DOCUMENT RESUME

ED 063 339 TM 001 357

AUTHOR Walker, Jerry P.
TITLE Installing an Evaluation Capability in an Educational Setting: Barriers and "Caveats."
PUB DATE Apr 72
NOTE 22p.; Paper presented at the annual meeting of the American Educational Research Association (Chicago, Illinois, April 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Decision Making; *Educational Quality; *Evaluation Criteria; *Evaluation Techniques

ABSTRACT

This paper establishes and explains a series of propositions related to the installation of an evaluation system in an educational setting. The propositions are drawn from the problems identified by the Phi Delta Kappa National Study Committee on Evaluation. The paper treats five key problems and selected strategies for their recognition and avoidance. The problems are those of defining evaluation, identifying decision-makers, eliciting the values and criteria on which decisions are based, serving multiple decision levels, and distinguishing between research and evaluation. (Author)

INSTALLING AN EVALUATION CAPABILITY IN AN EDUCATIONAL SETTING: BARRIERS AND CAVEATS

BY

JERRY P. WALKER

THE OHIO STATE UNIVERSITY

A PAPER PRESENTED TO

DIVISION H

THE AMERICAN EDUCATIONAL RESEARCH ASSOCIATION

ANNUAL MEETING

APRIL 1972

INTRODUCTION

The problems of educational evaluation become very real when one attempts to apply an evaluation system to the ongoing operations of an educational institution. They are real because the same problems facing educational evaluation at any theoretical or conceptual level become barriers to the installation of an evaluation system. The recently published PDK monograph, Educational Evaluation and Decision-Making,* is centered around five macro-problems of educational evaluation which provide a meaningful framework for the topic here--barriers in the application and installation of an evaluation capability in an educational setting. The five macro-problems identified in the PDK monograph are those concerning: (1) the definition of evaluation, (2) decision-making, (3) values and criteria, (4) levels problems, and (5) the research model.

The purpose of this discussion is to examine each of these five problem areas: (1) as they are depicted in the PDK monograph, (2) as they appear as barriers to the application and installation of an evaluation system, and (3) in terms of selected caveats and suggestions for their recognition and avoidance.

EVALUATION'S PROBLEMS: AN OVERVIEW

Several chapters of the PDK monograph specifically address the five major problems of educational evaluation. Each will be very

---

*Stufflebeam, Daniel L., et al. Educational Evaluation and Decision-Making, 1971, Peacock Press, Bloomington, Indiana, 368 p.

briefly surveyed for the purpose of acquainting the reader with the general tone and dimensions of the treatment given them in the monograph.

The Problem of Definition

The definitional problem is much more than one of lexicographical confusion and ambiguity. Definitional problems are profound and far-reaching because as one defines evaluation so will he act and react.

Most important are these widespread definitions of evaluation which are inappropriate, inaccurate, or incomplete. The PDK monograph discusses three such definitions which cast evaluation as only: (1) measurement, (2) a means of determining congruence between performance and objectives, and (3) professional judgment. While each definition may have certain pragmatic advantages, each has unique disadvantages and all have the disadvantage of being incomplete. Evaluation may be forced by circumstances to be any one of these three some of the time, but it should never be just one of them all of the time.

Problems of Decision-Making

The problems of decision-making are integral to evaluation because evaluation as defined by the monograph's authors is the process of delineating, obtaining, and providing useful information for judging decision alternatives. Thus, evaluation has no meaning or utility apart from decision-making. The problems of decision-making center around the assumption (and premise of the monograph) that decisions ought to be rationally based on evaluative information. While the fact that they are often not data-based is itself a problem, other problems stem from our collective ignorance of decision-making phenomena; the

difficulties in linking evaluators and evaluative information with decision-makers; the proper role of the evaluator in maintaining a non-decision-making position while still influencing and altering the decision-makers' behavior; and the fact that decision-making settings are often obscure, situational, multiple, and incompatible.

Problem of Values and Criteria

Related very closely to the problems of decision-making are the problems concerning the values and criteria on which decisions are based. Their origin is often unknown or obscure; they shift from situation to situation and from decision-maker to decision-maker; they reflect the myriad and often contradictory values applied to education itself; in addition, the values of the evaluator himself may be inconsistent with those of the decision-maker.

The Problem of Levels

The so-called "levels" problem stems from the traditionally microscopic view held by educational evaluators in contrast to the increasingly macroscopic perspective of the educational decision-makers. Thus, because their units of observation, analysis, and interpretation are incompatible, the evaluator often uses inappropriate techniques. He attempts to "measure mountains with micrometers," which results in evaluative findings that are inconsistent with the information needs of decision-makers. The multiplicity of decision-making levels confound the problem of microscopic evaluations. The levels problem is also represented by the traditional tendency of evaluation to be concerned with determining the congruence between performance and standards while ignoring the examination of alternative

futures and strategies for their attainment (contingency evaluation). A final consequence of the level problem is the inappropriateness of data bases for serving macro-level, programmed decisions. The ad hoc, micro-level investigations simply do not provide the systematic and cumulative data base relevant to the information needs of contemporary educational decision-makers.

The Problem of the Research Model

A final ailment of educational evaluation discussed by the monograph's authors is the persistence of some to equate evaluation with research. The authors agree that to do so results in a variety of consequences which become barriers to sound evaluation. The dysfunctional consequences of equating evaluation with research (or of trying to "do" evaluation with only research assumptions, methods, and tools) include: working under unrealistic controlled conditions which are antithetical to the dynamism and complexity of many educational programs which must be evaluated; obtaining results only at the completion of the program (treatment); assuming or forcing comparability of subjects; and producing information which, even if relevant to decision-makers' information needs, restricts their interpretation to the dichotomous acceptance or rejection of hypotheses.

The reader is directed to the comprehensive treatment given to each of the five problems in the monograph. They have been briefly described here in order to provide a setting for discussing them in relation to attempts to apply and install an evaluation system in an educational R & D agency.

EVALUATION PROBLEMS AS BARRIERS TO THE INSTALLATION OF AN EVALUATION SYSTEM

The basis for much of the discussion to follow is, ironically, a series of subjective, experiential, and intuitive judgments about the barriers to the installation of a supposedly objective and rational evaluation system. However, the barriers take on a bit more credence in that they are based on the framework provided by the PDK monograph. Using the framework, each of the major problems will be discussed as it presents barriers to the application and installation of an evaluation capability in an educational setting. The barriers will be discussed in terms of symptoms, consequences, and possible guidelines for recognizing and circumventing them.

Problems of Definition

Quite possibly, the most important and subsuming of all barriers is the difficulty of arriving at an adequate definition of the evaluation system to be installed. It is helpful to look at the definition problems in terms of the appropriateness and consistency of the expectations held by the actors in an educational setting for the evaluation system which is to be implemented. These expectations can and will vary widely, yet it is essential that they be consistent.

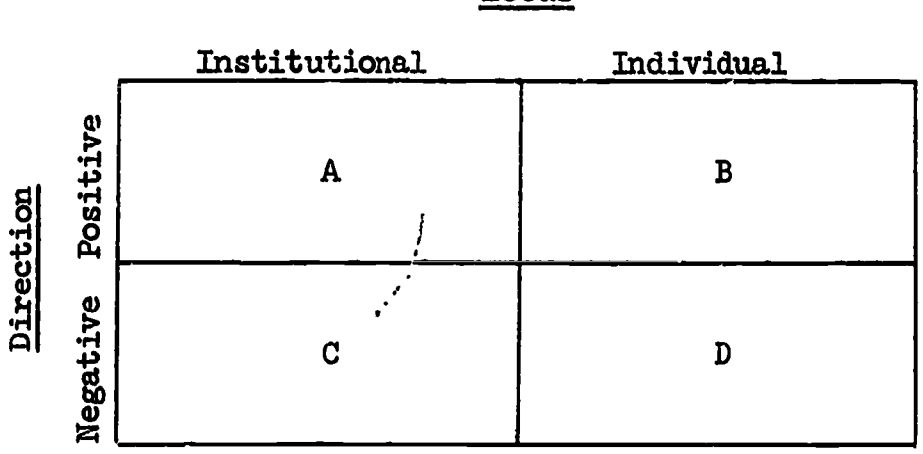
Foremost, the expectations must be consistent in terms of direction and locus. Figure 1 depicts the alternative expectations for these two dimensions.

Direction refers to the attitude held by an individual--simplistically expressed here as either a positive or negative attitude--toward the evaluation system. Locus refers to the question: "Positive or negative for whom?"--for the individual himself and/or for the institution in

which the evaluation system is to be implemented. Stated simply, individuals in an educational institution can define (hold expectations for) an evaluation system that are positive or negative in terms of the consequences which they think will accrue to themselves and/or the institution.

Figure 1

Directions and Locus of Expectations
Toward an Evaluation System



Using the cell designation from Figure 1, the logical combinations of definition sets for any given individual or group are: A and B; A and D; C and B; C and D. The optimum definition set is of course A and B. Each of the other three (if held in significant numbers by key individuals) become barriers to the installation of an evaluation system. For convenience (but with full awareness of and due caution to the semantic jumps being made) we may label the three dysfunctional definition sets: the "laggard" for A and D; the "opportunist" for C and B; and the "devil's advocate" for C and D.

To the "laggard," the evaluation system will be threatening because he sees the institution benefiting from it while he is penalized.

What is good for the institution is not good for the laggard. In fact, he is so labeled because his behavior is not consistent with the organizational goals and expectations. He has been "getting by" because (as he perceives it) evaluation in the past has not provided information about his organizational behavior to those decision-makers who, having such information, could penalize him in a variety of ways--from evoking negative sanctions to the withholding of positive ones.

The "opportunist," on the other hand, sees himself as benefiting from an evaluation capability but to the detriment of the institution housing them. If an evaluation system could in fact benefit the role functionaries in any organization but not the organization itself, then something is very wrong with either the organization or the evaluation system.

Whether the "devil's advocate" is a substantial barrier depends on how many there are and the underlying motivations for their definition sets. If their negativism is a relatively temporary expression of "let's wait and see" pragmatism (or "show me Missourism"), it becomes legitimately and feasibly incumbent upon the individual(s) responsible for installing the evaluation system to convince this type of devil's advocate of the system's merits. If, however, the motivation is one of deep-seated and dogmatic pessimism, then the numbers of such persons becomes a critical concern. The extent to which they are widespread and in key positions is directly inverse to the probability of the successful installation of an evaluation system.

The single most important component of an installed evaluation system (and therefore, potentially, the most critical barrier to its successful installation) is the collective attitude of the role

functionaries within the organization toward the evaluation system. Because this is true, it is critical that evaluation be defined appropriately and consistently from the outset. It is essential that the individuals responsible for installing the evaluation system be fully aware of the expectations (definitions) held for that system by individuals within the organization. It would seem to be entirely appropriate to construct, validate, and administer attitudinal scales which measure the direction and locus of an individual's expectations toward a proposed evaluation system. Then, if necessary, intensive dialogue through a variety of techniques (e.g. conferences, workshops, sensitivity groups, training courses, demonstrations, etc.) could be conducted to the extent necessary to assure the extinction and replacement of expectations held by the laggards, opportunists, devil's advocates, and others ("others" because the simple 2x2 classification ignored a variety of intermediate or neutral definition sets).

We have seen, then, that the problem of definition can be considered the most critical and subsuming of all evaluation problems that might act as barriers to the successful installation of an evaluation system. It has also been pointed out that a variety of definitions (expectations) can be held toward a proposed evaluation system in terms of direction and locus. Any patterned definition set other than "A and B" has been shown to be dysfunctional to the installation of a self-evaluation system. Further, the individuals responsible for installing an evaluation system have the responsibility for assessing the expectations toward that system and, as necessary, altering those dysfunctional expectations. Unfortunately, little attention is given to the assessment and altering of the definition sets held by individuals within the adopting organization.

This lack of attention is ironic because the evaluator, who should be familiar with elementary concepts of planned change, should know that any innovation will be successful only to the extent that it is perceived as favorable by those adopting the innovation.

Space will not permit an exploration of the other types of definitional barriers to the successful installation of an evaluation system. It is, however, important to briefly mention the variety of definitions that are often held toward an evaluation system. These might include defining evaluation as:

- A management information system (only).
- A panacea for all organizational ailments.
- A window dressing ritual to appease funding sources.
- A research methodology unit to function as a service for the organization's programs and projects.
- Just measurement, or judgment, or a post facto means of determining congruence between objectives and performance.

These and many other misdefinitions prevail. Their recognition and treatment should be very similar to the dysfunctional definition sets discussed above. Finally, the evaluator has the ethical responsibility to be certain that the prevailing definition of the system for which he has the installation responsibility is consistent with his: (1) expectations for that system, (2) ability and competence to install such a system, and (3) philosophy as to what ends should be served by an evaluation system.

Problem of Decision-Making

"Identify the decision-maker(s)"--is a necessary and reasonable initial step in any attempt to install an evaluation system. They must be identified because, as has been pointed out, evaluation has

no meaning or utility apart from its function of increasing the rationality of decisions. But the simplicity of this precept is misleading; it masks a host of very subtle but real barriers to the installation of an evaluation system. For example, the most common method for identifying decision-makers is likely to be the most inaccurate and hazardous--that of analyzing organizational charts depicting lines of formal authority and communication. It is inaccurate because of the fluidity of any formal organizational chart and, much more importantly, because of the inevitable discrepancy in any organization between the formal and informal lines of communication and authority. Thus, to identify the covert decision points is as important as the identification of formal decision points; it is incumbent on the evaluator to identify both.

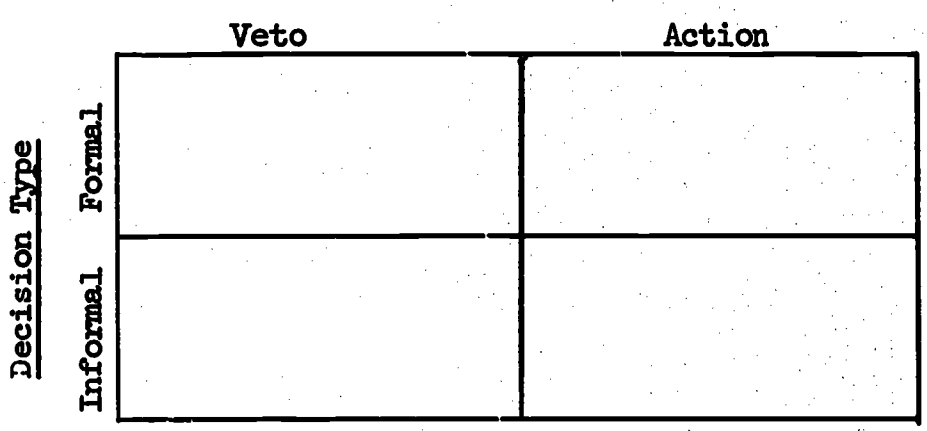
The almost paradoxical dilemma inherent in identifying all decision-makers lies in the fact that the informal decision points are, by definition, covert. Thus, they are at best difficult to identify. Even if they can be identified, they must then be incorporated in the design for installing the evaluation system; but by virtue of incorporating them in a design, they are no longer covert. An organization cannot formalize that which is informal.

A similar problem exists in differentiating the "action" decision-makers from the "veto" decision-makers. Action decision-makers are those having the authority (or more generically, "power") to commit human, monetary, and time resources to the continuation or change in the course of events. The veto decision-maker has the power to block the flow of resources which have been redirected by the action decision-maker. While it is realized that all decision-makers cannot be cast as pure types--either action or veto, the evaluator should be aware that some decision-makers operate primarily in one of these modes.

Figure 2 depicts a simple classification scheme for identifying (or at least recognizing the existence of) the formal and informal, action and veto decision-makers within an organization.

Figure 2

A Classification Scheme for Classifying and Identifying Decision-Makers



The utility of such a scheme may lie primarily in creating awareness of these different decision types. It is naive and defeating to assume and act as if only the "formal - action" decision-making mode exists.

In short, the problem of decision-making as it presents a barrier to the installation of an evaluation system centers on the difficulties in identifying decision-makers. Identification is difficult because: (1) the decision-making structure of an organization is fluid, (2) the structure always operates on both formal and informal levels and (3) there are different types of power (e.g. action or veto) which decision-makers exercise.

Problems of Values and Criteria

Only artificially separated from decision-making problems are those barriers which present themselves in the form of the multiplicity of values and criteria on which decisions are based. As there are overt and covert decision-makers so are there overt and covert criteria on which their decisions are based. Often, the covert criteria are weighed more heavily in decision-making than the overt. In addition, there exist a host of criteria which are unintentionally covert; that is, they are unknown to the decision-maker at the time that criteria are being elicited by the evaluator. The decision-maker would express them *if* he knew what they were. If at all, they are finally expressed only *after* the evaluative information is provided to the decision-maker; he then can recognize the inadequacies and omissions as he faces a choice among alternatives. In designing an evaluation system, the evaluator must delineate the information sources, methods for information retrieval, and techniques for providing information. It is imperative, then, that he be aware of the criteria on which the decisions will be made--particularly those programmed decisions which are serviced by information flowing from an established data base.

An additional question which the evaluator must answer is: "Whose criteria?" Assuming that the formal and informal, veto and action decision-makers have been identified and all possible covert and overt criteria have been elicited, the evaluator still must know *which* decision-makers will base *what* criteria on the information he is to provide. He will find often that the criteria and the associated costs of obtaining information to meet them vary greatly among the different decision-makers to be served by a single evaluation.

It is difficult and probably inappropriate for him to arbitrarily select which information will be obtained and provided for which decision-makers. Rather, he should simultaneously provide all (appropriate) decision-makers with information such as that represented in the simplified version presented in Figure 3. He should request that they resolve any incompatabilities and arrive at the questions to be answered (criteria are best phrased as questions to be answered by the evaluation).

Figure 3

Questions (Criteria)	Sources **DM_1	DM_2	DM_3	DM_4	Estimated Cost
* Q_1	X		X		$
Q_2		X			$
Q_3	X				$
Q_4				X	$

** DM: Decision-maker *Q: Question

At the least, this technique "protects" the evaluation from post-evaluation charges that the wrong or not enough questions were answered by the evaluation. More positively (and hopefully) the technique properly places the decision as to which decision-makers and which criteria will be served and at what cost with those decision-makers themselves. In addition, this technique makes it very clear to the decision-makers that (1) pertinent information is expensive, (2) it is ultimately their responsibility to determine what information is to be obtained and to whom it is to be provided, and (3) that an evaluation can have no meaning or utility apart from the criteria (questions) which they set forth.

The Levels Problem

Related closely to the problems of definitions, decision-making, and values and criteria is the so-called levels problem. Its origin as a problem for educational evaluation lies in the fact that traditionally the evaluator's unit of analysis is not compatible with the decision-maker's unit of interpretation. The incompatability is often based on the micro-perspective of the evaluator and the macro-perspective of the decision-maker. In regard to the barriers this problem presents for installing an evaluation system, the dilemma becomes one of assuring that the evaluative information is of sufficient time and content scope.

The evaluator must determine when information is needed and the content of the information. As for content, the evaluator and decision-makers must come to a clear agreement on the types of generalizations that the decision-maker must make. The decision-maker must state clearly (1) the population with which he is concerned, (2) the degree of confidence he must attach to evaluative findings, and (3) the necessary generalizability of the information. Each of these has clear implications for the sampling methodology, analysis techniques, and resources required by the evaluation. The time scope must also be specified by the decision-maker. If he needs certain information daily and other information semiannually, then these needs should be clearly communicated to (elicited by) the evaluator. This is particularly important for establishing an efficient cycle for retrieval from the data base which serves the regularized or programmed decisions. It is as inefficient and ineffective to burden the decision-maker with information too often or too soon as it is to retrieve and provide information too late.

An additional levels barrier is represented in the difficulty of maintaining an optimum "dross rate" in the data base. Ideally, this base would serve all programmed and ad hoc decisions. But the sporatic and unpredictable information needs of decision-makers cannot always be met by rapid retrieval from an extant data base. Thus, the evaluator will knowlingly store information which has little chance of ever being retrieved. He will do so because of the realization that sometime someone's ad hoc information need may be met by this information, but he also realizes that all ad hoc decisions cannot be met all of the time by any data base regardless of its scope. Thus, the balance to be maintained is one of determining the marginal utility of each additional information unit to be stored vis a vis the probabilities of it being eventually used to meet ad hoc decisions. The tendency to store all information often results in (literally) rooms full of information that has virtually no possible utility or retrievability.

A final levels problem in installing an evaluation system is in designing methods for obtaining "contingency" information. Contingency information might be thought of as information about alternative future states and desirable strategies for their attainment. The extent to which resources will be committed to obtaining such information is a function of the autonomy and philosophy of the organization. If its direction is, or must be, one of minimizing risks while maintaining the status quo, contingency information is relatively unimportant. However, such information is important if the organization is future-oriented, risk-taking, and has the freedom to maintain this position.

Whether contingency information is to be obtained is particularly important in a needs assessment type of context evaluation. It is one

thing to identify existing needs and quite another to anticipate future needs. Both the difficulties in obtaining and the apparently low priorities attached to contingency information are reflected in the virtual absence of solutions awaiting the emergence of educational problems.

It has been pointed out that the levels problem is a barrier to the installation of an evaluation system in terms of (1) the discrepancies between the scope of the evaluator's time and content units of obtaining information and the scope of the decision-maker's information needs, (2) the difficulties of maintaining a data base to serve programmed and ad hoc decisions, and (3) the need to indicate the extent to which contingency information is to be provided by the proposed evaluation system.

Problems of the Research Model

Finally, the barriers to installing an evaluation system that stem from the traditional view of evaluation as synonymous with research will be discussed. The dysfunctional consequences resulting from equating the two have been presented earlier. These same consequences are manifest if one attempts to apply an evaluation system to an educational organization which assumes that evaluation can be done with research tools. From the outset the decision-makers must recognize the unique and desirable differences between the two. It is the evaluator's responsibility to clarify and maintain the distinctions.

In short, the application of an evaluation system cannot be adequately designed using the research model as the guiding method for obtaining and providing information for decision-makers. It is but

coincidental when the decision-maker's information needs can be met using the research model; when they cannot, it becomes essential for the evaluator to have a repertoire of tools and techniques available for obtaining information to meet these needs.

SUMMARY AND CONCLUSIONS

Using the format of the 11th annual PDK monograph, Educational Evaluation and Decision-Making which examines five problems central to the theory and practice of educational evaluation, this paper has examined these problems as they appear as barriers in applying a "CIPP" type evaluation system to the ongoing operations of an educational organization. The five problems were those of (1) definitions, (2) decision-making, (3) values and criteria, (4) levels, and (5) the research model.

A Recap of Evaluation Problems

The definition problem embodies the collective expectations toward an evaluation system. These expectations determine whether benefits or penalties (are thought to) accrue to the individual and/or the institution as a consequence of an installed evaluation system. It is the evaluator's responsibility to identify the differential definition sets and to systematically extinguish and replace those which are dysfunctional to the success potential of a proposed evaluation system.

Decision-making problems require that the evaluator carefully identify the formal and informal decision-making patterns as well as the different types of decision-makers--especially the distinction between those having the power to redirect resources and those having the power to block the actions of other decision-makers.

As both definition and decision-making problems require that the evaluator carefully study the organizational context in which an evaluation system is to be installed, so do the barriers stemming from the problem of multiple, often contradictory, values and criteria on which decisions are based. It is imperative that the evaluator, before obtaining information, have clear directions from all decision-makers as to which criteria will guide the evaluation.

The levels problem centers on the difficulties involved in providing timely and pertinent information to decision-makers. If information is to be timely and pertinent, then its scope must be atuned to the time cycles and content units required by the decision-makers. The problem is particularly critical in designing a storage and retrieval format for a data base to serve both programmed and ad hoc decisions.

Finally, the problems of the research model were discussed as they present barriers to the installation of an evaluation system. It was suggested that it is inappropriate to use the research model as the paradigm for evaluation because evaluation, by our definition, has meaning only to the extent that it serves decisions.

Some Final Propositions

Finally, an attempt will be made to condense the foregoing discussion to a series of propositions which will serve as a capsule summary and a set of hypotheses to provoke additional investigation of the barriers to the installation and successful adoption of an evaluation system in an educational organization. Thus, while the statements which follow bring this discussion to a close, it is hoped that they will provide an additional impetus to expand our knowledge of educational evaluation and decision-making.

A. Definition

The probability of successful installation and adoption of an evaluation system:

1. Increases as key actors within the adoptive organization define (expect) the evaluation system:
   - 1.1 - to result in positive consequences for themselves and the organization.
   - 1.2 - consistent with the evaluator(s) expectations for and abilities to install the system.
2. Decreases as key actors within the adopting organization define (expect) the evaluation system:
   - 2.1 - to result in positive consequences for themselves or the organization.
   - 2.2 - to result in negative consequences for themselves and/or the organization.
   - 2.3 - to be only:
     - 2.3.1 - a means for evaluating according to professional judgment,
     - 2.3.2 - a system to measure individual performances,
     - 2.3.3 - a means for determining congruence between planned and actual means and ends, or
     - 2.3.4 - a means for meeting the evaluation mandates of external sources.
   - 2.4 - to be:
     - 2.4.1 - a panacea for the organization.
     - 2.4.2 - a complete management information system.
     - 2.4.3 - a service pool of methodological skills and tools.

B. Decision-Making

The probability of successful installations and adoption of an evaluation system:

1. Increases if the organization's decision-making patterns:

   1.1 - can be accurately identified by the evaluator.

   1.2 - are more formal than informal.

   1.3 - are more programmed than ad hoc.

   1.4 - operate more in an action than a veto mode.

   1.5 - are relatively stable across circumstances and over time.

2. Increases as the relationship between evaluator(s) and decision-maker(s) is:

   2.1 - characterized by mutual trust and respect.

   2.2 - one having unobscured communication channels.

   2.3 - one in which the evaluator is not also a decision-maker.

C. Values and Criteria

The probability of successful installation and adoption of an evaluation system:

1. Increases as an organization's decision-makers:

   1.1 - are able and willing to express the criteria on which their decisions will be based before evaluative information is obtained.

   1.2 - Agree on the values and criteria which are to guide an evaluation--especially those decisions which:

      1.2.1 - are programmed.

      1.2.2 - involve many decision-makers.

      1.2.3 - require substantial evaluation resources.

2. Decreases as the values and criteria on which decisions are based:

2.1 - are covert.

2.2 - are contradictory among different decision-makers.

2.3 - are primarily those of any agency or individual other than that organization's decision-makers.

D. Levels

The probability of successful installation and adoption of an evaluation system:

1. Increases as the evaluator's unit of analysis and the decision-makers unit of interpretation become similar.

2. Increases as the evaluative information provided to the decision-makers:

2.1 - is provided when they need it.

2.2 - allows them to generalize with the level of confidence necessary to select among alternatives.

2.3 - is retrievable from an in-house data base.

E. The Research Model

The probability of successful installation and adoption of an evaluation system:

1. Increases as both the evaluators and the decision-makers maintain the distinctions between research and evaluation.

2. Increases as the evaluator has skills and tools in addition to those required by the research model for meeting the information needs of decision-makers.