in providing a tangible, clearly understood objective tailored to his specific needs. On the other hand, for the counselor or the agency the 26-closure may be equally useful as a basis for terminating services -- either because a client's goal has been reached or because there appears to be too little hope of ever doing so.

Two comparisons may be useful here. One is with manpower programs, where in general there is no effort made to maintain a relationship between the client and the agency, other than that established by the client filling a "slot" in a specific training program. Even more fleeting contact is not uncommon, with the client simply being referred by the agency to a listed job opening. This clearly contrasts with the approach in rehabilitation programs: the planning of a sequence of services fitted to the needs of the individual client -- rather than the choosing of clients to match the available openings for training or employment.

A similar comparison could be made with welfare and other income transfer programs, which are rarely (in practice) committed to an active role in "human resource development." These programs often lack both an identifiable process for helping the client to improve his skills or life conditions, and an explicitly stated goal that could be used to measure program impact.

In spite of these advantages of the 26-closure, its use as the primary measure of program performance, combined with administrative pressures on rehab counselors to produce a minimum "quota" of closures each year, have generated objections from a wide range of program participants. As suggested earlier, these objections have to do both with the quality of information provided by the 26-closure, and with its effects on the behavior of those within and outside the agency. Let us examine each type of problem in greater detail.

Information problems. One basic complaint about the use of the 26-closure as a performance measure is that it leads to an emphasis within the agency on quantity (total number served, number of 26-closures) over service quality (the adequacy and appropriateness of services for each individual client). An equally serious problem is that the 26-closure-plus-quota system encourages outside observers of the rehab program to look almost exclusively at the number of closures, when assessing agency performance. As an example, the legislative analysis of the proposed FY 1974 budget for California's State Department of Rehabilitation asserted that: "The basic output of the Department of Rehabilitation is the number of disabled persons successfully rehabilitated during a year" (Calif. State, Legis. Analyst, 1972). The report also noted the "apparent decline in performance" of the agency -- measured in terms of the small percent of agency-projected 26-closures that had actually been obtained in each of the preceding years. The analysis called for annual reports from the State agency, to explain any future discrepancy between the number of rehabilitations projected in the agency's budget statement and the number actually achieved.

This outside focus on the total number of 26-closures can only reinforce pressures to use similar, simplistic performance measures within the agency -- a clear case of organizational "goal displacement." Fortunately, many individuals in the rehabilitation field have consistently opposed any trends toward using numerical production as the sole output criterion. They have correctly perceived that to yield completely to a single-number criterion would have meant replacing the program's traditional emphasis on individualized, high-quality services with the same "mass treatment" approach that has often characterized other social service and human resource programs.

It should be clear that paying primary attention to the total number of 26-closures can also create problems for clients, making it more difficult for the "severe" cases to be accepted for services at all and less likely that any client, regardless of severity, will receive many of the more expensive or time-consuming rehabilitation services that he might need.

The strongest objections to a single numerical output measure, however, have come from rehabilitation counselors themselves. There are at least three reasons for this. One is that many of the less experienced (or less effective) counselors, or those working with difficult caseloads, see the annual quotas for 26-closures as a direct threat. It is commonly believed -- correctly or not -- that a counselor's numerical record of closures has a direct and primary impact on his chances for promotion.

Similarly, many of the "better" counselors also object to the use of 26-closure quotas, because it interferes with their expressed

desires to provide clients with high quality counseling and other services that are increasingly defined by professional standards. In particular, the need to meet a production quota leaves them less time (often a more critical constraint than money) to deal with their "challenging" cases. Finally, both those counselors who have difficulty meeting their quotas and those who are merely inconvenienced or distracted by quotas tend to view the imposition of any administrative criteria as a slap at their professional judgment and desire for autonomy.

A different level of concern over the 26-closure is that it is misleading or inadequate as a measure of many aspects of agency (or counselor) performance. Both administrators and some outside observers of the program tend to be suspicious of statistics on total closures, because they provide no basis for distinguishing "easy" cases from "difficult" ones. On the one hand, lower level staff may be suspected of actually "fudging" the numbers. On the other hand, administrators may simply be concerned that comparisons among counselors, district offices, or entire State agencies are "unfair" because they penalize those: (a) who work under more difficult conditions, (b) who are faced with (or select) more difficult caseloads, or (c) who are more honest in closing (and recording) cases as successful 26's.

Counselors express similar concerns over the "unfairness" of the 26-closure in a program subject to so much inherent variation. In practice, though, most proposed methods of adjusting the 26-closure to take case difficulty and other factors into account would probably prove unwelcomed to at least some counselors -- those who would gain

little or nothing in their relative rating, and yet at the same time lost the convenient excuse that their "unique" situation cannot be adequately described by the (unadjusted) 26-closure.

The 26-closure has also been denounced for ignoring the subtler aspects of program performance and the full extent of its benefits to clients. Not only are all "successful" closures counted the same regardless of the client's gains, but any case closure status other than 26 is assumed to represent no benefit to the client. In fact, though a client may have gained substantially in non-vocational areas, or he may have developed vocational skills but for some reason been unsuccessful in locating immediate employment.

Recent informal studies by the California agency staff revealed that many clients who were closed as "unsuccessful" after receiving services (status 28) in fact credited the program as providing them with valuable services. In a second study, it was found that a number of former clients with unsuccessful 28-closure later turned up as contributors to the State's disability insurance fund, indicating that they had somehow obtained jobs (perhaps due in part to the agency's services*).

Findings of this sort suggest that performance measures might pay more attention to those clients closed in statuses other than 26 --

*Moreover, observers have noted that some clients closed as 28's may be "drop-outs" from the program precisely because they were able to locate employment on their own. The problem of attributing observed effects to the program rather than to other influences remains a serious one, both for "successful" (26) closures and for the so-called "unsuccessful" closures. The issue will never be resolved through cross-sectional correlation studies alone. Research based on some form of longitudinal, quasi-experimental design will be necessary -- even though expensive and subject to ethical and political, as well as technical complexities.

for example, the existing status 22 (services completed and ready for employment). Alternatively, new closure statuses might be added, to give the agency some credit (for example) for a client who has obtained employment plus other benefits as a result of rehab services -- but without completing his initial plan.

While benefits to non-26-closures may at times be understated, gains to clients closed as 26's might also be exaggerated -- either because benefits were obtained without the agency's help (several studies have suggested that many clients locate employment with little or no placement effort by their counselor) or because employment and other gains reported at closure may not last much longer than the minimum thirty days. Possibilities that the 26-closure may seriously understate or overstate actual client gains due to the program only underscore the need for more complete and sensitive measures -- as well as the often-repeated call for periodic follow-up of closed cases on a sample basis.

A further, more fundamental objection is that the 26-closure may simply be inappropriate for some clients. Even with the addition of homemaking and sheltered employment as acceptable "occupations" (along with competitive employment) at closure, there are still clients who may be able to benefit enormously from rehabilitation services but who cannot meet the criteria for a 26-closure, and thus have little chance of being accepted for service.

The recently-vetoed "Rehabilitation Act of 1972" took some important steps toward further broadening the eligibility and closure requirements for those severely disabled clients who could be helped toward "independent living" goals, even if not toward employment per se. The thrust is one that is likely to be continued.

Another concept gaining support within the rehabilitation field (and also included in the 1972 legislation) involves a shift away from the idea that services should always be terminated after the client reaches some pre-determined goal. For at least some clients, including those with severe, multiple, or "developmental" disabilities, it may be more appropriate to think in terms of "follow-along services." This would allow the agency to retain some degree of active involvement with the client more or less permanently (Conley, 1972). Performance measures might then be monitored at regular intervals, to determine the client's improvement, or at least the stability of his gains.

Incentive problems. Each of the issues mentioned above has been concerned with the quality of information provided by the 26-closure. Many observers of the rehab program are also questioning the effects of the current performance measure in altering the incentives (and thus the behavior) of agency participants -- and of counselors in particular.

The fact that all 26-closures have equal weight in the formal recording system tends to create incentives for at least some counselors: (a) to select relatively easier (or more certain) cases, (b) to encourage their clients to agree to easier vocational goals, (c) to shorten the period of service-delivery if possible, and (d) to close the case as soon as is legally allowed -- or alternatively, to delay the closure date in order to help meet the quota for an anticipated "lean period" in the future (Conley, 1972; Hawryluk, 1972; NCACVR, 1972; Silver, 1969; Viaille, 1968).

While it has been widely argued that the above incentives exist, there is little hard evidence that counselors actually do perceive and respond to them in making decisions about who will receive services,

what services will be provided, and when a case will be closed. Intuitively, it seems reasonable to assume that different counselors react somewhat differently to incentives created by the 26-closure-plus-quota system. There are a number of other factors (professional self-image, peer approval, administrative regulations and direct supervision, commitment to helping the client) that also strongly influence a counselor's decisions. As will be suggested below, efforts to deal with this problem of undesirable incentives of the 26-closure can take a variety of forms -- ranging from the attempt to deliberately shift counselor behavior in carefully specified directions, to the far less demanding task of trying to reduce some of the worst dis-incentives of the current system, but without specifying in detail an ideal pattern of incentives and behavior.

In response to the two sets of issues raised by the 26-closure -- quality of information and effects on incentives -- a large number of proposals have been advanced in recent years to modify, supplement, or replace the 26-closure as the program's basic performance measure. In the following sections, several of these proposals are reviewed and critiqued. The examples were chosen with a bias toward more recent proposals, and toward those that best illustrate alternative strategies and the limits of each. Critiques of the proposed measures emphasize their projected use in an agency rather than the details of their technical design.

Weighted Case Closure

Currently the most popular approach to changing the performance measures in rehabilitation is to apply suitable "weights" to the basic 26-closure to make the total number of (adjusted) closures a more

accurate comparison among counselors or districts. It is usually assumed that such weighting will dramatically change the current pattern of incentives and behavior presumably caused by the unweighted 26-closure.

Some of the literature* on weighted case closures focuses on the need to explain observed differences in numerical production -- among counselors, districts, States, and over time -- and to determine what proportion of this variation is due to: client characteristics, service variables, or "environmental" factors beyond the control of the counselor or the program. Improved understanding would then allow administrative adjustments in the expected number of 26-closures, better allocation of the program's resources to correspond to the level of need in each community, and so forth. Any residual, "unexplained" variation (other than chance) would then be the responsibility of administrators and supervisors to deal with at the field level.

In other writings, however, the focus is primarily on the deliberate use of weighted case closures as incentives, to change the client-selection and service-delivery decisions of individual counselors (for an explicit statement of this view, see the testimony of former RSA Commissioner Newman during the recent Senate hearings on H.R. 8395: U.S. Senate, 1972, p. 239).

Most of the proposals for weighting 26-closures are based on one or more of the following factors (Conley, 1972):

--counselor time and "effort."

--use of program resources.

---

*An extensive, up-to-date review of this literature may be found in: Oklahoma State DR, 1972.

--"case difficulty"- based on empirical rates of success for similar cases, on time spent in the rehabilitation process, or on a priori judgments -- for each case or by categories of similar cases.

--degree of client change - measured by changes in earnings, shifts in vocational skill categories, improved performance of daily tasks, or changes in psychometric test scores.

A fifth criterion for weighting case closures, based on the degree of public policy priority attached to serving various client groups, is seldom mentioned explicitly (Collignon, 1973). Nevertheless, it is precisely this concern for priorities in allocating resources among potential clients that presumably underlies the issue of "creaming," the most serious dis-incentive attributed to the unweighted 26-closure. In any case, client selection and resource allocation decisions will continue to be made implicitly -- if not openly. As will be suggested below, proposals for weighted case closures could either have the effect of raising these fundamental policy issues and helping to structure the debates -- or they could further obscure the critical questions of equity and program coverage.

It is generally suggested that the actual numerical values for a weighted closure system be derived from one of three sources:

(1) statistical correlation studies, to determine which client, system, or environmental characteristics are associated with observed success rates (or with some other criterion variable, such as changes in earnings).

(2) expert judgment by counselors, supervisors, outside authorities, etc.

(3) agency performance norms (past or current), including average success rates, average time spent receiving services, etc.

In practice, any given weighting proposal may draw on more than one of these data sources. For each category, some representative examples from the recent literature are summarized in the sections to follow.

Statistical correlation studies. A number of studies have attempted to identify factors which could be used to predict client success. Lawlis and Bozarth (1971) performed regression analyses using a number of client characteristics and "administrative" variables. With the exception of age and education, they found that the administrative factors (average expenditures per case, total spending by each counselor) were better predictors of rehabilitation success than were characteristics of the clients themselves. In an earlier study, DeMann (1963) found that a combination of eight variables (characteristics of clients and of services) could identify potential rehabilitants vs. non-rehabilitants about two times out of three, a score that has yet to be beaten using correlation techniques. Aiduk and Langmeyer (1972) reported that the same demographic indicators that had proven statistically significant in studies dealing with general rehabilitation caseloads were less effective predictors for a specific sub-group of rehab clients: mentally ill patients in a State hospital.

Clearly, the pattern that emerges is one of serious inconsistency across studies -- even where similar variables were used. A recent literature-review by the Oklahoma State agency (1972) discusses this and other problems with the correlation-prediction approach. Not only was there a lack of consistency among studies (which might have resulted from actual differences among programs or client populations -- as well as from poor analytic techniques), but even within each study it was common to find only a small percentage of the total variance in client success rates "explained," by all factors taken together. On the whole, these two findings do not suggest optimism about the use of statistical techniques, either for identifying client

sub-groups with similar chances of success, or for specifying the actual weights that should be attached to a successful closure for each sub-group.

The Oklahoma paper concludes that:

> Given the complexity of the problem, the tools presently available have limited usefulness. Demographic characteristics have been shown to have low predictive relationship to outcome, and psychometric data adds little knowledge in terms of the difficulty of the individual case. (p. 19)

Other criticisms have been made of the statistical approach to predicting rehabilitation success. For one thing, the lack of any coherent theory of the underlying causal relationships increases the risk of false inferences from observed correlation. A further problem is that some of the correlation studies have used data that in practice can only be collected after a case is accepted, and in some instances, only after closure. This makes such studies "predictive" only in the statistical sense -- they can contribute little to weighting schemes that attempt to shift counselor incentives at the point of accepting new clients for services.* A third set of problems stems from the fact that some of the data used in correlation studies would in reality prove too difficult, time-consuming, or unreliable if it were collected for individual cases.

Other technical issues arise in particular correlation studies: the same sample may have been used both to fit and to validate the parameters, or the attention of researchers may have been focused more

---

*In a strict sense, this statement is true only if the weights are to be calculated for individual cases, prior to acceptance. However, historical correlation data might be used to develop weighting factors for classes of "similar" cases (see below). Or the weights might be used only for post hoc adjustments to improve the "fairness" of closure data -- rather than to change counselor incentives directly.

on the consistency of a relationship than on the magnitude of change in the criterion variable (i.e. -- the size of regression co-efficients). Moreover, even when regression co-efficients are large enough to be significant in programmatic as well as statistical terms, there is no guarantee that these variables will also be the ones that might respond to deliberate manipulation (Etzioni, 1971).

A final important question is how the findings of correlation studies would be used -- to the extent that one is able to accurately identify (client) characteristics associated with rehabilitation success. Some alternative uses (DeMann, 1963; Perlman and Hylbert, 1969; Salomone, 1972) might include:

> --adjusting the pattern of services to better fit the needs of a given client (i.e., predictive variables used to anticipate special problems).
>
> --changing the allocation of resources or the number of closures expected of each counselor, depending on his caseload characteristics.
>
> --using the prediction tools to more effectively screen out those clients whose characteristics would indicate a low probability of success.

If the latter approach were chosen -- attempts to change which clients are served, rather than the techniques, resources, or expectations -- there would still remain urgent questions to resolve: which of these statistically significant correlates of success should nevertheless be disregarded, for ethical reasons? For example, if it did turn out that race or religious background served as a useful "predictor," should either one automatically be accepted as a criterion for selecting clients who will or will not be served? A less dramatic but perhaps more current example would be the case of severely disabled clients: would a better understanding of what "severity" means, in terms of

reduced chances for 26-closure, lead to new strategies for serving these groups -- or rather to increased incentives to screen them out?

Expert ("clinical") judgments. This category includes two broad sub-groups: (a) those studies that use judgmental techniques to define groups of cases that will be treated similarly for weighting (where the assignment of a particular case to its group can then be a relatively routine task), and (b) those studies that apply separate judgments to each case in assigning closure weights.

An example of the first approach is a method proposed by Silver (1969) to establish a few broad groupings by disability type, based on the experienced judgments of counselors. A second survey of counselors would then be undertaken to actually assign relative weights to each group, according to subjective estimates of "difficulty." Silver indicated that later refinements might further disaggregate each disability group according to age, education, work experience, etc. Additional counselor surveys would then be needed to estimate relative difficulty for each of the sub-groups.

A study by Harper (1972) actually attempted such a survey of counselors, to identify those client characteristics that seemed to reflect the greatest "barriers" to successful 26-closure. Unfortunately, the format chosen for presenting questions to the counselor-respondents, as well as the techniques used by the author for aggregating these separate responses do not appear to be valid.* With modifications, though, this general approach might yield some interesting results.

---

* Harper selected three values (or ranges) for each client variable (except sex) and then asked respondents to evaluate each separate value in terms of how serious a barrier it might pose to successful rehabilitation. For example, the "age" was broken into three values: under 27 years, 27-40 years, and over 40 years. The median scores for each separate item were then combined into an unweighted score for the

The two studies referred to above attempted to use judgmental techniques to establish categories of cases that could then be treated alike for purposes of weighting. In contrast, a study by Salomone (1972) used a panel of three counselors to judge the "motivation" of each individual client, based on case records. The results were not promising; judgmental ratings of client motivation showed only "weak, uneven relationships" with actual client success as measured by the 26-closure.

A more ambitious attempt to apply judgmental techniques to individual cases is described by Westerheide (1972), as part of a major research and demonstration project in the Oklahoma State agency. The Oklahoma study was aimed at developing and testing a single index for both "case difficulty" and "client gain," using a composite of 23 five-point rating scales. Each scale was to be scored separately by a judge, for every case. The rating scales were defined within five "functional areas": educational status, economic/vocational status, physical functioning, adjustment to disability, and "social competency." Particular attention was paid to the provision of anchoring definitions for each of the scales, along with a manual to assist the judges (who, in the case of the pilot study, were counselors not familiar with the clients, working only from case records).

---

variable. A comparison of these scores (the means of medians) was then the basis for selecting six out of the thirteen original variables for the construction of an aggregate scale.

This methodology led to some unusual choices. For example, race was excluded from the final scale -- because the high difficulty rating that respondents gave to a client being "Black/Brown/Red" was drowned out by the low ratings given to "White" and "Yellow" (sic!). Similarly, the perceived difficulty of rehabilitating clients in the "over 40" category was offset by the ease of dealing with younger clients. This meant that the combined variable, age, was totally excluded from the final scale.

Other conceptual and technical difficulties were due to the arbitrary break-points imposed on the continuous variables and to the use of simple aggregates of each respondent's ratings -- without attempts being made to normalize these for across-the-board biases (i.e., "halo" or "anti-halo" effects).

While the Oklahoma approach appears to be somewhat more sophisticated than other proposals to use judgmental techniques in individual cases, some questions might still be raised on both technical and operational grounds. One of these concerns the method of aggregating scores on the separate scales into average "area scores" for each client (it is these average scores which account for a good portion of the inter-judge reliability that is claimed for the rating scheme). The study suffers from another common drawback, the tendency to focus at the pilot testing stage more on issues of reliability ("Is the rating consistent?") than on those of validity ("Does the rating mean anything important?"). The latter criterion is, of course, more demanding -- it requires that both the separate indices and the method of aggregating them be tied to some solid conceptual framework. A weaker position, of course, is to simply trust that the reliability (consistency) of a judgment-based scale will represent at least a lower bound for its conceptual validity (Oklahoma State DR, 1972). This assumption is not always justified -- it rests on the argument that: "If there weren't something sensible about the question we're asking, our judges wouldn't all be responding in the same way."

An article by Bolton, Butler, and Wright (1968) deals in general with the relative advantages of "clinical" vs. "statistical" methods for predicting client success in rehabilitation. Their review of a number of studies once again points to conflicting findings. Several of the statistical studies claimed that they could accurately predict an eventual 26-closure in 65% or more of the cases. This is a showing notably better than that of the "average" conselor's clinical judgments.

However, Bolton et al. point out that to simply compare the "best" statistical-prediction methods with the judgment of the average counselor may be misleading -- counselors themselves were observed to vary widely in their ability to accurately predict a client's success in achieving 26-closure. Moreover, there seems to be much more potential for improving the accuracy of counselor judgments than that of the statistical methods. These possibilities include training, providing counselors with better information on clients, and the use of a check-list with "leading questions" to help organize and guide the counselors' judgments. The authors' conclusion was that the ideal approach would probably combine both statistical methods and some guided clinical judgments. Of course, the issue introduced above still remains: whether the purpose of improved prediction should be to better screen out some clients, or to better tailor the available services and resources to meet a given client's unique needs.

It should also be pointed out that rehabilitation programs have a strong tradition of relying on the professional judgment of counselors. Any attempt to administratively impose statistical methods as restraints on judgment -- rather than as sources of assistance and guidance -- would encounter strong resistance from the field. On the other hand, the decisions demanded of the counselor have also become more difficult and time-consuming, due to greater variety in the potential rehab caseload, a broader set of available services, and a wider range of goals and closure criteria. By easing -- but not replacing -- this decision-making burden, statistical prediction techniques could play a potentially valuable role.

A final difficulty with the clinical approach is that if counselors are to make judgments regarding the weighting of their own cases (which they know best), they would be confronted once again with the same old incentives to "fudge" results in order to meet production demands -- even if these are based on "weighted" production. On the other hand, if weighting judgments are to be made by others (counselors or supervisors, for example), additional problems arise, due to inadequate information in case files and the lack of incentives to take time and trouble for careful judgments.

Agency performance studies. Under this third approach to weighted case closure, the relative weights (usually reflecting "case difficulty") are derived from data on the system's past or current performance. For example, a study by Goff (1969) assumed that the total time elapsed from acceptance to closure could be used as a proxy for case "difficulty." Multiple regression techniques were empoyed to identify client and service characteristics that correlated with case difficulty (time spent) and with ultimate success (26-closure). Unfortunately, only a small percentage of the total variance was accounted for by the sixteen variables used. The author suggested "further research" prior to any efforts to actually construct a weighting system based on time-in-services.

Miller and Barillas (1967) used the rate of success for different client groups, rather than time from acceptance to closure, as the basic criterion for a "difficulty" weighting. Clients accepted for services were grouped on the basis of variables assumed to affect chances for successful rehabilitation. The proportion of each group that historically had been closed in status 26 (vs. other closures:

08,28,30) was then used to construct weights for each group (more credit was given to those groups with lower rates of success).

The same basic approach was adapted by Sermon (1971, 1972) to design and test an operational index using the rate of success as a proxy to weight groups of cases for difficulty. One basic problem with the "difficulty index" as defined by Sermon is the method used to translate group success rates into relative "difficulty." The procedure was to determine the percent of cases from a given group that were not closed as 26's, and divide this percent by ten, producing a weighting factor with a theoretical range from zero to ten. However, this also meant that the "credit" to be given for closing a case approached zero, as the rate of success for that group rose toward 100%. Extremely high or low success rates would thus threaten to distort incentives.

Other difficulties were pointed out by Sermon himself. One is the assumption that case difficulty and complexity are infact associated with -- and validly proxied by -- low rates of success. A second is the bias introduced when weights are computed using success rates for those clients who have already been accepted for services -- in other words, after extensive screening, based in part on estimated case difficulty, has already occurred. A third problem with the index is that the weights themselves are likely to need periodic updating, thus increasing the costs of the system and decreasing the chances that it will be perceived and responded to as a stable source of counselor incentives (see below).

Sermon suggested that the basic difficulty index be used to generate two additional performance indicators. These would then be used along with the actual (unweighted) number of 26-closures to provide

four complementary sources of information. The other two indicators would be: (1) an index of the "total difficulty" of a counselor's closed cases over a given time period, and (2) an index of "average difficulty" (i.e., the weighting factor used for that counselor's "average" closed case).

A more complex scheme has been proposed by Serot (1972), involving three indices that are normalized to reflect the average performance for a State, district, or individual field office. One index is the ratio of "actual" 26-closures to those that would be "predicted" based on the counselor's or district's caseload mix. For each category of clients, the predicted number of closures is equal to the total number actually served from that group, multiplied by the group's average success rate in the system as a whole.

Using a similar case-mix adjustment, the second ratio compares actual to predicted costs of services. The third index is a ratio of the first two -- roughly a cost-effectiveness measure.

Note that the Serot approach does not attempt to develop a weighting co-efficient for the total number of 26-closures, but instead converts the number of closures itself to a normalized index, which then becomes one measure of performance.

One advantage of the Serot indices is that their method of aggregation makes it relatively easy to determine the extent to which differences in performance are probably due to case-mix effects rather than to different rates of success achieved by counselors with similar clients. For more detailed analyses of case-mix effects, the client-groupings could be based not only on disability type, but on other

characteristics as well. The degree of disaggregation could also be increased, subject only to computer capacity. Since the indices are formed by summing across a number of client groups, it would also be possible to include additional weighting factors, to adjust the straight success votes for changes in client earnings (a proxy for client benefits), or for the relative priority attached to each group by agency policy.

On the other hand, the Serot index shares one problem of the Sermon measure: it is based on the average experience of clients who have already been accepted for services.* A second possible problem is that even with a fairly high degree of disaggregation, the particular client characteristics chosen to define sub-groups may still be inappropriate. In other words, there might remain considerable variation in success rates or costs within each group. This means that not only would the case-mix adjustment fail to have the intended effect, but also that under such a system counselors would again be encouraged to selectively accept ("cream") the easier clients, this time from within each sub-group. As long as they could identify such easy cases they would produce more closures while still receiving disproportionately high "credit," based on the average experience for clients in that group.

One recent study (Armstrong, 1972) actually applied the Serot index in assessing the relative performance of districts in California.

---

*Note that this is a data problem more than a conceptual one. Unfortunately, the only way to obtain such data would be through an experimental study design that involved accepting all applicants to an agency and then comparing the experience of those who "ordinarily would have been rejected" with those who would have been accepted (Bolton et al., 1968). Of course, even this complicated data-gathering effort would not necessarily adjust for the effects of "informal" screening prior to the application stage.

Not only was there a wide degree of variation within each disability-defined sub-group, but (as expected) the case-mix adjustments were themselves so aggregated that they made little difference in the relative performance of districts.

General critique of weighted closures. A number of other issues related to weighted case closure proposals should be raised at this point, especially since there is widespread interest in the idea and a strong tendency on the part of some observers to see weighted 26-closures as "The Solution" to all the information and incentive problems associated with the current measure of performance.

One fundamental criticism applies to all of the weighted case closure approaches: they are derived from the 26-closure itself and thus retain its two central shortcomings: (1) the possibility of arbitrary decisions on the part of counselors and (2) the primary emphasis on vocational goals as opposed to those other aspects of client gain that are less easily measured.

First, consider problems associated with a weighting formula that would be applied separately to each case. Whether statistical correlation or clinical judgments are used, it is always more difficult to assess an individual's chances than those of a group of similar cases. On the other hand, the "group-average" approach also creates problems, since on top of probabilistic variations within each group some additional error will be created in assigning each case to its appropriate group.

If the purpose of the weighted case closure is in part to alter incentives faced by the counselor in selecting clients and in planning services, the information used for weighting will have to be

collected and the weights computed at intake -- rather than after services have been delivered.* Collecting information in advance, though, raises questions of expense and counselor-time. The information and calculations would probably have to be provided for all clients referred to the agency, since even a preliminary screening process would undermine the purpose of weighted closures as incentive-shifting devices.

Nationwide, there are approximately twice as many (recorded) new referrals as there are cases accepted for services, and three times as many referrals as 26-closures (U.S. RSA, 1972). For the 50%+ of referrals who would not end up accepted for services, agency resources used in computing "weights" will have been wasted. The trade-off here is that between the value of more accurate prediction and screening (but value to whom?!) and the costs of additional data collection and analysis.

There is a further serious question of self-fulfilling expectations that may be created by a weighting system that formalizes, prior to acceptance, the degree of difficulty expected from an individual case. Almost any counselor is able to provide examples of successfully rehabilitated clients who ordinarily would have been predicted to "fail." Perhaps some of these clients were accepted because a perceptive counselor had an intuitive "hunch" that proved to be right. In other cases, though, the saving factor may simply have been that the lack of information about a client's probable chances of success made the counselor ignorant of the "statistical risk" involved in deciding to accept that client.**

---

*Although this latter approach can still be used if the purpose is merely to assign the proper post hoc "credit" to the counselors.

**Of course, a valid argument could also be made that intuition and "conventional wisdom" can themselves be outdated, biased, or simply wrong: note Conley's findings (1969) that -- contrary to the conventional

An alternative to using client characteristics to establish groups of "similar cases" might be to assign the same weight to each administrative grouping. For example, a single weighting factor for difficulty or client gain (etc.) might be applied to the entire caseload of a counselor or even of a district office. The loss in accuracy using these case-mix adjustments might be made up by the administrative convenience of using only one average value. However, it should again be noted that both the Armstrong (1972) study mentioned above and a California agency staff report that attempted to adjust district-level 26-closures for the average skill and experience of the district's counseling staff (Calif. State DR, 1972-A) found that such gross adjustments produced almost no change in the relative rankings of districts.

Even assuming that an appropriate and practical weighting formula can be devised, there is the question raised earlier of how it should be used: to give additional "credit" for closing a fixed number of (more difficult) cases, or alternatively to lower the expected quota of 26-closures for more difficult caseloads.* The incentive of additional credit may be a weak one if it is true that counselors do not try to maximize their statistical performance records, but instead to achieve only the minimum quota to avoid administrative pressures -- while still

---

wisdom -- rehabilitating the more severe and expensive cases may have a higher economic "pay-off" to society than rehabilitating a larger number of quick-and-easy cases. Properly used, though, statistics on success rates might help to inform and reform intuitive judgments or experiential knowledge.

*It is assumed that these "positive" forms of reinforcement would be more effective than any "negative" ones -- i.e., taking away credit or raising quotas for easier or lower-priority caseloads. For at least some counselors, the latter would represent a net loss compared to their present position. For all counselors, it would probably be perceived as an administrative "speed-up," and resisted.

following the norms of "good professional practice" (Pflueger, 1972-B).

On the other hand, the idea of lowering numerical quotas for difficult or high-quality closures would be fought by agency administrators, since it would probably reduce the pressure for overall (numerical) production. It is one thing to "informally" allow supervisors and field administrators to demand different levels of performance from different counselors; it is another to formally acknowledge an easing of the present standards.

Another type of problem has to do with the use of success rates, as a proxy for severity or case difficulty. It is possible that what a low success rate for a specific group of clients really reflects is as much the poor predictability of closure for that group as its severity or difficulty per se. In fact, it was exactly this argument that led to the recent provisions for up to eighteen months of "extended evaluation" for severely disabled clients, during which time services may be delivered even without a formal agency commitment to "accept" the client.

A low level of predictability may also be the result of: wide variations within an identified client group, an initial lack of the information necessary to predict success, or simply poor judgment due to counselor inexperience or an agency's unfamiliarity with a particular type of client. Low success rates due to a counselor's inability to selectively screen clients might also arise for other reasons: an administrative requirement that all applicants from a certain referral source be accepted; the need of a new counselor to build up his caseload

quickly; or the separation of eligibility-determination and service-delivery functions (i.e., the counselor responsible for services and closure does not make the initial decision to accept a client).

Of course, this whole issue of client success rates has been discussed up to now from the viewpoint of the agency or counselor interested in achieving a good performance record. From the client's point of view, any kind of screening may be seen as a device for denying services, rather than for achieving a "good record." The client might thus attribute a low rate of success not to poor prediction at intake, but to inadequate services delivered by the agency, built-in expectations of failure, or other biases on the part of the counselor.*

A further set of issues needs to be raised: whether counselors can in fact be expected to perceive and accurately respond to rather subtle differences in the degree of credit (or quota-reduction) attached to different categories of cases. In the first place, many other factors such as a counselor's personal preferences for (or against) working with some types of clients may enter into his selection and service decisions. Such influences are difficult to even identify, let alone to control. Even if changes in performance measures are intended simply to obtain more accurate and fair information, rather than to shift incentives directly, weighting-factors that are theoretically correct but small in magnitude

---

*There may be considerable truth to this view. One of the central findings of a recent demonstration project in Wood County, Wisconsin, was that rehabilitation services can be effectively extended to many clients who would normally be screened out as "not feasible." Moreover, indicators of client gains, average costs per client, and success rates in the demonstration program were comparable to those experienced with more "traditional" client groups (Hammond et al., 1968; Katz et al., 1971).

may not in fact make enough difference in the comparative rankings to be worth the effort. It is possible that either for purposes of improving information or for changing incentives, the weights themselves would have to be rather extreme.

This in turn raises two questions: first, that of the overall range of the weighting scale. What should be the ratio of the lowest to the highest allowable value? To take one extreme example, should program policy be such that a single case of extraordinary severity could justify a counselor's efforts for an entire year -- compared with thirty or more cases of more modest severity?

The second question is whether counselors' reactions to a weighting system can in fact be adequately anticipated and controlled. Pilot or demonstration projects are not necessarily helpful in predicting these reactions; if field staff do not perceive the changes in performance measures as permanent and "real" they are unlikely to take them seriously, and respond as they would under full-scale implementation. A trial-and-error approach has comparable drawbacks: constant adjustments in the weights, based on experience with the system, would begin to introduce too much change and uncertainty. The counselor, and his supervisor as well, are likely to simply ignore a weighting system where the rules change too often. Finally, if a weighting system is too complex to be intuitively understood, it may simply be dismissed as irrelevant by everyone except the agency's evaluation staff.

Along these lines, a recent survey of counselors in several State agencies (Abt Assoc., 1972) reported that about half of the respondents were, in the abstract, in favor of some type of weighted case

closure system. A similar proportion (not necessarily the same individuals) thought that weighted closures would in fact result in counselors serving different clients (those with severe disabilities). A few of the respondents, however, thought that some counselors, when faced with the complexity and uncertainty of a weighted closure system, might actually react by serving easier cases. A similar view was echoed during informal conversations with rehabilitation counselors in the California agency.

Finally, even if weights for different categories of cases could be explicitly defined and did prove powerful in altering incentives and behavior, it is unlikely that they would long remain to be determined solely on "technical" grounds. Client groups would quickly perceive their own advantages or disadvantages, and begin actively lobbying for readjustments of the weights. Of course, as will be suggested later, one important purpose of performance measures may be precisely that -- to focus explicit attention on underlying issues of value, choice, and resource-allocation; to stimulate widespread debate; and hopefully to provide at least the language and the forum for debate and then a new consensus on program priorities.

Other Approaches to Performance Measurement

While the previous sections were concerned specifically with weighted case closures, there are a number of other proposals aimed at refining or replacing the 26-closure as a performance measure rather than modifying it through a weighting process. Some examples of these other approaches are described below, along with a few of the issues they pose.

Refine/disaggregate the 26-closure. Several strategies have been proposed for adding new closure categories or making finer distinctions among 26-closures, but without attempting to re-aggregate different types of closures through a weighting process.

Hansen (1970) suggests a separate category of closures (along with different eligibility criteria) for those clients for whom the existing standard of "potential for gainful employment" is too restrictive, even though homemaking and unpaid work are now accepted as "gainful occupation" at closure. A similar theme is pursued by Collignon and Thompson (1972), who suggest that additional statuses be recognized as "points of closure" for some clients -- or as intermediate stages of progress for others who need long-term "follow-along" services for their chronic disabilities. Informal discussions with several counselors in the California agency also suggested the need for some additional closure statuses, reflecting "success" in terms of some but not all of the usual 26-closure criteria, for example:

--provision of services, benefits to client, but without completion of plan or achievement of vocational/employment goals.

--achievement of vocational goal (employment for 30+ days), but not directly through the efforts of the rehab agency.

--achievement of vocational goals plus non-vocational gains: a "26-plus" closure.

Still a different approach would be to record 26-closures in separate categories according to the client's wage level or type of occupation at closure, or by client disability. In a limited sense, the latter is already done, through the routine R-300 statistics on "characteristics of clients closed as rehabilitated." What is still needed,

though, is some way of making these separate types of closure less directly additive.

If these multiple closure categories are to be maintained separately, rather than allowed to fuse into a single, aggregate measure of performance, it will be necessary for both the agencies and their outside audiences to develop new capacities for dealing with multiple, non-commensurable measures. One simple first step might be for the agencies to set target quotas for each category of closures or each group of clients. While this is already done on a limited basis in some State agencies, such quotas tend to vary from year to year and are seldom tied to explicit agency policies. Perhaps a more serious problem, though, is that efforts to focus attention on performance measures within separate categories give way too easily to the persistent habit of looking at the total number of 26-closures. The possibilities for using multiple and non-commensurable performance measures will be explored further in the concluding chapter.

Replace the 26-closure. An alternative way of dealing with the problems associated with the 26-closure is to simply replace it -- or attempt to -- with a different performance measure.* Hawryluk (1972) reviewed the usual complaints about the 26-closure system, and offered a list of alternative indicators (with literature references) incorporating

*It is worth asking how feasible this would really be. Not only is the 26-closure familiar and well-established, but it is likely that whatever the future of rehabilitation programs something very similar to the 26-closure will continue to provide the basis for terminating services, to successful as well as unsuccessful clients. Data on such terminations will be available; the question is not how to stop it from being used, but rather how to alter its form or interpretation, how to supplement it -- in short, how to avoid reductionism in evaluating performance.

both vocational and non-vocational benefits to clients. He recommended that more effort be directed to the development of multiple scales of rehabilitation gain as basic measures that might reasonably replace the 26-closure.

One direct approach to an alternative measure has been explored by Reagles et. al. (1970-B). In this pilot study, the attempt was made to construct a single index of "client satisfaction," and then to test its relationship with other more readily observed aspects of the client's status at closure. It was found that the client satisfaction index correlated significantly with: his level of earnings at closure, competitive employment at closure, the cost of services, and the amount of counselor-client contact.

This study of client satisfaction was part of a larger-scale effort within the Wood County demonstration project (mentioned above), aimed at developing a single scale that would represent the degree of client gain along several different dimensions (Reagles et al., 1970-A, 1972). The scale was computed using data from a client questionnaire; the twenty items included in each client's composite score dealt with several aspects of his ability to function in "daily living" situations as well as vocational ones.

The use of changes in client earnings before and after rehabilitation, while earlier offered as one basis for weighting case closures, could also be considered in itself as a distinct performance measure. Changes in earnings have routinely been used as a major element of benefit-cost studies in rehabilitation. In principle, this criterion could also be applied to clients who are not necessarily closed as 26's. In either case, the problems of experimental controls and attributing

gains to rehab services would remain. Of course, the basic difficulty with using changes in earnings as the primary performance measure is that it addresses only a limited set of rehabilitation goals. In the case of some of the newer client groups and closure standards (e.g., homemaking or sheltered employment) it may not be relevant at all.*

Analysis of selection/attrition. Another strategy that does not involve either replacing or (necessarily) changing the 26-closure would simply help people to look in different ways at some of the routinely-collected data. Data from the R-300 records would be used to construct indicators that focus attention on issues traditionally ignored at the policy-making level in rehab programs, such as resource allocation and priorities among clients. One promising approach in this direction involves calculating differences in the rates of attrition among sub-populations of clients. Detailed studies would then be undertaken to determine the source, causes, and remedies for those specific instances of observed attrition judged to be the most serious (for a more detailed account see: Harris, 1973).**

There is little doubt that despite its relative effectiveness -- compared with other programs -- in serving those who do manage to get "into the system," only a small proportion of the potential pool of need for rehabilitation services is actually reached by the State agencies

---

* In some cases, non-wage proxies for "economic productivity" might be developed -- but this then begins to cloud the intuitive appeal and methodological simplicity of the earnings concept.

** Note that this type of analysis would still address only the first of the two sets of performance questions mentioned earlier: who is in fact served and successfully rehabilitated by the program? The questions of how well clients' needs are met and the overall impact of the program would still remain, for other evaluation efforts to grapple with.

in any given year. Estimates of this pool of need vary widely, but even the official (and for once, possibly conservative*) estimates by the Rehabilitation Services Administration (U.S. RSA, 1972) show that only 9.3% (1.1 million) of the 11.9 million individuals considered in need of services and potentially eligible were actually served by rehab agencies during Fiscal 1972. And only 2.7% (326,100) of the total number in need were successfully rehabilitated (26-closures) in that year. Clearly, a great deal of de facto selection of clients is taking place, both before and during the service-delivery process.

The question of who is screened into or out of the program needs to be distinguished from an even more basic one: the total amount of selection that will continue to be necessary -- given needs that exceed resources. To the limited extent that the who-gets-served issue has even been perceived, attention has been confined largely to "creaming," the deliberate selection by a counselor of easier cases -- usually identified in terms of disability.

This concept of creaming, with its implicit negative overtones, needs to be broadened in several ways. First, client selection or attrition actually occurs at several stages of the rehab process, and results from a large number of decisions (or "non-decisions") by various actors within and outside the system. Second, many characteristics of the client (in addition to disability type), of the agency, and of the services delivered should be examined as possible correlates with --

*Although definitions of "the disabled" vary considerably, other nation-wide estimates have run as high as 17+ million in 1966 (Social Security, 1968) and 19.6 million in 1970 (Ridge, 1972). Clearly, higher estimates of need would translate into even lower ratios of program "coverage."

or criteria actually used for -- client selection and attrition.* Finally, the fact remains that some priority in selecting clients will continue to be necessary, and choices will in fact continue to be made. Thus three major tasks of analyzing selection and attrition are: (1) to reveal those choices that are being made implicitly, (2) to encourage the formation of more explicit policies, and (3) to support their adoption within the program.

The analysis of client selection and attrition could take place in two stages. First, data already available from the R-300 reporting system would be first disaggregated and then combined in new ways to form indices of attrition for many sub-groups of the client population. For example, sub-groups defined along one or more of the following dimensions might prove to be relevant:

--disability type and severity

--race, age, sex, income level

--education, training, skills

--prior work experience

--place of residence (urban/rural; city size)

--source of referral

--services provided; type of special program

For each of these sub-groups, a "base population" would be identified, preferably at the stage where an individual is simply considered to be "in need of services." A sample cohort from each of these initial groups would then be monitored through successive stages

*A number of such "screening factors" -- as they are perceived by State agency directors -- are discussed in: National Rehabilitation Association (1964), Chap. 1.

in the service-delivery process, to determine what proportion fail for whatever reason to advance to the next stage, and then ultimately to "successful rehabilitation" (however this is defined).* Comparisons of attrition rates among sub-groups, or for the same group at different stages of the rehabilitation process, would then allow identification of the extent and possible sources of client attrition, as well as a determination of who is most seriously affected.**

This proposed approach for identifying the extent, location, and nature of selection and attrition goes considerably beyond current practice in most State agencies and at the Federal level, which at best involves simply describing the characteristics of rehabilitants compared with those of the total working-age population. One basic difficulty with such a limited approach to attrition analysis is that the demographic composition of the disabled population may differ significantly from that of the total population (Ridge, 1972). A second

---

*This "cohort" approach is preferred to the usual practice of cross-sectional comparisons, which compare the total number of clients in an early stage with the number in a later stage during a given program year. The ratio between the two is then interpreted as a "success rate." Given the (varying) time elapsed from acceptance to closure for each client, such ratios are in fact comparing different individuals in the two stages. This will change the results significantly whenever a program is expanding in size or when the composition of its caseload is changing over time.

**Of course, some of those who voluntarily leave or who are screened out of the program may not need the agency's services, either because: their disabilities are not serious enough, they are receiving appropriate services through other public or private channels, or they have successfully obtained employment (or other goals) on their own. In any case, the proposed analysis would help to identify such clients and to clearly distinguish them -- statistically and then programmatically -- from the much larger group of disabled who need services but fail to receive them.

problem is that this single, cross-sectional comparison gives no clues as to how or why selection and attrition actually take place.

The analysis discussed here could be further extended by applying criteria other than (or in addition to) the 26-closure. For example, one might also ask how many 26-closures from each client sub-group were placed in competitive employment, sheltered employment, homemaking, etc. Or, those closed into competitive employment from each sub-group might be the subjects of a one-year follow-up study, to determine how many in each group had been subject to "post-closure attrition."

After the major sources of attrition in each group ("major" in absolute or percentage terms) have been identified using statistical indicators, a second round of more detailed analyses would be undertaken to determine whether: (a) the observed attrition should in fact be considered a "problem;" (b) if so, what its underlying causes are; and (c) how these can be affected by agency policies or administrative actions.

It would clearly be expensive to conduct thorough "process" analyses of this sort throughout the program -- at least it would be costly compared to the present patterns of monitoring and evaluation. But it is also important to ask whether such an in-depth analysis of attrition might not prove "cost-effective," i.e., yield gains to the agency or perhaps to some other group that would justify the costs. Since this would be difficult to determine in advance, a reasonable strategy might be to undertake only a limited number of such studies, dealing with a few major problem areas identified through the routine scanning of the attrition statistics. It should also be kept in mind that, unlike

the (relatively cheap) monitoring of statistical indicators, such in-depth analyses would not have to be repeated regularly. They could instead be undertaken on a one-shot basis, as staff and funds are available and as problems and policy issues concerning who is served become salient.

Introducing New Performance Measures -- Some Further Issues

One question that persists throughout all of the proposals discussed above is whether performance measures should ideally take the form of a single, composite index covering several key aspects of program performance, or instead a series of indicators intentionally kept separate yet complementary.* The choice of one composite index would assume that acceptable trade-offs -- among different objectives of the program as well as among various client groups -- have somehow been settled in operational terms, through a weighting formula or some other mechanism. As will be suggested later in this paper, resolution of the issues in this form seems unlikely, and perhaps undesirable.

The single composite index might be useful, however, in the role of providing information to outside audiences in order to justify and defend the program. An attempt to use multiple indicators for this purpose might create an information-overload that would mask the essential features of even a successful rehab program. On the other hand, where the primary need is less to provide routine information than to raise new issues and re-educate the program's publics, a careful use of multiple indicators may be the most effective course.

---

*The report of the study group on evaluation for the tenth Institute on Rehabilitation Services (1972) does stress the need for multiple indicators, but without tracing out any of their implications.

A number of other interesting problems may be created by various actors' perceptions of the new performance measures or of the way they are introduced. For example, the absolute size and range of a new measure may need to bear some relationship to the existing 26-closure, if it is to avoid appearing totally alien and "radical" (Pflueger, 1972-B). This is particularly a problem for indices such as those proposed by Sermon (1972) and Serot (1972), which in format have little in common with the familiar "annual quota of 26-closures" (currently around 30 in California).

It is also uncertain how well counselors would respond to the small absolute changes in indices such as Serot's, which in most cases would hover within a few decimal points of 1.00. Fractional changes in such an index, no matter how valid theoretically, lack the intuitive, one-to-one link with real-world events (and counselor actions) of the 26-closure. A further dilemma mentioned by Pflueger (1972-B) is that while any new performance measure cannot be too different from the 26-closure and still be readily accepted -- it may also encounter risks by being too similar! The problem is that the new measure might then highlight the degree to which administrators or counselors had previously been stretching or "gaming" the original performance measures.

It should not be assumed that all of the problems in computing a new performance measure, nor all details of the organization's response to it, can be fully anticipated in advance. This means that <u>continuing</u> adjustments will have to be made, either in the level or interpretation of the measure(s), the methods of collecting and analyzing data, or the uses to which new information is put. Such "fine-tuning" of a new performance measure, as pointed out previously in the case of

weighted case closures, may present a welcomed challenge to the staff analyst but a genuine threat to the counselor and others out in the field. Why play a new game if the rules themselves keep changing? Especially if they change during the period between one's decision to accept a client and the date when a case is successfully closed and one can claim the credit.

Re-thinking the Need for New Measures

The general thrust of this chapter has been that it is a difficult task to design a unitary system of performance measurement -- using the concept of weighted closures or any of the others that have been suggested -- and be able to count on it providing either more accurate and "fairer" information or a stable pattern of incentives and behavior, preferable to those now produced by the 26-closure.

There is also the danger that the introduction of the "right" performance measure will be sought as a panacea, with the implicit assumption that it will provide an automatic way of controlling the system through its pre-determined effects on incentives -- a way of avoiding the difficult, continuing tasks of managing the agency and resolving basic policy issues.

If new performance measures for rehabilitation programs will inevitably fall short of this idealized image, they nonetheless can play a significant role in helping to "loosen up" the system, removing at least some of the worst dis-incentives of the present 26-closure. New performance measures, particularly in the form of disaggregated, multiple indicators and new combinations of existing data, can also have a major impact in raising new issues, educating program participants,

and providing a framework for the building of consensus on the program's goals and priorities. None of this happens "automatically" though; performance measures are at best a tool -- not a machine.

Whether performance measures are to be used to create positive change in rehab programs directly (through new patterns of incentives or the reduction of dis-incentives) or indirectly (through the issues they raise or the management information they provide), the development and use of more effective measures requires some prior understanding of organizational structures and the power relationships within a typical rehabilitation agency, along with the potentials for planned intervention. These topics will be taken up in the following chapter.

CHAPTER 4

THE REHABILITATION AGENCY - STRUCTURE, CONTROL, AND INTERVENTION

Introduction. Theory: Organizational structure and process: Typologies and models; individual and organizational goals; power relationships; control mechanisms; control through incentives. Empirical studies: Power and decision-making in rehabilitation agencies: counselor autonomy; goals and goal-conflicts; "action instruments"; issuances; monitoring; research; funding; technical assistance; personal contacts and interagency coordination. Potential means of intervention: Incentives and sanctions; legislative and administrative regulation; supervision and management; structural reforms; counselor training and recruitment; long-run strategies.

> "...there is little guarantee that authority delegated and re-delegated down from the highest level of the system will necessarily be exercised in the manner intended. ...practical control of the system is not as tight as the formal structure would suggest..."
>
> "...Counselors are able to proceed in the conduct of their work as they deem appropriate and report essentially whatever they deem necessary in order to satisfy the information and evaluative demands of the system."
>
> --Abt Assoc., 1972

Introduction

If performance measures and evaluation in general are to be used in rehabilitation programs, rather than merely written about, we need to understand as much as possible about their potential for generating changes that are intended as well as some that are not. Collignon et al. (1972-B) refer to this as the need for "contingency

planning": asking before the evaluation instruments are designed (and certainly before they are put into use) what the possible responses to a given evaluation finding might be. Even if the analysis does lead to a clear understanding of the problem and to the discovery of a plausible solution, do the necessary skills, resources, and commitment to implement it exist?

An understanding of the forces currently at work in the rehabilitation system thus seems to be important for the present topic, especially since each of the potential roles for performance measures (see Chapter 1) is tied in some way to the larger issue of what changes are desirable and how they can be achieved. The links between performance measures and program change may either be direct (introduce new measures aimed at changing the pattern of incentives) or indirect (provide information to policy-makers, raise new issues, or re-educate program participants and observers).

In spite of the apparent need, the literature contains very few empirical studies or appropriate theoretical models that deal with patterns of relationship and formal or informal mechanisms of control in rehab agencies -- with perhaps two exceptions. One is a major Federally-sponsored study by Abt Associates (1972), which attempts to build a broad policy framework for understanding the service-delivery process, the "support system," and the overall impact and effectiveness of the rehab program nationwide. Much of the material in this chapter is drawn from the Abt study and its supporting survey data.

A second exception is a recent analysis of linkages between organizational characteristics and indicators of performance within

the Washington State program (Collignon and Serot, 1973). The study emphasized statistical (cross-sectional) techniques rather than in-depth analyses of the actual processes at work or observations of structural changes within a single agency over time. Nevertheless, the topic itself -- what features of the program or the agency's structure are associated with different levels of performance -- is one that merits further attention.

In the sections to follow, we will first review some of the relevant literature in organization theory and discuss its applicability to the rehabilitation program. Next, the major findings from the few empirical studies of structure and control in rehab agencies will be summarized. Finally, some concluding thoughts will be offered on alternative strategies for intervening in the system to bring about intended change -- and on the potential and limits of each. The purpose of this chapter is to provide a general sense of what changes in the rehabilitation program are possible, and what some of the change-levers might be. In the concluding chapter, this material will be integrated with the earlier discussion of performance measures -- which can help to determine what changes are needed and desired.

Theory: Organizational Structure and Process

Typologies and models. The literature on organization theory has spawned an impressive number of schemes for classifying organizations. For the present purpose, one of the more useful typologies is that offered by Etzioni (1959) based on an organization's authority and goal structures. He characterizes a "professional organization" as one that creates, spreads, and applies knowledge. Such an organization not only

depends on the expertise of professional staff members, it actually inverts the usual relationship between the administrative and expert functions, between line and staff. Experts, according to Etzioni, tend to occupy the major positions of authority while managers play more of a supporting role to assure an even flow of resources and a suitable work environment for the professionals. While this model might apply to a university or a research hospital, it is appropriate only in a limited sense for rehabilitation agencies: it does suggest the powerful role of the "experts" at lower levels of the organization, rehab counselors in the field.

Perhaps a more appropriate model is found in a later work by Etzioni (1964). Here he defines an intermediate level between professional and bureaucratic organizations: the "semi-professional organization." In contrast with the professional organization, the latter employs professional staff whose training and expertise are somewhat less specialized and who are subject to closer supervision and regulation (but often by former professionals). The examples given are primary schools, nursing staffs, and welfare agencies. Most rehab agencies would also fit the pattern.*

Another theoretical model that provides some insight to rehabilitation agencies is offered by Vinter (1963). He defines "people-changing" organizations as those that seek to alter the attitudes and behavior of specific individuals ("clients") as their central goal,

---

*Perhaps it would be more accurate to describe the occupation of rehabilitation counseling as involved in a process of "professionalization" (Lambert, 1972; Muthard, 1969; Sussman, 1966). Different observers have judged the consequences of this as either good or bad; the important point is that the process itself has helped to raise issues of who is to play the leading role in establishing policies and carrying them out.

rather than as a consequence or prerequisite to other activities. Such organizations are also concerned with the scope and permanence of the changes they hope to bring about. Finally, they attach to their efforts a specific moral quality: the intended changes will be good for those served.

Within this broad category, Vinter further distinguishes between "socialization" and "treatment" organizations. The former provide limited assistance to help "normal" individuals progress toward socially-defined roles, while treatment organizations are more concerned with working major changes in (on?) those individuals defined as "deviant" or "defective" by societal standards.* It is likely that while those inside a rehab agency would prefer to view it as a treatment organization, at least some of its clients may perceive the program as a mechanism of socialization and social control.

Vinter also points out that all people-changing organizations tend to adopt a deliberate structure of staff-client relationships. The exact pattern might vary according to:

--the available "technologies" and resources; the balance between the use of positive and negative sanctions toward the client.

--assumptions about the "causes and cures of deviance" in the client, his motivation, and how much (permanent) change is possible.

--the degree of standardization: whether all clients and all treatment methods are viewed as essentially similar.

Using Vinter's model, the complex, non-standardized, and relatively resource-rich "treatment technologies" that have traditionally characterized rehab programs would seem to fit best with the actual

---

*Also see Biscamp et. al. (1972) and Friedson (1966) on the relationship among concepts of disability, deviance, and stigma.

patterns of decentralized decision-making, relative counselor autonomy, and highly individualized services provided through a one-to-one relationship between counselor and client. It might be hypothesized that any attempt to alter this basic pattern would also require (or provoke) fundamental changes in the treatment methodologies and in assumptions about the uniqueness of each client and his needs.

Hasenfeld (1972) proposes a further refinement of Vinter's model, defining a "people-processing" organization as one that attempts to change -- not the individual but his formal status, social role, or his access to resources. Thus the activities of a people-processing organization might involve "labeling" individuals (as in a university admissions committee) or matching them with available resources or roles (an employment agency). There are some elements of this function in rehab agencies, particularly in the sense that counselors are expected to act as "brokers" and advocates for their clients, improving their information and access to jobs and community resources. Aside from the rather oppressive terminology (people-processing), the problem with this model is that such organizations can too easily become merely formalistic in their labeling or matching methods. In responding to narrowly-defined efficiency demands, they might ignore the diverse needs of clients such as the disabled.

A final model of interest is that proposed by White (1969). He distinguishes between the traditional "bureaucratic" organization and a newer type of "dialectical" organization on the basis of each organization's relationship with its clients rather than its internal structure or goals. In theory, the dialectical organization (towards which White's

bias is clear, if unstated) is unique because it does not subordinate the client to the professional, nor invest the client with unique responsibility ("blame") for his condition or with the sole duty to change. The role of the professional becomes less specialized and his relationship with the client less confined, segmented, and (presumably) "impartial." Clients and staff (both professional and support) are all involved in a democratic, participatory, therapeutic, "open" process. The organization's style is characterized as "dionysian," in contrast with the "apollonian" style of the typical bureaucracy. An example of such a dialectical organization would be the (idealized) open ward in a mental health facility. However, the degree to which this model, even if feasible elsewhere, can be transferred to rehab agencies remains open to question. Also in doubt is whether the model should be transferred, given its leaning toward the "total institution."

Individual and organizational goals. One recurrent theme in the organization models described above is the conflict between "administrative" and "professional" goals. This often translates into a conflict between the needs of the organization as a whole and the desires of its professional participants. In the case of rehabilitation, counselors are demanding independence from regulations and from close supervision, leaving them free to follow (hopefully) the norms of "good professional practice" (Rosengren, 1964). The strength of this movement toward even greater professional autonomy is apparent both in the literature and from even a brief discussion with participants at any level of a rehab agency.

An over-emphasis on either administrative regulation or professional independence can be harmful, though, to the organization as well

as to its clients. Vinter (1963) points out that too much concern with providing a protective environment for autonomous professional practice may distort the organization's original goals -- analogies to the medical profession are perhaps the most obvious. But it is equally clear that too much emphasis on administrative control can also lead to undesired consequences (for example, the practice of "creaming" during client-selection in response to administratively imposed quotas for 26-closures).

On a more general level, a distinction has been made between the "system-maintenance" and "goal-achievement" functions of organizations. Etzioni (1960) criticizes those researchers who evaluate an organization by looking only at its explicit goal-seeking efforts, while overlooking the importance of other functions such as routine administration, internal "integration" and morale-building, resource-seeking, and external defense. These tend to be seen either as a diversion of the organization's resources or as a subversion of its "real" goals. In contrast, Etzioni proposes a "system-view" for evaluating an organization, which openly recognizes these other functions as legitimate and necessary if the organization is to continue functioning effectively -- let alone survive -- long enough to achieve its goals.

This view is shared and amplified somewhat by Thompson (1967). His approach is derived in part from a distinction made by Parsons of three basic functions performed within complex organizations:

--technical functions - the central activity of an organization, using its available resources to produce some output of value: physical commodities, services, information, etc.

--managerial functions - providing administrative services and management control, obtaining resources, and disposing of outputs from the "technical core" functions.

--institutional functions - dealing with the outside environment; providing for legitimation and continued support for the organization as a whole.

One of an organization's most critical needs, in Thompson's model, is for effective managerial and institutional functions that can "buffer" its core technology from the uncertainties and other pressures of its environment -- and can also mediate conflicts within the organization itself. Without these other two functions, the core technology cannot operate effectively and the organization will not achieve its goals.

It was interesting to note in interviews with field-level administrators from the California State agency that they were very conscious of their roles in "buffering" counselors from the demands of higher-level administrators (who in turn, they recognized, had to deal with outside constituencies to obtain resources and protect the program's image). One administrator insisted that he protected his counseling staff from "unnecessary" demands for data and written justifications whenever feasible. Another pointed out that his policy was to instruct counselors to worry about the quality of their casework, and to let him deal with the pressures for numerical production from the office.

Power relationships. Another interesting theme in the literature concerns the relative power and influence at "lower" levels of an organization, including both professional participants and others. Blau and Scott (1962) have theorized that the power of professionals at lower levels of formal organizations stems from their acknowledged ability to deal with non-routine, specialized, and often unpredictable tasks (for example, the determination of client eligibility and the delivery of individualized rehabilitation services). However, this

source of professional power is also perceived as "constantly self-diminishing," since the agency's repeated experience with similar problems over time tends to change at least some areas of uncertainty into routine, "technical" functions (ibid., p. 176).

Vinter (1963) posits a somewhat different view of the source of professional power in treatment organizations. He suggests that it derives from the fact that the "output" of such organizations, changes in people, is by its very nature diverse, poorly-defined, and difficult to measure or to value (compared with the physical output of a factory or the sales of a business firm). Such organizations therefore have difficulty in demonstrating their effectiveness to outside audiences, and must rely in part on the public's faith in the competence and legitimacy of the professionals to justify the organizations' continued existence and support.

This theoretical view offers some interesting insight to the current struggle in rehab over who will take the lead (counselors, administrators, clients, or outside legislators) in defining measures of output and setting standards of performance -- or indeed, whether these will be done at all. According to Vinter's model, it would clearly be to the advantage of professionalizing rehab counselors to keep the definitions of program output and performance somewhat vague in the minds of outside audiences. (At present, though, such outside perceptions tend to be over-simplified rather than vague!)

A somewhat different view is taken by Mechanic (1962). He perceives the power of "lower participants" (not necessarily professionals) as stemming from their superior access to information, their knowledge of the "rules of the game" within a small area of the agency's

operations, and their network of personal contacts and informal coalitions. Mechanic sees all of this in negative terms, as a process that subverts the proper control of the organization from the top and limits its effectiveness. Similar resentments are often found among the top administrators of any formal organization -- rehab agencies being no exception. As will be suggested later, though, there may be alternative ways of responding to the whole question of power and control in rehab agencies, ways that are more realistic and ultimately more productive.

Control mechanisms. Etzioni (1961) suggests a three-way classification of "compliance mechanisms" in organizations, reflecting both the type of power used to control lower-level participants and the kinds of commitment to the organization that are in turn elicited:

--coercive controls - including physical force (or the threat of it) and other direct means of coercion. Lower-level participants tend to be alienated from the organization.

--utilitarian controls - based on the manipulation of material rewards. Participants tend to relate to the organization in terms of "rational calculation" of their own gains and losses (or some "disjointed-incremental" approximation of it).

--normative controls - the appeal here is to personal loyalty and moral commitment. Lower-level participants tend to relate more positively, intensely, and personally to the organization.

Of course, one question that could immediately be raised is whether any of these concepts are entirely appropriate for a professional (or semi-professional) organization like a rehabilitation agency. Using Etzioni's own model, the concept of a compliance structure in such an organization should perhaps also be inverted: as seen by the field-level professionals, the problem would be how to elicit from the administrative structure the necessary support, "buffering," and compliance to their professional norms, without stimulating excessive

regulation. Etzioni himself adds in a later discussion (1964) that the problem of leadership and control need only arise when there is substantial goal-conflict between the organization and its lower participants, and that to some extent recruitment and socialization can reduce the need for continuing, direct control. Whatever the direction of flow of "compliance demands" in a rehab agency, in practice the controls that are used include both utilitarian and normative ones. The exact mix depends on the agency, the individual participants...and one's own view of the processes at work.

Along similar lines, Rosengren (1967) draws a basic distinction between organization control: (a) through regulation and structural changes, and (b) through supervisory "style" and the manipulation of indirect incentives. He suggests that regulation and structure may be appropriate means of control in those organizations with relatively predictable tasks, resources, and levels of output -- as well as a core technology that is essentially rational and controllable. Supervision and incentives, on the other hand, may be more appropriate in organizations where the basic tasks are less specialized, the problems encountered less uniform and predictable, and where lower-level staff in turn seldom exercise close control over their clients. This second set of conditions seems to apply more often to rehabilitation agencies, but as will be suggested later in this chapter and in the next, structural changes may also be useful means of control in some cases. And finally, the deliberate manipulation of incentives is not without its problems.

A similar distinction is offered by Rosner (1968), who identifies two general classes of administrative control structures:

--activity control - specifying procedures, or at least resource inputs.

--visibility of consequences -- an emphasis on measuring and monitoring the outcomes of participant behavior; specifying goals rather than means.

In general, Rosner suggests that either type of administrative control would conflict with innovation and adaptability within an organization. A limited test of this hypothesis, based on the observed rate of acceptance of new drugs in a sample of hospitals (a proxy for "innovation"), showed that the first approach (activity control) was in fact associated with reduced innovation, but that the latter controls (visibility of consequences) actually tended to accompany increased levels of innovation.

The second of Rosner's control structures is elaborated in a model proposed by Scott et al. (1967). They suggest an unusual definition of authority systems, based on the processes of evaluating the performance of an organization's participants. The authors stress that one key assumption underlies their model of authority and control: that in fact participants are aware of and concerned about such evaluations of their performance -- either because of the sanctions they carry, their inherent "symbolic" value, or their contribution to the individuals' self-esteem.

The Scott model posits the existence of sets of "authority rights" linking one individual to another. Each authority right may be tied either to a position or an individual, or the link might change from one task to the next. The aggregate pattern of authority rights in part defines roles and relationships within the organization. Four basic types of authority rights are:

--allocating right - the right to assign a task or goal to an individual in the organization.

--criteria-setting right - the right to specify criteria for evaluating performance, determine their relative importance, and set standards of acceptability.

--sampling right - the right to determine which specific cases or pieces of information will be examined in evaluating performance.

--appraising right - the right to actually apply established criteria, judge performance, and attempt to influence the distribution of organizational sanctions.

These basic concepts may be useful in pointing to some organizational dilemmas that can arise when an authority system based on evaluation linkages is "incompatible" -- that is, when a significant number of participants can no longer maintain their own evaluations at levels they consider acceptable. The authors list four reasons why this might occur:

(1) contradictory evaluations - these may arise when separate, conflicting evaluations are performed for different aspects of one's performance, or by different evaluators.

(2) uncontrollable evaluations - those that focus on processes or require resources that the individual cannot (completely) control, including chance, natural forces, other actors, etc.

(3) unpredictable evaluations - these may occur when the individual isn't aware of the criteria being used to evaluate him, when the criteria are constantly shifting, or when the sampling of his work isn't representative.

(4) unattainable evaluations - here the problem is that the standards themselves (or the individual's personal aspirations) are so high that he cannot achieve a satisfactory evaluation with the resources available to him.

Authority systems that are incompatible for any of the above reasons also tend to be unstable, generating pressures for change within the organization. This might lead to efforts to change the organization's basic structure and processes, its evaluation procedures, or even its goals. Alternatively, the pressures may simply encourage individuals

to leave the organization, or to remain there but lower their personal standards of expectation. Examples of all of these responses can be found in rehab agencies, in the context of a continuing debate over how evaluation should be done, and by whom.

Control through incentives. A strong case is made by Rivlin (1971) for the advantages of decentralized administration when combined with performance measurement and appropriate incentives. She cites the example of manpower programs as fertile ground for applying these principles. In the process of discussing them, though, she develops (at least by implication) a list of assumptions and pre-conditions required before performance measures and incentives can serve as an effective control device:

--goals and priorities must be reasonably well- established.

--there must exist performance measures (or proxies) that adequately reflect these goals.

--participants at all levels of the decentralized system must accurately (and similarly) perceive these measures, pay attention to them, and respond in predictable ways.

--the basic "technology" (strategy) of the program must work, and adequate resources must be on hand or obtainable. In other words, the primary obstacles to improved performance must be a lack of motivation and guidance for lower-level participants.

This list of requirements presents some difficulties, if performance measures are to be used with incentives for the control of a complex program like rehabilitation.

A similar faith in the feasibility of incentives-as-controls is apparent in an article by Clark and Wilson (1961). The authors speak of incentives as the basic "glue" that can bind the motives and behavior

of the individual with those of the organization.* Using a breakdown of incentive types similar to those of Etzioni (above), the authors discuss the relative advantages of incentives that are: material, "solidary" (such as personal relations, or status), or "purposive" (either ideological or selfless).

Perhaps the most emphatic statement of the case for using performance measures plus incentives as control devices is found in the work of Blau and Scott. In his initial development of these concepts, Blau (1956) studied a State employment agency to determine how its processes and actual performance varied in response to deliberate changes in performance measures and the reward structures attached to them.

The situation had some clear parallels to current issues over the use of the 26-closure to measure rehabilitation agency performance. Initially, Blau found that the only measures examined or recorded by the employment agency were related to the number of client interviews conducted by a counselor each month. The result was that counselors attempted to maximize their number of interviews, rather than the number of job placements (in a tight labor market). The agency's response to this was to introduce a new set of eight separate indices, including the number of job placements along with several ratios comparing program "inputs" to proximate "outputs."

The reaction from the field staff was vehement: they charged that program managers were forcing them to "sacrifice quality for quantity," to violate professional standards, and so forth. Shortly,

---

*For an earlier statement of this, using the concept of "organizational equilibrium," see Barnard (1938), Chaps. xi and xvi; and Simon (1957), Chap. VI.

though, counselors not only began to respond to the new measures, but, as the author concluded: "By altering the performance record or the relative emphasis on various factors administrators could induce lower echelons to change their practices immediately" (page 233). Another interesting result according to Blau was an improvement in counselor-supervisor relations. The supervisor could now refer to "the record," rather than being forced to invoke his formal authority or personal judgment alone when called on to criticize a subordinate. However, it also turned out that supervisors did not rely totally on the statistics; to do so would have undermined counselors' respect for their position and judgment.

In some cases, the new performance measures led to fierce competition among counselors in a field unit; in other units (where overall performance levels tended to be higher) counselors actually began to cooperate more, to raise everyone's total score and "beat the system." Interestingly enough, in these more cooperative, high-performance units, supervisors tended to rate individuals mostly on the basis of non-statistical, judgmental factors. In addition, informal peer judgments often inverted the administrative performance standard -- it was assumed that a counselor with a low number of job placements was probably providing the highest "quality of services" to individual clients!*

In their more recent work, Blau and Scott (1962) re-emphasize the advantages of "statistical performance records," as one of several "impersonal" mechanisms of control that allow supervisors to assume the

---

*Similar views were expressed by counselors in the California State rehab agency, but at the same time, both field-level and central office administrators were convinced that qulaity and quantity of a counselor's casework tended to go together.

role of helper rather than constant watchdog and critic. Performance statistics also allow the professional to better understand and accept the organization's goals as "given," and to concern himself more with means. Of course, this goes to the core of some of the current issues in rehab agencies: the ends too often have not been "given" -- or at least not generally agreed-upon.

Another case study involving the deliberate use of performance measures to alter incentives and organizational behavior also comes from the manpower field, in this case the Work Incentive (WIN) program in California (Calif. DHRD, 1971; Siebert, 1972). Once again, it was observed that counselors were concentrating on keeping a large caseload and filling all of the available "slots" in the program -- rather than on assisting clients to complete their training and work preparation and secure permanent jobs. It was inferred that the basic cause for this was that no one higher up in the agency monitored the actual number of job placements.

The solution was to introduce a series of new output measures, which included one composite index along with a separate listing of its components. Factors included in the basic index were: the number of enrollees successfully placed in jobs, the average wage at termination from the program, and the reported welfare savings (one of the program's legislative mandates).

These "output" factors were then adjusted for the corresponding resource inputs to each WIN office: staff hours spent on the program, incentive payments to enrollees, and enrollee man-months (an overall adjustment for caseload size). Each month a new index was computed for

every office in the State. Comparisons were made both with the Statewide average and with that office's level in prior months. The composite index was useful for pointing out if problems existed; the separate indicators were more helpful on the management level for discovering the precise nature and the possible reasons for poor performance.

According to Siebert (interview, 4/73), the introduction of the new performance measures had powerful (and, in general, intended) effects on the behavior of the local WIN staffs. One effect that was not sought was a (suspected) tendency on the part of the staff to select from among the available applicants those who seemed to have a better chance to successfully and rapidly complete the program, thus contributing to the unit's performance record.*

It is worth summarizing some of the key differences between the situation in rehabilitation programs and in the two manpower programs described above -- which seemed to be generally successful in using deliberate changes in performance measures to shift incentives and, ultimately, program performance. One difference is that in the manpower examples the basic problem was fairly well-defined, measurable, and agreed-upon (at least by those initiating the changes): how to direct staff efforts to actual job placements, rather than to "processing" activities alone. In rehabilitation, that initial problem has already

---

*The counterpart of "creaming" in rehab programs. In this case, however, the practice was not uniformly considered to be a problem since the original concept of the program was in part to "cream": accepting those welfare recipients with the greatest chances of being "permanently" removed from the welfare rolls after a fairly brief exposure to training, placement, and related services.

been dealt with by the 26-closure -- an effort which in turn has produced the more subtle and complex problems of who is to be served (in what proportions), how well, and toward what goals? In other words, these second-order issues involve not only basic performance but priorities: the allocation of limited agency resources among potential client groups and among competing program objectives.

It should be pointed out that recent criticisms of manpower programs have been moving in these same directions, questioning: (a) who is actually being served (the hard-core unemployed or those who probably would have found employment on their own?), and (b) the actual degree of impact (some "placements" may consist of temporary jobs lasting only a few days or even a few hours). Attempts to refine the job-placement performance measures in manpower programs to deal with these new issues may be considerably more difficult than introducing the measures in the first place. Of course, a final distinction between the two programs is that rehab, in its range of services as well as client goals, has begun to expand well beyond vocational concerns alone.

Empirical Studies: Power and Decision-making in Rehabilitation Agencies

Many administrators and outside analysts share a common view of the structure of State rehab agencies: their image is one of decentralized administration and autonomy on the part of professional counselors in deciding who is to be served, what mix of services is to be provided, how much will be spent on a given client, and when a case should be closed (as "rehabilitated" or not). For administrators, at least, the undertone is often one of frustration: "There's so much independence out in the

field that we can't get anything done." Many counselors, of course, will emphatically argue exactly the opposite: that there is far too much restriction and control from above to allow them to do their jobs properly.* In spite of these strongly-held opinions, there is surprisingly little data on the nature of power and decision-making in rehab agencies.

One exception to this is the consultant study mentioned earlier, prepared by Abt Associates (1972) under contract with SRS/RSA. One major part of this study dealt with the "support system" (as opposed to the actual service-delivery processes), in terms of perceived patterns of power and influence. The Abt study was based primarily on a series of detailed, structured interviews with a sample of counselors, field office administrators, and State agency directors in ten States (18 separate agencies, including general vocational rehabilitation and blind).** In addition, a sample of more than 350 clients were interviewed in two States: California and Pennsylvania. Unless otherwise indicated, the tentative findings below regarding power relationships in rehab draw heavily upon the Abt study and its supporting data.

Counselor autonomy. Responses to the Abt surveys provide some impressive indicators of the degree to which counselors exercise control over individual management decisions -- and thus, in the aggregate, over

---

*Nevertheless, even they will generally acknowledge that rehab agencies are far less centralized and regulated than most others in the social service field.

**Note that only the directors of State agencies and the top administrator in each field office were interviewed. Obviously there are dangers in using their responses to draw inferences about the views of the agency's central office staff, or of field administrators in general. Nonetheless, we will tentatively draw such inferences below, for lack of other data.

larger policy issues such as priorities among clients or among services.* In one set of questions, actors at each level of the agency were asked who they thought had "final authority" over three types of decisions: (1) priorities for which clients to serve, (2) the maximum amount to be spent on each client (with or without some overall agency-imposed limit), and (3) the allocation of funds among different types of services. The following table shows, for each of these decisions, the percent and number of participants at each level (State directors, local administrators, and counselors) who thought that final authority was in fact highly decentralized (resting either with counselors or with field level administrators).

Table 1

Question, and Respondent Group	% indicating authority at field or counselor levels
1. Priorities for which clients to serve?	
--State directors (n = 12)	34%
--Local administrators (n = 39)	43
--Counselors (n = 65)	61
2. Client expenditures (with limit)?	
--State directors (n = 16)	75%
--Local administrators (n = 52)	98
--Counselors (n = 83)	96
3. Client expenditures (without limit)?	
--State directors (n = 16)	57%
--Local administrators (n = 52)	74
--Counselors (n = 73)	90
4. Allocation among types of services?	
--State directors (n = 12)	25%
--Local administrators (n = 51)	69
--Counselors (n = 72)	77

*Data from an earlier nationwide survey of 90 rehab agencies revealed similar patterns of field autonomy in acceptance, closure, and expenditure decisions (Nat. Rehab. Assn., 1964).

the services provided, rather than to the question of who was to be served by the program. This suggests once again that many of the fundamental value and allocation questions remain to be recognized in some quarters -- let along resolved.

"Action Instruments." The Abt study proposed seven categories of "action instruments" that might serve to link participants at various levels within the rehab system (including those outside of the agency itself). Since the client of the study was RSA, the viewpoint adopted in the analysis was to ask which instruments seem to be the most effective in getting Federal policies and priorities transmitted down to the State, local office, and counselor levels.

One important point not brought out in the report is that those instruments which are most effective in transmitting policies from the Federal to the State central office level (from one administrator to another) may not be the same ones that are best for relaying State agency policies out to field offices and ultimately to the counselors themselves (i.e., from administrators to professional staff). The former task may be fairly easy, using changes in overall funding, earmarking and administrative regulation, or informal contacts and persuasion. However, control exercised from the central office of activities out in the field may be much more difficult, nor is it clear what specific instruments might work best. Above all, there is a serious question of the degree to which activities in the field should be centrally controlled -- regardless of the feasibility of doing so. (See chapter 5.) The following paragraphs summarize the Abt findings regarding each proposed action instrument.

"Issuances." This category includes regulations, guidelines, legislation, and other forms of written information (handbooks, releases, etc.). State directors tend to see issuances as having some influence; local administrators and counselors consider them very weak as policy instruments.

Monitoring. Included here are: data generated routinely through the R-300 reporting system (and State supplements to it), administrative reviews, site visits, and audits. One notable finding is that some actors at each level of the agency see the basic information system as either incomplete or inaccurate. Supplemental requests for data and the use of "informal" channels of information are common. In addition, the Abt study turned up a number of cases where the specific data items that higher-level administrators said they were monitoring turned out to be very different from those that the counselors and the administrative field staff thought they were being monitored upon.*

Research. Questions in this section were addressed to the utilization of research findings in State agencies. A number of the States had "research utilization specialists" on their staffs, but often very little sense of how much "utilization" was actually taking place -- or of what the concept really meant and how to "measure" it. Moreover, it would seem that research inherently has a limited role to play in communicating or enforcing policies. Perhaps it is useful only in those cases where there already exists some strong motivation to change, and all that is needed is a suggestion of what to do plus at least some assurance that a proposed alternative has been successfully tried elsewhere.

*Recall the comments of Scott et. al., (1967) on the problems of incompatible evaluation criteria (see pp. 100-101, above).

Funding. Actions in this category might include: increasing or reducing funds, earmarking or otherwise restricting their use, and adding specific accounting and reporting requirements. All respondents, but particularly State directors, were emphatic in stating that new funds would be a powerful instrument for carrying out Federal policies -- but that threatened reductions would be a very weak tool. It is possible that these responses were somewhat biased by the Directors' suspicions that to identify funding cut-backs as effective might eventually encourage their use!

A different pattern emerged in response to the question about earmarking and other restrictions: State directors tended to see these as strong means, field administrators and counselors as weak ones. A further distinction regarding new funding was made by State directors: between funds for special projects ("innovation and expansion" grants), and additions to the regular "Section 2" allocaticn for each State, which currently provides most of the case-service funds. Funding for special projects was seen as a strong tool for altering priorities; add-ons to Section 2 funds were considered weak, perhaps because they are automatically assumed to be the turf of the well-organized, "traditional" client groups. On the other hand, the influence of new special-project grants might also be limited, since it was generally considered more difficult to obtain matching funds for them from State Legislatures.

Technical assistance. Both counselors and field administrators felt that not enough technical assistance was being provided from higher levels of the agency, but that the TA that was available was generally helpful. Once again, though, it is not clear that technical assistance represents a meaningful way of trying to alter priorities, except in

very limited circumstances. For example, it might be more effective in changing decisions over what services are provided than in altering those concerned with client selection. Technical assistance might also play a useful supporting role, to accompany changes in the information and reporting system.

Personal contacts and interagency coordination. The findings for these two instruments were inconclusive, either because of the wide variations in their perceived effectiveness, the rarity with which they are used at all, or the fact that respondents didn't understand what was meant by the question. A reasonable assumption might be that personal relationships across levels of an organization -- in rehab as elsewhere -- are potentially valuable means of communicating in both directions. The problem with such ties is that they are difficult to deliberately introduce into a large organization. Formal processes of interagency coordination may offer even more tenuous links for policy implementation -- unless they are backed up by the formal authority and personal interest of a top administrator in each agency, and reinforced with strong institutional commitments such as transfers of funds.

Potential Means of Intervention

Using a somewhat different set of categories from those of the Abt study (above) this final section will review several possible mechanisms of control and deliberate change in rehabilitation agencies. Once again, it should be emphasized that the basic question itself may not be entirely appropriate: How to control the system, from the top down, toward pre-determined goals? In response to the material presented below, an alternative way of framing this question will be offered in the final chapter.

Incentives and sanctions. As pointed out previously, incentives for guiding the behavior of lower-level participants can operate in two general ways: (1) the attempt can be made to tie them directly and "automatically" to measures of performance (as with weighted case closures), or (2) the link can be made indirectly, with performance data providing a basis for administrative judgments that in turn can lead to the creation of incentives or to other administrative actions that fit a particular situation. In either case, there is the possibility that the intended incentives will not in fact be accurately perceived and responded to by those at whom they are aimed. This problem might be dealt with more easily under the second approach, where administrative judgment is expected to intervene between performance measures and the response to any incentives they may create.

For incentives to work effectively under either pattern, however, they need to be clearly linked with some form of positive or negative sanction. Three types of sanctions are considered below: recognition, promotion, and budget changes.

The use of recognition as a sanction to change behavior covers a wide range of possible approaches, from formally publishing comparative performance records of counselors (or field offices, etc.) to less formal means of generating and conveying approval or disapproval. One problem with using recognition as a sanction is that over time it may become more difficult to maintain the legitimacy of administratively-defined performance standards -- particularly statistical ones. As rehab counselors become more professionalized, they also tend to seek approval more often from non-administrative sources: fellow-counselors, current journal literature, professional conferences, and the established counselor training programs.

Promotion has been widely discussed as a means of reinforcing incentives at the counselor level, particularly with respect to statistical performance measures. Program lore states that a counselor's chances of promotion are closely linked to his performance record of 26-closures. In fact, the relationship may be considerably more complex (Neuberg, 1973), or at least less consistent. Regardless of the actual significance of a counselor's statistical performance record in decisions regarding his promotion, it is clear that other "judgmental" factors also enter. Any effort to change this decision-making process by making promotion more directly and "mechanically" tied to performance measures, whether new or old, can be expected to encounter sharp resistance not only from counselors but from supervisors and higher administrators, as well.

One further point to be noted is that the promotional ladders in most State agencies presently provide for only one pattern of advancement: from counselor into supervisory and then administrative positions (Sussman, 1966). For many counselors with a "professional orientation," this may represent an extremely limited gain -- advancement in salary and status but at the cost of losing direct contact with counseling work in the field. One widely proposed reform that was strongly supported by several counselors in California was the creation of a genuine "senior counselor" position -- with appropriate recognition and salary levels, but retaining some degree of contact with clients, community resources, and prospective employers. Changes such as this might strengthen the value of promotion as a reward, and thus as a tool for implementing policies.

Budgetary rewards and penalties are also commonly mentioned as useful sanctions for enforcing policies at all levels of an agency. In

spite of the contrary assertions by State agency directors (see page 99), it is possible that budget reductions, or at least the threat of them, would provide some leverage from Federal to State agencies, or from the State directors to the top administrators in each field office. At the supervisory or counselor level, however, the threatened use of budget reductions would probably have little or no effect -- except to damage staff morale. Counselors and other field staff are simply not accustomed to assuming responsibility for obtaining their own resources or for dealing with higher levels of the agency. Instead, they tend to look to their supervisors and other local administrators to provide the resources and "buffering" that allow them to concentrate on actual service-delivery.

On the other hand, for similar reasons the use of budgetary rewards as positive reinforcement may be of limited relevance to counselors, and at intermediate levels in the agency might also be perceived as a mixed blessing. More money can bring with it the expectation of increased "production." This creates a special problem in cases where the barrier to greater production in the first place may not have been funds (alone), but the availability of experienced staff, suitable rehabilitation facilities and service vendors, additional clients who are eligible (and "feasible") for services, or enough job placements for successful rehabilitants.

A further problem with using budgetary rewards is, of course, that there often is not very much unallocated money to begin with, either at the Federal or the State level. About three-fourths of the total RSA budget is inflexible, in the form of formula-allocations to the State agencies. A considerable portion of the remainder may also have strings on it, either because of prior long-term commitments or policy

objectives -- other than that of sanctioning performance incentives -- that require support. Even a change in the grant-matching ratio has become less relevant as a positive incentive, now that the basic rehab services themselves are 80% Federally-funded.

Legislative and administrative regulation. In traditional terms, one obvious means of controlling a program and enforcing policy decisions is through legislation and administrative regulations. To the extent that these remain useful tools in rehab programs they are better for linking administrators at different levels than for enforcing administrative priorities among professional counselors. At the field level, both counselors and support staff have consistently proven ingenious at finding ways to ignore, bypass, or modify most administrative rulings, to match their own needs and perceived responsibilities.

Supervision and management. The two sets of strategies described above (incentives/sanctions and legislative/administrative regulation) are based on the implicit assumption that if "the right intervention" is made once, at a limited number of points in the system, "all else will follow." In other words, the assumption is that a single change in the rules, specifying the reward or penalty for each level of measured performance in the future, will cause the system to continue responding in stable, predictable ways.

An alternative set of assumptions might point to the need for sensitive, continuing supervision that improves the contact between administrators and professional counselors in the field. In this view, performance measures would be seen not as replacements for administrative judgment and action, but as one source of support for them.

the earmarking of funds or of counselor time to serve particular clients, the development of functional specialists within field offices, and the use by each counselor of specialized referral sources. Each of these approaches has already been tried in a number of agencies, but they have seldom been linked together with any explicit policy theme or means of mutual reinforcement. Such structural reforms might potentially be effective in shifting priorities at any level of the agency. But they also have serious shortcomings, in that they risk fragmenting both: (a) the service-delivery process (i.e. program specialization at the cost of flexibility and comprehensiveness) and (b) the program's political constituency (even more than it already is -- witness the rivalry evident in the recent Senate and House hearings on H.R. 8395, especially among the better-established client lobbies).

Counselor training and recruitment. On the surface, it would seem that the RSA-supported graduate training programs for counselors have a potential long-term role to play, in deliberately shifting current patterns of who is trained and recruited into the State agencies, as well as the mix of skills, attitudes, and expectations that new counselors bring with them to the agencies.

However, two recent studies (Biscamp et al., 1972; Lambert, 1972) point to a number of facts that may weaken this assumption. One is an earlier finding by Dishart (1965) that only about 40% of the training programs' graduates are in fact hired by State rehab agencies (although this proportion varies from two-thirds of all graduates in the South to one-third in other regions). An additional one-third of the graduates nationwide are hired by "rehab-related" public or private agencies. Several observers have pointed to another constraint on the use of the

training programs as instruments of (Federal) policy: the lack of contact, and in some cases the outright hostility, between some of the agencies and the training programs.

A third set of problems is related to the proper role of the counselor; differing views on this are often held by the training programs and the agencies (Patterson, 1957). The training programs tend to emphasize the counselor's role as therapist and change-agent. In contrast, the State agencies (and at least some of the militant client groups) may be more interested in a counselor who can act as coordinator and broker, based on his thorough knowledge of community resources, current opportunities in a wide range of vocations, and the specific functional problems faced by disabled individuals (along with ways that these can be overcome).*

In spite of its substantial contribution to funding the graduate programs and its traditional support for many students, RSA apparently hasn't made much use of the training programs as means of carrying out its long-term policy goals. The Federal agency seems to have no well-established procedure for regularly monitoring the training programs -- let alone for in-depth evaluations or site visits (Lambert, 1972). Nevertheless, one writer on the subject (Ayers, 1970) sees strong potential in them. He concludes that the training programs could contribute to a shift in counselor attitudes (at least towards those clients who are disadvantaged) through increased exposure to such individuals in class

---

* A contrasting view of this agency-program split, as seen by faculty members of one of the graduate training programs, is that the agencies really prefer "passive production-workers," while the training programs are trying to produce innovative, sensitive, bureaucracy-resistant advocates for disabled clients...

sessions and practice field work, through "human sensitivity" training, and through the recruitment of more disadvantaged and minority students.

Long-run strategies. There is a broad range of possibilities in this category, but they may be grouped for convenience into two major clusters: (1) strategies concerned with changing values and attitudes, and (2) those aimed at altering the basic balance of power in the rehab program. In the first category, long-run changes in the values and attitudes of counselors, administrators, Congress, the public, and even clients themselves should continue to be pressed through the conventional channels: education, public information campaigns, and active advocacy by the agencies and organized client groups. In addition, new performance measures might be introduced and widely publicized, for the purpose of increasing public awareness of the program's goals, accomplishments, and the key issues that remain to be faced.

Before the necessary changes in attitudes can occur, though, there must be a greater awareness of: the nature and prevalence of disabling conditions; the characteristics shared by physical, mental, and socio-culturally defined handicaps; and the degree to which such handicaps can be reduced or overcome through rehabilitation, allowing individuals to function successfully in the larger society or in sub-cultures of their own choice. An ultimate goal might be to begin redefining the accepted image of rehabilitation programs -- away from that of a "residual," corrective function and towards a "mainstream" one, a normal means for society to meet its accepted obligations to a substantial cross-section of the population (Campbell, 1970).

A second set of long-range strategies is concerned with bringing about fundamental changes in the balances of power in the program --

between counselors and administrators, between clients and the agency, and among client groups themselves. All of these are deeply normative questions; the answers finally arrived at will affect not only the program's effectiveness but its distributional consequences, as well.

My own stance on the counselor-administrator balance is that there is still room for strengthening influence at the lower (field) levels of most rehab agencies, compared with their central administrators. At the same time, this will intensify the need for safeguards and countervailing forces, so that growing strength and professionalization among rehab counselors might avoid -- insofar as possible -- the closed, restrictive patterns that have characterized many of the older professions. Perhaps this means that rehabilitation counselors should remain at some stage of semi-professionalism -- or perhaps it simply underscores the need for a new definition of professionalism, and for greater diversity among the individuals who share that label.

One such safeguard would involve strengthening the influence of clients in the program -- a desirable effort for other reasons as well. A number of means have been proposed for doing this; some appear to be more speculative than others at the present time:

--client review panels.

--client representation on advisory boards.

--independent evaluation studies by client groups.

--a voucher system to encourage "market-type choices" for rehabilitation services.

--direct funding of alternative (client-run) services and facilities.

--client sign-offs required before a case can be closed (accompanied perhaps by an "inverted appeals procedure" for the counselor).

--recruitment of disabled individuals for agency staff positions.

--effective political organizing by additional groups of disabled clients (or potential clients).

A third set of concerns has to do with the relative influence of different organized client groups. There is little doubt that at present the better-established, "traditional" client groups have a more pervasive influence in the program than do those groups only recently organized -- or the even larger number of potential groupings of clients who are now poorly represented, but who share similar needs and interests. Perhaps the only way to approach this issue (other than as an advocate for some specific group) is to argue: (a) for supporting most of those clients who are the most under-represented at any particular time; and (b) for promoting more effective coalitions among separate groups, where this will help to increase the total level of program resources available to all.

CHAPTER 5

TOWARD A GUIDANCE PROCESS FOR REHABILITATION PROGRAMS

Overview. Some strategic considerations. The suggested approach -- key features: determining who is served; determining how well clients are served. Structure and participants: political and organizational responses. Use of results. Conclusion: program evaluation and the future of rehabilitation: dimensions of change; sources of concern; implications for rehab evaluation.

Overview

The preceding chapters have dealt with three major topics: (a) what information is needed and available for evaluating the performance of rehabilitation programs, (b) how it can be collected and analyzed, and (c) what might be done in response to the findings -- given the means available for intervention and the potential for change in the system.

Unfortunately, the bulk of writing and thinking on program evaluation has maintained the traditional boundary between those first two functions and the third -- between understanding the program and taking appropriate actions based on that understanding. This distinction is by now a familiar one -- in Western cultures, at least. It takes a variety of forms in the literature and lore of program analysis: research divorced from its "utilization," evaluation studies without proper "implementation," the role of the expert as opposed to that of the manager.

One of the themes implicit in this final chapter is the importance of trying to break down this distinction with respect to major social programs such as rehabilitation.* Rather than perpetuate this self-imposed and unnecessary barrier between evaluation and active response -- a barrier expressed in terms of the activities and methods, timing, and actual participants in each -- we need to invent and then deliberately experiment with new ways of linking information and action, shaping these finally into coherent, ongoing processes of "program guidance."

A second, related theme involves the need to free ourselves from several assumptions that are enticing but often dangerous in a complex program like rehab, assumptions of: stability, predictability, reliability of control mechanisms, and universal agreement on goals and priorities. While it is easy to think of the cases where these assumptions have not worked in the past, it is far more difficult to actually abandon them in the present -- let alone to specify the alternative assumptions that we should adopt, or to understand what these would mean for our efforts to manage and change programs.

A few of these alternative assumptions, however, can at least be perceived in vague outline. They assume, first of all, a guidance process for rehabilitation agencies that does not stress one single source of control and innovation. Instead, it would favor processes that:

--are decentralized and pluralistic.

---

*A similar view is taken by Rusalem (1969), in exploring the concept of the "rehabilitation practitioner as researcher."

--look predominantly to the field for new ideas and for the necessary commitments to change.

--reward a greater degree of risk-taking and diversity.

--rely on learning and adaptation as well as the pursuit of pre-planned policies.

The intent here is not to advocate an organization without coherence or common purpose, but rather one that can begin to draw its sense of direction and its strength from new and multiple sources. It is also accepted that any public agency must continue to recognize and deal with other very real organizational needs -- for example, those of public accountability and continued public support. To the extent that these conflict with the agency's internal needs for survival and for adaptive change, such conflicts must first be recognized before they can be dealt with. In particular, there is a need to more clearly define the distinct roles that performance measures and evaluation processes can play in serving internal as well as external needs, and then to strike some acceptable balance where these roles conflict.

A further important conclusion, in the context of the present study, is that the most significant roles for performance measures in rehab programs -- and the most critical reasons why change is needed in the current measures -- are perhaps not those most commonly thought of. Given the program's complexity and continued evolution, the multiplicity of its (often unstated) goals, and the diverse needs of its clientele, no single composite index of performance can provide a meaningful picture of either: (a) what is happening in the program or (b) what parameters to change in order to produce different outcomes. Nor can any change in performance measures be effective, as an isolated strategy, in deliberately shifting the priorities of program participants in their decisions about who will be served and in what ways.

The internal information role of performance measures can best be served through more sensitive, disaggregated, multiple measures that are flexible enough to evolve over time, and at all times are intimately linked to participatory processes of evaluation and program guidance. What is needed in the guidance role is the use of appropriate performance indicators -- not as "automatic" cues for incentives and sanctions -- but as aids and complements to sound administrative judgment and response.*

Finally, among the most important but difficult-to-specify roles of performance measures are those of raising issues, provoking new insights, and re-educating both inside participants and outside observers to the questions that need to be faced.

This chapter, then, represents a first effort to describe how performance measures might function in various roles, and the possible consequences of introducing new measures as one element of an effective guidance process for a rehabilitation agency.

The aim in these final pages is not to provide a detailed recipe for the "model" system, but rather a broad view of the points that should be considered and a general framework for thinking about them. The details themselves should remain to be developed by those individuals directly involved in trying to strengthen the monitoring, evaluation, and administrative processes in a particular State agency. Hopefully, this would mean, once again, a larger number and variety of participants

---

*Initially, most performance indicators may not be intended as mechanical substitutes for judgment and administrative action. The problems arise in their actual use. If a new indicator is accepted at all, it is more than likely to be over-accepted -- probably because it was over-sold. There seems to be a natural impulse to view any new indicator as a comprehensive cure for all of the program's ills, a tendency that clearly leads to early disillusionment.

than in the past. It should also be clear that I am not suggesting any fundamentally new techniques for performance measurement, but instead new ways of thinking about the roles that such measures can play as part of an active process of program guidance.

Some Strategic Considerations

Two basic features of the information-and-evaluation functions in a rehab program should be multiplicity in the measures used and pluralist participation in the processes themselves. Multiple measures can express complexity and variety more easily than can a single measure; they are also easier to change and to supplement over time; and they can encourage participants to begin looking at the program from a variety of new points of view. For similar reasons, there is a need for more diverse actors to participate in analyzing data and drawing judgments from it. This not only allows a wider range of opinions to be considered, but even more importantly it helps to limit the influence of any single "expert." A further advantage of introducing multiple measures and broader participation in program evaluation is the increased probability of generating new alternatives and innovative ideas -- a task equally as important as the assessment of existing ones.*

As suggested above, the designer/implementer of new program evaluation processes would be wise not to attempt to define the details of "the" perfect system -- there will always be too much uncertainty over

---

* Of course, multiple indicators can also be used to resist innovation, since criticism based on one of the measures may be countered by reference to another, more favorable one. This only underscores the need to accompany new, multiple indicators with an effort to re-educate program participants to make them more comfortable in dealing with multiple dimensions of good (or poor) performance.

who will actually be conducting analyses in the future, for what purposes, and under what differing conditions. Instead, efforts might be more profitably directed toward: (a) clearly spelling out a few of the most essential features of a new system, (b) attempting to build in dependable balancing mechanisms, and (c) providing for adequate channels of access for new ideas. In short, the design emphasis should be less on trying to optimize than on controlling for possible "damage" -- by anticipating and heading off some possible mis-uses of the proposed evaluation system (for example, the "reductionist" tendencies to over-simplify and over-aggregate performance measures in practice).

Having observed this warning, there are at least three general strategies that might be used to improve the current monitoring and evaluation practices in a State agency:

(1) Refine the measurement techniques, along with the concepts underlying them. One might, for example, attempt to improve the accuracy, validity, and comprehensiveness of the data itself, as well as the manner in which it is made available to users (clarity, format, timeliness, etc.).

(2) Change the specific individuals involved (or re-educate the same ones): those who make judgments and take action based on the performance information. Included here would be strategies for routinely involving clients, counselors, supervisors, and administrators (etc.) in monitoring and evaluation -- through review panels, contracts for special studies, or a broadening of the staff recruitment process.

(3) A significant departure from either of the first two alternatives might be to abandon the whole notion of evaluation-and-response as a way of improving the program's guidance process, and instead try to build in an "internalized" means of feedback and "self-regulation." Examples of this strategy might include the creation of markets or near-markets for rehabilitation services (vouchers or direct income transfers to clients), agency contracts with "counselor-entrepreneurs," the shift of greater control over service-termination to clients (a required sign-off process), or the badly-needed efforts to strengthen the appeals procedures for clients (Pflueger, 1972-A).

The strategies within each of these clusters might measure up differently in terms of various goals: improving program effectiveness, providing for "external defense," raising key issues, and assuring a minimal level of equity in allocating the rehab program's limited resources. Strategies from the first two clusters (refine the techniques, change the individuals) are recognizable in several of the proposals outlined below. The third set of more fundamental changes is not well represented here, although many of its specific proposals may eventually emerge from the general "opening up" of the system to new ideas and new participants.

The Suggested Approach -- Key Features

Determining who is served. The basic analytic tool for dealing with this set of questions should be a systematic study of the attrition and selection processes at work in the State agency as a whole and within each of its districts or operating units (see Chapter 3). In most cases, an analysis of this sort could be performed using data already available within the Federal/State reporting system. The first stage of the analysis would use statistical data to investigate two types of questions. One is the total number (and percentage) of clients, from an initial "cohort group" of those in need of services, who for any reason fail to get into the rehab program or who become program drop-outs somewhere along the way to "successful" closure.

A second set of questions concerns the characteristics of those who are selected out or drop out at each stage of the rehabilitation process. In addition to the usual concerns for success rates by disability, proper examination of the data should uncover any significant differences in attrition rates that correspond to other groupings of

clients (by race, age, sex, urban vs. rural residence, etc.) or to features of the service process (source of referral, type of services provided, etc.).

Once a scanning of this statistical data has identified possible problem areas, more detailed analytical studies will be needed to understand the actual mechanisms at work and how to intervene to change the results. At this stage, some approaches might include: follow-up studies of a sample of clients, detailed case studies of the decision (or non-decision) processes that influence attrition, direct field observation, interviews with counselors and supervisors, and so forth. In particular, in-depth surveys of a sample of program drop-outs to obtain _their_ views of the program may provide unique insights to the entire screening and attrition process (Rabinowitz, _et al._, 1964).

Some of these "process" studies would involve considerable expense. In each case, judgments will have to be made whether the expected pay-offs are worth the investment of agency resources (political and organizational resources as well as financial ones) and whether in fact there seem likely to be adequate motivation and resources to implement a solution if one were generated.

All of these approaches emphasize the raising of issues -- involving resource-allocation, program priorities, variations in the extent of client need actually served -- that have up to now tended to be ignored in explicit policy discussions for rehabilitation programs. Of course, _de facto_ allocation and priority decisions have continued to be made as a result of literally hundreds of thousands of separate decisions by counselors and others out in the field. Raising these issues more openly will _not_ produce either instant consensus on goals

and priorities nor (even where limited consensus may emerge) automatic compliance by all those who make or influence actual decisions. But it can begin to shape some limited agreements on what to do in the worst situations, and on the general directions of change. More importantly, it will help to make under- or over-representation of some client groups a new focus for attention, analysis, and policy choice, and hopefully encourage an early identification of emerging problems before they become critical.

Determining how well clients are served. Here the questions for analysis fall into two broad categories: (a) What are the actual processes of service delivery and how can they be improved in quality, scope, and efficiency? (b) What are the outcomes of the rehab program; what is the overall impact of services on the well-being of clients, their families, and the broader community? The answers to either set of questions may lead either toward incremental, "administrative" changes in the present program or toward the need to plan and initiate entirely new strategies and service-delivery mechanisms.

The process questions can best be dealt with through a continuation and strengthening of practices already evident in most State agencies:

--review of case records.

--periodic checks on the caseflow process (to identify bottlenecks, monitor lead indicators such as referrals or new plans, and spot individual cases that are moving more slowly than anticipated).

--monitoring of an (expanded) array of "management indicators" and operating ratios.

As already suggested, though, there would still need to be some major differences in the processes of monitoring and evaluation, the participants, and the ways in which the resultant information is used.

In addition to broadening participation at all levels of the agency in both the analysis of data and its interpretation, the findings themselves should be disseminated rapidly and in a format that is easy to understand and use. The dangers of information overload must be recognized, however (particularly as multiple indicators come into wider use), along with the fact that different participants may have very different information needs (for a recent effort to deal with these issues, see: Calif. State DR, 1972-C). Finally, it should be clearly understood that information on service-delivery processes must be used to supplement rather than to supplant experienced judgments.

Adequate treatment of the program impact questions, on the other hand, will require more far-reaching changes in the current monitoring-evaluation practices of most State agencies. There is little doubt that the 26-closure is inadequate as a single measure of program outcome. It may in fact be harmful, to the extent that it creates dis-incentives for providing costly or time-consuming services to many of those clients who may be in greatest "need" -- but who also tend to have lower chances of success.

Many of these same problems, however, might persist under some of the proposals for adjusting or replacing the 26-closure -- especially those that call for the calculation of a single index of "weighted case closures." The weighted closure approaches have an additional shortcoming: most of them seem to further submerge -- rather than to help expose -- the underlying issues of resource-allocation and program priorities. For example, the concept of weighting 26-closures by success rate (or "risk") manages to sidestep the entire question of whether those client groups whose cases are inherently less "predictable" should

receive less emphasis from the program -- and if so, which clients should be favored, and to what degree?

Instead, it would seem that the best alternative to the current 26-closure for measuring program outcome might be a series of performance indicators, intentionally kept disaggregated and non-commensurable so that they cannot be simplistically reduced to one composite index. Indeed, under this approach several alternative weighted closure indices might be calculated -- but only if they are explicitly presented as part of a set of multiple indicators. Similarly, separate counts might be kept of case closures according to: the type of services provided, occupation or other status at closure (competitive employment vs. other activities), the client group or special program involved, and so forth. Clearly, this attempt to disaggregate the basic 26-closure totals (without weighting and then re-aggregating them for direct comparison) will have to overcome strong resistance and reverse some well-established trends.

Both program managers and outside observers might tend to resist the introduction of multiple measures and disaggregated indices of this sort since they would make direct comparisons among States, districts, or counselors much less mechanical. This objection is not without some validity; there is a need for social programs to remain accountable to their outside publics in terms that can be understood -- and simple comparisons among agencies and operating units can provide some useful standards for judging what levels of performance are possible as well as desirable. But, on the other hand, such outside judgments and comparisons have too often in the past been made on the basis of a narrow range of information, indicators that are over-simplified or too aggregate,

and a limited view of the program's goals. Or the data itself has been used merely to legitimate a decision already made on other grounds.

As long as multiple indicators of performance provide for some links among programs and allow for valid, limited-purpose comparisons, they can avoid most criticisms of irrelevance or confusion -- and at the same time serve to re-educate outside observers and inside participants in the actual complexity of a program like rehabilitation.

While more complex, sensitive, multiple indicators of the type suggested above can provide a continued flow of information for both internal and external uses, there will also remain a significant role for one-time, special purpose studies investigating particular problems in depth. These might include:

--benefit-cost studies similar to those that have long been popular in rehab programs at both Federal and State levels (but too often only to provide justification for the program's funding requests).

--continued research into the factors associated with (and those actually responsible for) successful rehabilitation in individual cases and effective performance for the agency as a whole.

--demonstration projects, accompanied by experimental or quasi-experimental evaluation designs, for developing and testing new service strategies.

Structure and Participants

To summarize several points made previously, the purposes of broadening participation in program evaluation are:

--to introduce new, innovative ideas and to insure that problems are considered from a variety of viewpoints.

--to raise important issues, bringing out conflicts of interest and then helping to structure and guide them towards resolution.

--to involve those actors who will eventually share responsibility for carrying out policies, hopefully to gain their advance commitment to the recommendations that will emerge.

While increased participation is called for during both the evaluation and response stages of program guidance, merely inviting clients, field staff, and administrators to all become "involved" is no guarantee that such a process will actually occur -- or that it will be balanced, equitable, and non-trivial. One way of encouraging the latter (non-triviality) is to insist that new participants also participate in the initial staff-level work of collecting and analyzing the data -- rather than simply being appointed to advisory committees which meet occasionally to hear and react to prepared reports. Broader participation at these initial stages of analysis might be achieved through staff-exchanges, new recruitment policies, the funding of alternative (parallel) studies at different levels and locations in the agency (field units as well as the central office), or the use of consulting contracts with outside groups (client and counselor organizations). Once again, these strategies are expensive and time-consuming -- but the failure to adopt at least some of them may also be expensive in the long run.

It is equally important that the participation process be balanced rather than dominated by one or two groups, if a genuine interplay of interests and the building of consensus are to occur. Traditionally, the major problem may have been that the rehab agency's staff would threaten to override the voice of clients in the decision-making process. Now, however, there are additional possibilities for imbalance among client groups, even in a nominally "open" process: the better-established (and often more conservative) client organizations tend to have more influence than newly-formed advocacy groups -- even though the latter may potentially represent a much larger segment of the disabled.

Seen in this light, the commonly-raised questions of (a) whether program evaluation is best done by "insiders" (agency staff) or "outsiders," and (b) whether evaluation staff should come from "technical" or "programmatic" backgrounds can be seen in a different context. The response in either case is that both types are needed -- but as only one element of a considerably broader mix.

Political and organizational responses. The introduction of new performance measures and new processes of analysis and program guidance will inevitably encounter resistance (as well as generate support) from a variety of sources within and outside a State agency. These varying responses, and their net outcome, are difficult to predict in advance. They depend not only on the specific changes proposed in performance evaluation, but also on how these are perceived by program participants, and on the resources and motivations of each actor. In spite of this uncertainty, it may be interesting to speculate briefly on some possible responses of a few key actors to proposed changes in current performance measures and evaluation processes along the lines suggested earlier.

Federal agencies - RSA and SRS would be expected to have mixed reactions to proposals of this type. On the one hand, they would see the strengthening of State agency capabilities in program evaluation, policy-making, and administrative guidance as highly desirable. On the other hand, the specific reforms suggested here aim toward making the performance measures less standardized, less easily compared among States or with pre-determined norms, and probably less suitable for external use (including review at the Federal level) than for internal program guidance. Similarly, the approach here is based on the premise that rehabilitation programs in each state should take advantage of their existing degree of

diversity and decentralization, and not attempt to exert greater control from above in detailed, pre-determined directions. Clearly, this whole thrust seems to place limits on a more intensive policy role for the Federal agencies.

State agency directors - Some agency directors, whether "strong" or "weak" in conventional terms, might see the new emphasis on multiple performance indicators and participatory processes as a further erosion of their own positions in the agency. Many directors might also feel uncomfortable adopting policies that explicitly acknowledged change and ambiguity and attempted to deal with these using open, pluralist forms of decision-making and response. On the other hand, such changes might be welcomed by those directors who are anxious to evade their own leadership responsibilities, as well as by those who are confident enough to willingly broaden and multiply the channels of influence in their agencies.

Analytical staff - Program analysts in the agency's central office would generally stand to gain from any major increases in the level of complexity or influence of performance measures. They might also be likely to support such efforts purely on the basis of their professional concern for improving the technical quality of program evaluation. On the other hand, the specific approaches proposed here would require that staff analysts begin to share their functions with new participants -- thus weakening their professional claims to unique expertise. These new approaches would also require the staff to accept a greater role for normative and political factors in policy decisions -- as at least co-equals with the "technical" findings from evaluation and monitoring. While in fact this may have always been the case, acknowledging it openly is still another matter. Some analysts will be better-prepared to do so than will others.

Field-level administrators - Counselor-supervisors and field office administrators may have the greatest potential for either benefitting or being threatened by the proposed new approaches, depending on their own responses and on other conditions in the agency. On the one hand, the new multiplicity of performance measures will provide them with a far richer potential source of information on how well their counselors are actually performing, and with better clues as to what should be done to solve specific problems. At the same time, though, supervisors and field administrators will have to become more skillful in interpreting and using multiple, disaggregated data sources, and will have to avoid being overwhelmed by the volume of data and the absence of simple decision rules. While they will thus have new tools to work with, field-level administrators will also be faced with new pressures and with the need to deal with a wider array of participants, as well.

Counselors - In theory, counselors stand to gain a great deal from the proposals for more comprehensive, pluralist, and decentralized evaluation and decision-making functions. At the very least, they should benefit from any decline in the dominant position of the 26-closure -- even if the alternatives at first prove less than ideal. However, other factors may offset this advantage, raising some doubt about the degree of support that can actually be expected from counselors.

One such factor is the traditional aversion of many counselors to any form of systematic measurement and analysis. They tend to see this as an infringement on their professional prerogatives or as a challenge to basic humanitarian motives (i.e.- "You can't put a dollar value on helping another human being to live a full life!"). In addition, counselors are often (deservedly) wary of any proposed changes

in program monitoring and evaluation; in the past such changes seem to have led invariably to more restrictions and more paper-work. As much as counselors dislike the present system of 26-closures-plus-quotas, it is at least familiar and predictable. Either the new proposal might be worse (inaccurate, too complex, or too time-consuming), or else it might actually prove effective -- and thus provide a more powerful tool for administrative control from above!

Many counselors could also feel threatened by the idea of direct involvement by new participants (including clients) in evaluating agency performance -- which in practice often translates into evaluating counselor performance. The proposals here are aimed in part at raising issues that have previously been less visible, and at focusing widespread attention on tacit decisions and their consequences. This may represent a direct challenge to some counselors, particularly those who are less "competent" (under a variety of definitions) or who have simply become accustomed to making decisions (on client-selection, services, and closure) that they might find difficult to justify.

Client groups - On the whole, clients should also be major beneficiaries of the proposed evaluation approaches, given the increased emphases on the quality of services and on active client participation in analysis, decision-making, and administrative response. However, there would also be some powerful redistributive effects among groups of clients: for example, those arising from the systematic analysis of selection and attrition rates. The implications are that some client groups, already relatively well-off in terms of the proportion of their "need" that is being successfully served, might actively oppose either the analysis itself or any response to its findings. On still another level, the very process of analyzing attrition and publicizing the results

may itself encourage new groups or coalitions of clients to form -- whether along traditional lines of disability or according to some other shared characteristics or common goals. This would inevitably change the internal politics of existing client groups, as well as the relationships among them.

Congressional, Legislative, and other outside audiences - Some of the "outside publics" who are closest to the rehabilitation program may welcome the proposed changes. Others who have only a marginal interest or limited knowledge might find it more difficult to deal with the increased complexity of multiple indicators. Their response might be to simply shut out all information except for familiar and intuitively appealing indicators (such as "total number of 26-closures," or "dollars returned to the taxpayer for every dollar invested" -- a variant of the benefit-cost ratio). While data of this sort should continue to be provided precisely for the purpose of maintaining an adequate level of outside understanding and support, it is important to begin a process of re-educating outsiders to new ways of thinking about the program's performance using multiple indicators.

Finally, some opposition is likely to surface from Congressmembers and Legislators who are concerned about the extent of decentralization of policy-making and control functions that is inherent in these proposals. From this point of view, the new approaches would be objectionable not only because they explicitly acknowledge the existing degree of decentralization in rehab agencies -- but because they actually condone the situation and insist on helping it along. Perhaps the only way of responding to this attitude on the part of legislators -- other than through continuing efforts at re-education -- is to provide proof that

even under "traditional" measures of performance the proposed new guidance processes can be shown to work better.

Use of Results

As discussed previously, performance measures can potentially play several distinct roles in a rehabilitation agency. An effective monitoring and evaluation process must clearly recognize this, making allowances for the possibility of direct conflicts (or at least trade-offs) between the internal guidance roles and some of the other uses of performance indicators.

It has already been suggested that one way to do this might be to develop different types of measures for different roles. The information and guidance functions, for example, demand complex, subtle, multiple indicators, while the justification and "defense" of the program to outside audiences may have to rely (for the present, at least) on relatively simpler, more aggregated, and intuitively obvious measures (Collignon, 1973). One clear "macro-task" then, is for all those involved in program evaluation to guard against confusing the two sets of measures. The danger here is that the simplified, aggregated approaches might drive out the more complex ones. A sort of "Gresham's Law of Evaluation."

Underlying this preference for guarding "internal guidance" information from too much outside exposure is one assumption that may not be widely shared by social program analysts and "professional reformers" -- particularly those at the Federal level. This is the assumption that innovation and change in a program like rehab do not originate uniquely, or even primarily, at the higher levels of the system or from policy

experts on the outside. Instead, the emphasis here is far more on discovering how to use performance measures to encourage change from within and from below.*

This preference for internally-based change is not necessarily contradicted by another desire: that of deliberately shifting the balance of power towards one "outside" group, clients. Client involvement must not become just a surface manifestation of still another centrally imposed requirement -- but instead represent a fundamental change in the rules of the game, a way of bringing clients themselves "inside" the system.

In addition to the internal guidance and external justification roles discussed above, a potential role for performance measures that is currently in vogue is to use them as deliberate incentives to shift counselors' behavior in pre-determined ways. It has been argued elsewhere in this paper that this represents a doubtful undertaking -- particularly in the form in which it is most often proposed: a single weighted index of 26-closures. The point is not that such a performance measure might not affect in some way the incentives and behavior of counselors and others, but rather: (a) that it would be far from the only powerful influence on decisions made in client-selection, services, and closure; (b) that any given set of weights will be perceived and responded to in very different ways by different individuals in the agency; and (c) that such responses are not only hard to predict in advance, but difficult to modify once they are established (repeated "adjustments" create too much confusion).

---

* This does not mean to deny valuable roles for the Federal agency and for national client and professional groups, in helping to support and diffuse innovation, to assure a continued supply of resources, and to educate outside publics to accept -- or demand -- change and innovation in the program.

One promising alternative is simply to be more realistic in the demands placed on performance measures in their "incentive-creation" role. Rather than attempting to use them to achieve "positive" purposes (altering counselor decisions toward pre-determined patterns), they might be used for "corrective" purposes -- to ease the worst disincentives faced by counselors. A further advantage of the latter approach is that it does not require total consensus on goals and priorities from the beginning, but merely some limited agreement on what to do in the worst cases, and what the general directions of change should be.

A final role for performance measures is the issue-raising and educative one. In some cases, this role might best be filled by multiple and disaggregated performance measures. In other cases, new ways of combining and looking at existing data (such as in the proposed analysis of selection and attrition) may yield valuable insights or draw attention to questions that have previously been ignored. Still a third way of using performance measures to question and to educate involves inverting familiar patterns of inquiry. For example, rather than persisting in trying to tell counselors how their performance should be evaluated, one might begin by asking them what criteria seem to them to be the most meaningful. How, for example, do they judge how effective a job they are doing?

When actually confronted with this question in informal conversation, several counselors from the California State agency were rather taken aback. After constant exposure to being monitored and evaluated by others, they had not only built up the expected defenses, but had also, it appeared, avoided giving much thought to what effective performance really meant to them. More widespread, systematic efforts to promote such self-examination -- on the part of all those involved in

the rehabilitation program -- might prove to be one of the single most powerful sources for change and renewal.

Conclusion: Program Evaluation and the Future of Rehabilitation

Recently a number of signs have indicated that substantial changes may occur over the next several years in the Federal-State rehabilitation program, changes moving it in the direction of consolidation (or at least closer ties) with other social service and manpower programs. If these signs are accurate, it may be appropriate to question whether any of the issues and strategies for program evaluation outlined in this paper will remain relevant for long. The purpose of this concluding section is to argue that the expected changes may actually increase, rather than diminish, the need to strengthen evaluation and program guidance in rehab along the lines suggested. In particular, I would argue that the participatory, multi-centered strategies discussed earlier would remain relevant for most of the alternative futures that can now be imagined, as well as for the periods of transition that may lie immediately ahead.

Dimensions of change. There are currently signs of change in the rehabilitation program at both the Federal and State levels, different in kind but mutually reinforcing. At the Federal level, the possibilities for change are largely programmatic in nature, ranging from merely relabeling the program or its funding sources to more significant changes in the program's mandate and the use of its funds. Even more fundamental changes would result from proposals to "de-federalize" policy-making and perhaps much of the program's resource base, as well, by including rehab as part of a "human resources" package under special revenue-sharing. Given the decentralized nature of the existing program,

with many policy decisions already being made at the State level or below, a shift to revenue-sharing would have fewer direct effects on day-to-day operations than it would on the balance of political influence and the locus of decision-making.

In contrast, those changes underway or being contemplated at the State level are concerned less with programmatic variables (resources, regulations, service strategies) than with the structure of the agency itself -- although admittedly the line between structural and programmatic changes begins to blur over time. The possibilities here range from mere paper re-organizations at the top administrative levels (often in the form of "umbrellization"), to proposals for more fundamental restructuring that extend all the way down to the operating, field office level.* Either programmatic consolidation (Federal) or structural reorganization (within the States) could occur alone, as both have in the past. If the two do occur together, though, they are likely to interact strongly, reinforcing each other but at the same time perhaps imposing mutual constraints -- if the perception of change by program participants and their ability to cope with it become overloaded.

As of this writing, the patterns of impending change are not yet clear, either at the Federal level or within those States that have shown interest in agency re-organizations. Even so, it might be useful to examine briefly some possible dimensions of program and agency change, to stimulate thinking about how these might affect the roles of program evaluation in the future.

---

* For example, the recent proposal in California to merge rehabilitation and manpower into a new "human resources development" agency, from the top administrators down to the field office staff.

One important dimension involves the pace and timing of such changes: are they all to take place (or be attempted) within a single program year, or instead be spread out over several years? Similarly, will Federally-initiated and State-sponsored changes occur at the same time -- and if so, will they be consciously coordinated or not? Finally, are major restructurings to occur simultaneously in other programs or agencies closely related to rehab, program upon which the rehab program or its clients now depend for resources, supportive services, job placements, referrals, etc.?

These processes of transition are at least as critical as the actual end-goals of change. The period of transition could be characterized by massive disruption in services, a plummeting of agency morale, and a leap in staff turnover -- or it can be preceded by attempts to plan sensitively for change, taking into account the needs and capacity of clients as well as professional staff. If the latter choice is made, it once again calls for a participatory process of analysis, an expectation that the views of all those affected by the program will be taken into account as a matter of course (one characteristic notably absent in most agency re-organizations).

Another important set of variables relates to the scope and depth of the proposed changes: how radical a departure would they represent, whether by intent or as an unexpected consequence? What proportion of the program and its participants would be affected? It is one thing to change the administrative super-structure of an agency or its source of funds -- it is quite another to alter its basic service strategies, the goals and criteria for program success, or the methods of determining who is to be served.

A third set of issues have to do with the effect of program changes on the roles and relative autonomy of those at the Federal, State, and local (field office) levels. It is often assumed that any move to further reduce the Federal role (as under special revenue-sharing) would probably shift more policy influence to the States. In the case of rehab, with its existing pattern of decentralized administration, such a shift might actually mean less autonomy within the State agency for the field-level staff than they currently enjoy. It is also likely that a shift in attention by outside lobby groups toward the State capitols would significantly change the patterns of influence. For example, professional counselor associations might find it easier to organize effectively at the State level than would client groups. In effect, it is possible that any further decline in the current Federal role might have its greatest effect, not in easing restrictions on program activities, but in weakening some important sources of policy innovation.*

Sources of concern. A number of concerns are commonly voiced whenever the subject of structural or programmatic change in rehab is raised, particularly when the proposed change involves consolidation with other social programs. Some resistance naturally stems from the desire of the organization and its existing staff for self-protection -- not only in the sense of job security but also that of preserving the

---

*This stance is not inconsistent with the earlier argument that the origins of many such innovations may lie with clients or field-level staff -- rather than in the central offices of either Federal or State agencies. The point is that the Federal role provides an alternative channel of access for new ideas. Its value is best understood as a potential, rather than an immediate, visible one: with the assurance of multiple channels of access it becomes meaningless for powerful interests to try to capture any one of them.

the disabled could potentially play an important role in encouraging and assisting other client groups to mobilize* -- if they can at the same time avoid being swallowed up by these other groups in terms of sheer numbers.

Another possible problem with rehab being involved in a program consolidation is the threat of diffusing its current mandate: to provide services to specific groups, aiming toward broadly-defined goals, and using flexible strategies and decentralized decision-making. A closely related possibility is that the program might lose its separate "identity" and its generally favorable public image -- becoming "just another [manpower] program." This is most likely to occur if Federally-sponsored program consolidation were accompanied by State-initiated agency mergers; the effects of either one alone would be less powerful.

From the viewpoint of both the rehab agency and its clients, a still more serious situation might arise if consolidation were accompanied by a net reduction in the total resources available to the newly merged programs. Resource constraints would mean increased pressures to compare the effectiveness of the programs in order to develop allocation formulas -- even where programs drawing from the same resource pool in fact served different populations with highly diverse needs. The problem, then, would be similar to that expressed earlier within rehab itself: the danger of making comparisons using criteria that are inappropriate or at least incomplete.

Once again, the best response would be to encourage the use of more sensitive, multiple performance measures -- probably specialized

---

*Even though only a small proportion of the disabled who are potential rehab clients are actively involved in program politics, on the whole they are better-organized than groups like the unemployed -- in part because of the relatively permanent nature of disability and in part due to the tendency toward short-term, sporadic agency-client contact in most programs other than rehab.

current structure of their professional roles, which they perceive as necessary and important. These sources of resistance to change can be expected in most organizational settings; what may intensify them in rehab agencies is that the idea of centrally-planned re-organization runs counter to recent moves toward easing restrictions, enlarging the program's mandate, and allowing more latitude for decisions at the field level. These internal concerns would apply not only to the period of transition, but to the fact that under any form of eventual consolidation with another program, rehab would probably be forced to accept new restrictions and a decline in field level autonomy.

On a different plane, most of those committed to improving the level and quality of rehabilitation services -- including not only agency staff but clients and outside constituents, as well -- worry that re-organization in the form of a program merger might threaten the current role of rehab agencies as "advocates for the special needs of the disabled." The extent to which most rehab agencies are now functioning in such an advocacy role is arguable; perhaps it would be fairer to say that the independent status of the program at least provides a focal point for action and a forum for debate.

A related problem is that any re-organization would be likely to disrupt the existing pattern of client influence in the program -- influence that, while remaining well short of optimal, has grown significantly over the past several years. The extent of damage to the current position of rehab clients depends in part on whether the disabled have sufficient interests in common with (for example) the unemployed -- and on whether such common interests would in fact be perceived by both sides as an adequate basis for a coalition. In a more positive sense,

by program, as well as according to the role the performance measure is to play (as suggested earlier). The aim would be to create an array of indicators with enough linkages among the various programs to provoke thinking about their relative effectiveness -- while at the same time resisting the pressure to combine such indicators into a formula, that would replace judgment with so-called "objective comparisons."

The attempt to use multiple indicators in a judgment-supporting role only would not only benefit the new program cluster as a whole, but in particular would avoid placing unfair burdens on its rehab component. The problem is that, using most simple measures, rehabilitation programs appear to be more costly per client than manpower programs -- while their corresponding benefits to clients (especially the non-vocational benefits) may be more difficult to measure. A participatory, multiple-indicator strategy will not necessarily "solve" this problem but it can at least help to legitimate a broader set of measures -- and measurers.

Implications for rehab evaluation. The above discussion suggests that a minimal strategy for easing the problems that might accompany agency or programmatic consolidation would be for the rehab component to maintain separate or at least supplementary program accounts. Additional rehab-specific performance indicators could then be compiled as necessary. As suggested earlier, such indicators could deal with the use of resources, clients served (who, how many), and measures of impact. Within the rehab component, the presence of these specialized indicators might help to preserve (partially) distinct lines of administrative control -- as well as more autonomy for field-level operations. Outside the program, the existence of such indicators might also help rehab to maintain a separate identity.

An even stronger position is that the rehabilitation component of a consolidated program should attempt not only to preserve its own position, but to actually reform and strengthen the evaluation and guidance processes within the larger program cluster. This could take several forms -- one would be simply to broaden the range of questions asked in these other programs (i.e.- not just the number of clients served and direct program costs, but the "quality" and ultimate impact of services, the degree of change attributable to the programs, coverage of the target population, the choice of which clients to serve, and so forth). Each of these additional questions would require enlarging the inventory of performance indicators -- either by borrowing some of those already in use (or proposed) within rehab itself, or by inventing new ones. Once again, vigorous attempts would be made to assure that the array of separate indicators resisted improper aggregation, either within a component program or among several of them.

The use of multiple indicators for the new cluster of programs would also be compatible with more decentralized forms of decision-making, greater individuation of service strategies, and a balanced increase in the influence of both professional staff (as opposed to central office administrators) and clients. In short, many of the features of the "rehab model" (or the changes currently being recommended) might be effectively transferred in this way to other social programs. Program consolidation could represent an excellent opportunity to do so, and new performance measures may offer some of the necessary tools for change.

SELECTED REFERENCES

Abt Associates. "National Evaluation of the Vocational Rehabilitation Program - Final Report." (DRAFT) Cambridge, unpublished, 1972.

Aiduk, Robert and David Langmeyer. "Prediction of Client Success with Vocational Rehabilitation in a State Mental Hospital." Rehabilitation Counseling Bulletin, 16:1, 9/1972.

Allen, George H. "A Comparison of Processed and Unprocessed Applicants to the Iowa Division of Vocational Rehabilitation." Rehabilitation Counseling Bulletin, 11:3, 3/1968.

Armstrong, Philip A. "Program Analysis: Patterns of Cost, Output, and Productivity Among Districts in a State Rehabilitation Agency." Berkeley: Institute of Urban and Regional Development, WP #204, 11/1972.

Ayers, George E. "Training Counselors for Working with Disadvantaged Clients." Rehabilitation Record, 11:3, 9-10/1970.

Barnard, Chester I. The Functions of the Executive. Cambridge: Harvard University, 1938.

Bellante, Donald M. "A Cost-Benefit Analysis of the Florida Federal-State Vocational Rehabilitation Program." Florida State University, (unpublished Ph.D. dissertation), 1971.

Biscamp, Larry, Judy Taylor, and Herbert Willsmore, with Charles Cole. "An Evaluation of Rehabilitation Counselor Training Programs from the Perspective of Disabled Clients." Berkeley: Institute of Urban and Regional Development, WP #176, 5/1972.

Blau, Peter M. "The Dynamics of Bureaucracy," in William Petersen (ed.), American Social Patterns. Garden City: Doubleday, 1956.

Blau, Peter M. and W. Richard Scott. Formal Organizations. San Francisco: Chandler, 1962.

Bolton, Brian F., Alfred J. Butler, and George N. Wright. "Clinical vs. Statistical Prediction of Client Feasibility." U. of Wisconsin, RRRI, Monograph VII, 1968.

California, State Department of Human Resources Development (DHRD). "Work Incentive Program Production Summary." Project #70-8, 1/1971.

California, State Department of Rehabilitation (DR). "Average Number of Rehabilitations/Man-year." (internal staff memo, Field Services), 9/1972-A.

_____. "Estimated Number of Disabled of Working Age (18-64) in California by Major Disabling Condition." (internal staff report #FSS-72-9-13), 1972-B.

_____. "Making Routine Statistical Reports More Relevant." (internal task force report), 6/1972-C.

California, State. Office of the Legislative Analyst. "Analysis of the Proposed 1973-74 Rehabilitation Budget." 1972.

Campbell, Ann. "Outreach: New Dimensions in Rehabilitation Services for the Physically and Socially Handicapped." Rehabilitation Literature, 31, 6/1970.

Clark, Peter B. and James Q. Wilson. "Incentive Systems: A Theory of Organizations." Administrative Science Quarterly, 6:2, 9/1961.

Collignon, Frederick. "A Brief Review of Evaluation Activity in the Washington State Rehabilitation Program, with Suggestions for Future Directions." Berkeley: Berkeley Planning Associates, mimeo, 12/1972-A.

_____. "An Overview of Program Evaluation Activity in Rehabilitation Services Programs: Current Status and the Problems Ahead." Berkeley: Institute of Urban and Regional Development, WP #207, 2/1973.

Collignon, Frederick and Richard Dodson. "A Cost-Benefit Analysis of the Vocational Rehabilitation Programs of the State of Washington." Berkeley: Berkeley Planning Associates, 1/1973.

Collignon, Frederick and David E. Serot. "An Investigation of the Impact of Organizational Overstructure upon the Performance of State VR Programs." Berkeley: Berkeley Planning Associates, 1/1973.

Collignon, Frederick, Adam Zawada, Barbara Thompson, and Joel Markowitz. "Guidelines and Criteria for Evaluating Vocational Rehabilitation Programs: A Discussion Paper for the Prime Study Group on Program Evaluation, Tenth Institute on Rehabilitation Services," Berkeley: Institute of Urban and Regional Development, WP #173, 4/1972-B.

Collignon, Frederick and Barbara Thompson. "The Evaluation Process in State Vocational Rehabilitation Agencies: Comments for the Prime Study Group on Program Evaluation, Tenth Institute on Rehabilitation Services," Berkeley: Institute of Urban and Regional Development, WP #179, 6/1972-C.

Conley, Ronald W. "A Benefit-Cost Analysis of the Vocational Rehabilitation Program." Journal of Human Resources, IV:2, Spr 1969.

_____. The Economics of Vocational Rehabilitation. Baltimore: Johns Hopkins, 1965.

_____. "Weighted Case Closures - Concepts and Problems." (unpublished, mimeo), 8/1972.

DeMann, Michael M. "A Predictive Study of Rehabilitation Counseling Outcomes." Journal of Counseling Psychology, Win 1963.

DeNeufville, Judith I. "Social Indicators Design and Use: An Interactive Process." Cambridge: M.I.T., (unpublished Ph.D. dissertation), 12/1972.

Devine, John. "Case Weighting Devices as a Differential Measurement of Client Difficulty and Counselor Effort." Sacramento, Calif. State DR, mimeo, 9/1972.

Dieckmann, Betty. "A Suggested Method of Selecting Cases for Caseload Review." Calif. State DR, internal staff memo, (1972?).

Dishart, Martin. "Vital Issues and Recommendations from the 1965 National Institutes for Rehabilitation Research." Washington, National Rehabilitation Association (NRA), 1965.

Dodson, Richard and Charles B. Cole. "An Introduction to Cost-Benefit Analysis of Vocational Rehabilitation Programs: A Model for Use by State Agencies." Berkeley: Institute of Urban and Regional Development, WP #192, 10/1972.

Ehrle, Raymond A. "Quantification of Biographical Data for Predicting Vocational Rehabilitation Success." Journal of Applied Psychology, 48:3, 6/1964.

Etzioni, Amatai. A Comparative Analysis of Complex Organizations. Glencoe: Free Press, 1961.

_____. "Authority Structures and Organizational Effectiveness." Administrative Science Quarterly, 4:1, 6/1959.

_____. Modern Organizations. New Jersey: Prentice-Hall, 1964.

_____. "Policy Research," in American Sociologist, 6/1971.

_____. "Two Approaches to Organizational Analysis: A Critique and A Suggestion." Administrative Science Quarterly, 9/1960.

Friedson, Eliot C. "Disability as Social Deviance," in Marvin B. Sussman (ed.), Sociology and Rehabilitation. American Sociological Association, 1966.

Goff, Clinton C. "An Objective Index for Measuring the Vocational Rehabilitation Counselor's Caseload Difficulty." U. of Oklahoma, (unpublished Master's thesis), 1969.

Goldin, George J. "Rehabilitation Counseling Criteria." Rehabilitation Record, 9-10/1964.

Hammond, Clarence D., George N. Wright, and Alfred J. Butler. "Caseload Feasibility in an Expanded Vocational Rehabilitation Program." U. of Wisconsin, RRRI, Monograph VI, 1968.

Hamrick, Bill W. "Rehabilitation Services and Client Vertical Mobility." Rehabilitation Counseling Bulletin, 9:3, 3/1966.

Hansen, Carl E. "Eligibility - A New Approach." Journal of Rehabilitation, 36:6, 11-12/1970.

Harper, Ronald Beck. "The Design and Application of a Scale for Assigning Differential Weights to Rehabilitation Clients in State-Federal Rehabilitation Agencies." U. of So. Calif., (unpublished Ph.D. dissertation), 2/1972.

Harris, Jeffrey P. "Analyzing Attrition in the Delivery of Rehabilitation Services," U.C. Berkeley: Dept. of City and Regional Planning, unpublished, 3/1973.

Hasenfeld, Yaheskel. "People-processing Organizations: An Exchange." American Sociological Review, 37, 6/1972.

Hawryluk, Alex. "Rehabilitation Gain: A Better Indicator Needed." Journal of Rehabilitation, 38:5, 9-10/1972.

Hefferin, Elizabeth A. and Alfred H. Katz. "Issues and Orientations in the Evaluation of Rehabilitation Programs: A Review Article." Rehabilitation Literature, 32:3-4, March and April, 1971. (two parts)

Institute for Interdisciplinary Studies (I.I.S.). "Toward an Analysis of Federal Rehabilitation Policy." Minneapolis, (unpublished report for project #HEW-OS-71-57), 1971.

Institute on Rehabilitation Services (Tenth), Prime Study Group on Guidelines for Evaluating Vocational Rehabilitation Services and Programs. Program Evaluation: A Beginning Statement. Washington, Government Printing Office, 1972.

Katz, Solly, George N. Wright, and Kenneth W. Reagles. "The Impact of an Expanded Vocational Rehabilitation Program Upon Intra-Agency Processes and Procedures." U. of Wisconsin, RRRI, Monograph XIV, series 2, 1971.

Lambert, Leslie. "The Conflicts Caused by the Mismatch between the Emphases of Graduate Training Programs in Rehabilitation Counseling and Those of RSA." U.C. Berkeley: Dept. of City and Regional Planning, unpublished, 12/1972.

Lawlis, Frank and Jerold D. Bozarth. "Considerations for the Development of Weighting Systems for the Evaluation of Counselor Effectiveness." Rehabilitation Counseling Bulletin, 14:3, 3/1971.

Lesh, Terry V. "Prediction of Vocational Rehabilitation Success." Rehabilitation Counseling Bulletin, 12:1, 9/1968.

Levitan, Sar and Garth L. Mangum. Federal Training and Work Programs in the 1960's - Chapter 7: "Vocational Rehabilitation - The Flexible Approach." Ann Arbor: U. of Mich., 1969.

Markowitz, Joel. "Sheltered Workshops in Vocational Rehabilitation: A Background Paper." Berkeley: Institute of Urban and Regional Development, WP #166, 12/1971.

_____. "Central Policy Issues for the Evaluation of Sheltered Workshops." Berkeley: Institute of Urban and Regional Development, WP #185, 7/1972.

Markowitz, Joel and Frederick Collignon. "A Second Look at Accreditation of Rehabilitation Facilities: Some Questions and Cautions." Berkeley: Institute of Urban and Regional Development, WP #175, 1/1972.

Mechanic, David. "Sources of Power of Lower Participants in Complex Organizations." Administrative Science Quarterly, 12/1962.

Michigan, State Department of Education, Division of Vocational Rehabilitation (DVR). "A Benefit-Cost Analysis of Vocational Rehabilitation Programs in the State of Michigan." (mimeo), 4/1970.

Miller, Leonard A., and Mario G. Barillas. "Using Weighted 26-Closures as a More Adequate Measure of Counselor and Agency Effort in Rehabilitation." Rehabilitation Counseling Bulletin, 11:2, 12/1967.

Muthard, John E. "The State of the Profession," in David Malakin and Herbert Rusalem, Vocational Rehabilitation of the Disabled: An Overview; Chapter XIV. New York: New York University Press, 1969.

National Citizens Advisory Committee on Vocational Rehabilitation (NCACVR). Report... Washington: Government Printing Office, 6/1968.

National Rehabilitation Association (NRA). Patterns of Rehabilitation Services Provided by the 90 State Vocational Rehabilitation Agencies of the United States. (Martin Dishart, Project Director). Washington: National Rehabilitation Association, 12/1964.

Neuberg, Leland G. "Promotion Patterns and Job Satisfaction in the Field Services Section of the California State Department of Rehabilitation: Study Design and Progress Report." U.C. Berkeley: Dept. of City and Regional Planning, unpublished, 3/1973.

Newman, Edward. "Perspective on Rehabilitation Services: A Report by the Commissioner of the Rehabilitation Services Administration." Washington, DHEW/SRS/RSA, (mimeo), 12/1972.

Oklahoma, State. Department of Institutions, Social and Rehabilitation Services (DR). "State of the Art Monograph on Case Difficulty and Client Change." Oklahoma City (unpublished memo from the "Services Outcome Measurement Project," RSA project #EV-3-71, DRAFT), 1972.

Patterson, C. H., "Counselor or Coordinator?" Journal of Rehabilitation, 23:3, 1957.

Perlman, Leonard G., and Kenneth W. Hylbert. "Identifying Potential Dropouts at a Rehabilitation Center." Rehabilitation Counseling Bulletin, 13:2, 12/1969.

Pflueger, Sue. "Performance Measures in Vocational Rehabilitation: The Search for a Measure for Quality of Service." U.C. Berkeley: Dept. of City and Regional Planning, unpublished, 12/1972-A.

_____. "Weighted Case Closure: A Discussion of the Issues Involved in the Search for a New Measure." U.C. Berkeley: Dept. of City and Regional Planning, unpublished, 12/1972-B.

Rabinowitz, Herbert S., Derwood E. Johnson, and Anthony J. Reilly. "A Preliminary Study of Program Drop-outs from VR." Rehabilitation Counseling Bulletin, 9/1964.

Reagles, Kenneth W., George N. Wright, and Alfred J. Butler. "A Scale of Rehabilitation Gain for Clients of an Expanded Vocational Rehabilitation Program." U. of Wisconsin, RRRI, Monograph XIII, series 2, 1970-A.

_____. "Correlates of Client Satisfaction in an Expanded Vocational Rehabilitation Program." U. of Wisconsin, RRRI, Monograph XII, series 2, 1970-B.

_____. "Toward a New Criterion of Vocational Rehabilitation Success." Rehabilitation Counseling Bulletin, 15:4, 6/1972.

Ridge, Susan Shea. "Estimating Need for Rehabilitation Services." Berkeley: Institute of Urban and Regional Development, WP #182, 8/1972.

_____. "Survey of State Evaluation Practices." Berkeley: Institute of Urban and Regional Development, WP #209, 4/1973.

Ridgway, V. F., "Dysfunctional Consequences of Performance Measures." Administrative Science Quarterly, 9/1956.

Rivlin, Alice. Systematic Thinking for Social Action. Washington: Brookings, 1971.

Rosengren, William R. "Structure, Policy, and Style: Strategies of Organizational Control." Administrative Science Quarterly, 12:1, 6/1967.

Rosner, Martin M. "Administrative Controls and Innovation." Behavioral Science, 13:1, 1/1968.

Rossi, Peter H. "Practice, Methods, and Theory in Evaluating Social Action Programs," in James L. Sundquist (ed.), On Fighting Poverty: Perspectives from Experience. New York: Basic Books, 1969.

Rusalem, Herbert. "The Research Role," in David Malikin and Herbert Rusalem, Vocational Rehabilitation of the Disabled: An Overview, Chapter IX. New York: New York University Press, 1969.

Salomone, Paul R. "Client Motivation and Rehabilitation Counseling Outcome." Rehabilitation Counseling Bulletin, 16:1, 9/1972.

Schultze, Charles L. The Politics and Economics of Public Spending. Washington: Brookings, 1968.

Scott, W. Richard, Sanford M. Dornbusch, Bruce C. Busching, and James D. Laing. "Organizational Evaluation and Authority." Administrative Science Quarterly, 12:1, 6/1967.

Sermon, Duane T. "Measuring Caseload Management Effectiveness of Counselors in Cooperative School Programs in Minnesota." Paper presented to National Rehabilitation Association (NRA) Convention, 10/1971.

_____. "The Difficulty Index: An Expanded Measure of Counselor Performance." St. Paul, Minn. State DVR, Research Monograph #1, 3/1972.

Serot, David E. "Indices of Cost, Output, and Productivity for Use in Evaluating Rehabilitation Services Programs." Berkeley: Institute of Urban and Regional Development, WP #187, 8/1972.

Serot, David E., and Frederick C. Collignon. "An Exploratory Cost-Effectiveness Analysis of VR Services." Berkeley: Berkeley Planning Associates, 1/1973.

Siebert, Glenn. "Effectiveness Indicators for HRD Offices: A Systems Approach." U.C. Berkeley: Dept. of City and Regional Planning, unpublished, 5/1972.

_____. "Implementation of Evaluation and the Systems Approach in Government: A Literature Survey and Conceptual Model." Berkeley: Institute of Urban and Regional Development, WP #201, 1/1973.

Silver, Diana Lee. "A Look at the Evaluation of VR Counselor Performance." Journal of Rehabilitation, 35:6, 11-12/1969.

Simon, Herbert A. Administrative Behavior. New York: Free Press, 1957 (Second Edition).

Smits, Stanley J. "The Rehabilitation Supervisor: Within or Beyond Our Reach?" Journal of Rehabilitation, 38:4, 7-8/1972.

Suchman, Edward A. "A Model for Research and Evaluation on Rehabilitation," in Sussman, op. cit., 1966.

_____. "Action for What? A Critique of Evaluative Research," in Carol H. Weiss (ed.), Evaluating Action Programs. Boston: Allyn and Bacon, 1972.

Sussman, Marvin B. "Occupational Sociology and Rehabilitation," in Sussman, op. cit., 1966.

Teitz, Michael B. and Frederick C. Collignon. "Final Report of the Project for Cost Benefit Analysis and Evaluation of Rehabilitation Services." Berkeley: Institute of Urban and Regional Development (forthcoming, 1973).

Thompson, James D. Organizations in Action. New York: McGraw-Hill, 1967.

Thoreson, Richard W., Stanley J. Smits, Alfred J. Butler, and George N. Wright. "Counseling Problems Associated with Client Characteristics." U. of Wisconsin, RRRI, Monograph III, 1968.

U.S. Dept. of Health, Education, and Welfare (DHEW). Office of the Ass't Secretary for Program Coordination. Selected Human Investment Programs. Report #1966-10, 1966.

U.S. Department of Health, Education and Welfare (DHEW), Social and Rehabilitation Services (SRS). Vocational Rehabilitation Manual. (mimeo), 7/1969, as revised.

U.S. Department of Health, Education, and Welfare (DHEW), Vocational Rehabilitation Administration. "An Exploratory Cost-Benefits (sic!) Analysis of Vocational Rehabilitation." Washington, mimeo, 8/1967.

U.S. House of Representatives. Committee on Education and Labor, Select Sub-committee on Education. "Vocational Services to the Handicapped: Hearings on H.R. 8395..." Washington, Government Printing Office, 1972.

U.S. Rehabilitation Services Administration (RSA). "Caseload Statistics: State Vocational Rehabilitation Agencies, 1972." Washington: Government Printing Office, 11/1972.

U.S. Rehabilitation Services Administration (RSA). "Statistical History: Federal-State Program of Vocational Rehabilitation, 1920-1969 - 50 Years." Washington: Government Printing Office, 6/1970.

U.S. Senate. Committee on Labor and Public Welfare, Sub-committee on the Handicapped. "Rehabilitation Act of 1972 - Hearings ...on H.R. 8395..." Washington: G.P.O., 1972 (two parts)

U.S. Social Security Administration. Social Security Survey of the Disabled, 1966; Report #2: "Disability, Work, and Income Maintenance: Prevalence of Disability, 1966." Washington: G.P.O., 5/1968.

Viaille, H. "Operations Research Program in the Oklahoma Vocational Rehabilitation Agency." (Report for VRA grant #RO-946). Oklahoma State DVR, 1968.

Vinter, Robert D. "Analysis of Treatment Organizations." Social Work, 8:3, 7/1963.

Westerheide, William J. "Reliability of a Pretest-Postest [sic!] Services Outcome Measure." Oklahoma State DR, unpublished staff memo, 1972.

White, Orion F., Jr. "The Dialectical Organization: An Alternative to Bureaucracy." Public Administration Review, 29:1, 1-2/1969.

Williams, Walter. Social Policy Research and Analysis: The Experience in the Federal Social Agencies. New York: American Elsevier, 1971.

Wright, George N., Kenneth W. Reagles, and Alfred J. Butler. "The Wood County Project - An Expanded Program of Vocational Rehabilitation." Final Project Report, Volume 1. University of Wisconsin, RRRI, 9/1969.

INTERVIEW SESSIONS

(Interviews and informal discussions, lasting between one-half and two hours, were held with the following individuals:)

Prof. Robert Biller, School of Public Policy, U.C. Berkeley.

Ms. Shirley Bohleen, Calif. State DR, Hayward Office.

Ms. Betty Dieckmann, Regional Administrator, Calif. State DR, Sacramento.

Mr. Ken Englebach, Ass't. to the Regional Administrator, Calif. State DR, Berkeley.

Mr. Norman Evans, Regional Administrator, Calif. State DR, Berkeley.

Mr. Ronald Kaminski, District Administrator, Calif. State DR, San Francisco.

Mr. Phil Ladas, Chief - Program Review and Development, Calif. State DR.

Mr. Leon Lowrey, Ass't. District Administrator, Calif. State DR, San Francisco.

Prof. Lloyd Meadow, Assoc. Coordinator - Rehabilitation Counselor Training Program, Calif. State University, San Francisco (with other faculty members).

Ms. Lillian McDermitt, Counselor, Calif. State DR, San Francisco (with other counselors).

Dr. Paul Mueller, Chief - Research and Statistics, Calif. State DR.

Mr. John Parks, District Administrator, Calif. State DR, Oakland.

Mr. Ed Roberts, Chairman of the Board, Physically Disabled Students Project, U.C. Berkeley.

Mr. Glenn Siebert, U.C. Berkeley (formerly with Calif. State DHRD).

Mr. M. O. Slater, Assistant to the Director, Calif. State DR.

Mr. Morgan Vail, Research Utilization Specialist, Calif. State DR.

Mr. Ray Williams, Field Services Division, Calif. State DR.