TOWARD GUIDELINES FOR CONDUCTING R&D ON THE GUIDANCE-SYSTEM IMPROVEMENT APPROACH

Marvin B. Mandell
and
Barry Bozeman,
with the assistance of
Steven Loveless

Submitted to the Development Project
Management Center, TAD/OICD/USDA.

May, 1983

INTRODUCTION

It is becoming increasingly clear that improvements in the management of development projects, programs and policies are essential for improving the quality of life in developing nations.[1] Indeed, some observers go as far as to suggest that "management is a major, if not the major constraint, to development performance" (USAID, 1981: 15). Unfortunately, traditional approaches to management improvement (e.g., short- and long-term training) have been disappointing in terms of their lasting effects on development management (USAID, 1981).

Largely in response to the urgency of improved development management and the shortcomings of traditional approaches to management improvement, a number of innovative technologies for management improvement have evolved in recent years. One such technology, which has been generated by the Development Project Management Center (DPMC),[2] is the "Guidance System Improvement Approach" (GSIA). The GSIA has been applied in a variety of country and sectoral settings and has shown considerable promise for improving development management.

For our purposes, there are two key characteristics of the GSIA.[3] First, the approach is flexible in two respects. It is a "loose-bundled" innovation; i.e., it consists of a core of components (activities) which adopters can "mix and match" to fit local needs (see Rice and Rogers, 1980; Johnston and Clark, 1982). Thus, the specific activities taken in one site will not

necessarily be the same as those taken in another site in implementing the GSIA. Also, within any one site, the GSIA is adaptive; actions taken during the course of a GSIA application are not necessarily those envisioned prior to the start of that application. In short, while the GSIA can be said to be (at least, potentially) a standard approach to improving development management in the sense of being based on an identifiable core of components and underlying principles, it cannot be said to be a standard <u>treatment</u> (as the term is used in formal experimentation) in the sense of being represented by a single detailed blueprint which is formulated in advance of the GSIA application, slavishly adhered to during that application and invariant across sites (see Johnston and Clark, 1982).

The second key characteristic of the GSIA is that implementation of the approach capitalizes on its divisibility (Zaltman et al., 1973) by relying upon an incremental, "partialization" mode of implementation (see Rothman et al., 1981, for a discussion of partialization). That is, initially an attempt is made to implement the GSIA in only a small segment of the intended overall target system; implementation in this small segment is then followed by an attempt to implement the GSIA in a somewhat larger segment of the intended overall target system; and so on, until it has been implemented in the entire intended overall target system.[4]

Because it is flexible, a highly specific description of the GSIA is neither feasible nor desirable. However, several charac-

teristics of the approach, in addition to the two specified above, can be identified. Among these are:

1. it is participative; i.e., site participants are given substantial responsibility for choosing and carrying out interventions;
2. it focuses on improvements in system performance as a means of developing "capacity," rather than directly on capacity;
3. it focuses on work-teams composed of individuals within the target organization, rather than isolated individuals;
4. it utilizes project management concepts in improving system performance;
5. it emphasizes the need for target system members to learn technical skills by "doing" within the context in which those skills are to be applied; thus, technical procedures are integrated with the group dynamics in which these procedures are to be utilized;
6. it focuses on introducing changes in the "guidance sub-system" of some unit (as distinct from the production, support or regulation sub-systems); the guidance sub-system "coordinates and directs...the other sub-systems" (PROPLAN/DPMC, 1983).

To date, little systematic research concerning the GSIA has been conducted. It is reasonable to believe, however, that, if properly conceived and executed, applied research can play a

significant role in increasing the effectiveness and efficiency of the GSIA. This report is aimed at providing a foundation for designing and conducting applied research which will serve such a role. To establish this foundation, we consider the ways in which applied research can contribute to the approach reaching its full potential and the forms such research should take in order to be of maximum effectiveness. In particular, an integrated applied research system which we label a "performance-sensing and learning user-support system" (PLUSS), will be described and discussed. As we envision it, a separate PLUSS would be developed for each site in which the GSIA is applied. As discussed further below, in addition to sharing common principles of design and operation, both substantive information and appropriate research instruments developed would be exchanged among the various systems developed.

Information provided by a PLUSS is expected to be of use to a variety of actors. These actors as a whole are referred to here as users. Within the overall set of users, two important subsets need to be distinguished, namely: those directly involved in implementing the GSIA within a site, referred to as <u>site participants</u>[5]; and those expected to have a role in future recurrent adoption decisions within the target system, referred to as <u>future target system adopters</u>. Other users include those called upon to support the GSIA in one form or another and those expected to have a role in adoption decisions in other target

systems.

The remainder of this report is divided into three main sections. In the first of these, some background concepts and terminology drawn primarily from the literature on innovation, evaluation and organizational effectiveness are presented. Then, an overview of the design and operation of a PLUSS which is appropriate for a GSIA application is presented. In the final section of this report, a set of issues involved in the design and operation of a PLUSS is discussed in greater depth.

BACKGROUND

INNOVATION: ADOPTION AND REINVENTION

An important distinction is drawn in the innovation literature between diffusion and organizational innovation. The former refers to the transmission of some innovation through the population of potential adopters (target systems). Organizational innovation, in contrast, is the process in which an innovation is selected and implemented in a given target system. While these two processes are intimately related, our primary focus here will be on the process of organizational innovation.

In light of the two key characteristics of the GSIA specified above ("loose-bundled" and reliance on "partialization"), certain characteristics of the organizational innovation process of this approach are particularly noteworthy. First, as with other loose-bundled innovations, "reinvention," i.e., the alteration of the innovation during the organizational implementation process (Rice and Rogers, 1980), is likely to play an especially prominent role in the adoption of the GSIA. Reinvention (also termed adaptation) is generally viewed as contributing significantly to the continued use of an innovation in that it results in both a better fit with local needs and conditions and greater commitment to the innovation on the part of those involved in its modification (see Glaser and Backer, 1977; Larsen and Agarwala-Rogers, 1977; Rice and Rogers, 1980; Berman et al., 1975). Calsyn and his associates (1977) provide a dissenting view, however. They contend that while

reinvention occurs quite frequently, it tends to seriously compromise the value of the innovation. Rice and Rogers (1980) differentiate planned from reactive reinventions. While they limit the latter to changes in response to specific problems occurring during implementation, we believe it is important to recognize that reactive reinvention can occur in response to unforeseen opportunities as well as in response to unforeseen problems.

A noteworthy implication of its reliance on "partialization" is that the organizational innovation process for the GSIA is a recurrent, rather than one-shot, process within a single target system. Of particular concern is that instead of a single adoption decision, adoption of the GSIA within a single target system involves multiple, recurrent decisions concerning continuation (retention) and internal diffusion (see Kimberly, 1981; Walton, 1975), each of which is critical in the sense that it can potentially prevent full adoption of the approach within that target system.

The foregoing discussion of the nature of the GSIA suggests three important functions that a PLUSS should serve. First, because the approach is flexible, rather than a standard, rigid "blueprint," information concerning how actual applications of the GSIA operate is likely to be valuable in helping users distill the approach. Secondly, the expected importance of reinvention in the adoption of GSIA suggests that information which can

facilitate this process should be given high priority in developing a PLUSS. Third, information concerning the effectiveness of the GSIA to date in the target system being considered should be given high priority. The importance of this function stems from the pivotal role that recurrent decisions concerning retention and internal diffusion play in the overall adoption of this approach in a given target system and the likely influence of perceptions of "effectiveness" in each of these recurrent decisions (see Baldridge and Burnham, 1975; Zaltman et al., 1973).

A common feature of these three functions is that they are aimed at facilitating components of the adoption of the GSIA which occur after the initial adoption decision. Moreover, they are accomplished by studying applications of the GSIA which have taken, or are taking, place in the target system of concern. In short, these functions suggest that in order to make a significant contribution to the adoption of the GSIA in a given target system, a PLUSS should concentrate on monitoring and evaluation, rather than ex ante analysis. However, as we shall discuss below, conventional approaches to monitoring and evaluation are not well-suited to the purposes of a PLUSS outlined above and the nature of the GSIA.

WEAKNESSES OF TRADITIONAL APPROACHES TO MONITORING AND EVALUATION

The major source of the deficiencies of conventional approaches to monitoring and evaluation, given the purposes and focus of a PLUSS, is that conventional approaches implicitly assume that:

1. a "blueprint-type" intervention is being considered; and
2. there are strong a priori assumptions (hypotheses) concerning cause-effect relationships, i.e., the types and magnitudes of the effects the intervention will have, if properly implemented, and the mechanisms which will cause these effects to occur (Casley and Lury, 1982; Cronbach et al., 1981; Johnston and Clark, 1982).

While certain production technologies (e.g., dams) might satisfy these assumptions, most managerial technologies, in general, and the GSIA, in particular, are almost certain to involve considerably less structure and more uncertainty than the above assumptions permit.

The major weakness of traditional approaches to monitoring, given the purposes and focus of a PLUSS, is the narrowness of such approaches. That is, traditional approaches to monitoring tend to concentrate primarily on whether the intervention is being implemented in a manner that is consistent with its design specifications. Any shortcomings-- i.e., gaps or errors--in the set of a priori assumptions concerning cause-effect relationships underlying the intervention as designed which would cause it to be ineffective are left to be answered through summative evaluation. Often, however, by the time the summative evaluation is completed, it is too late to revise the intervention appropriately. Finally, traditional forms of monitoring generally fail to pro-

vide information which might be useful for opportunity-sensing.

The major weakness o traditional approaches to evaluation, given the purposes and focus of a PLUSS, is that such approaches, for the most part, have an overly limited aim, namely "...establishing, with as much certainty as possible, whether or not [a well-defined] intervention is producing its intended effects" (Freeman, Rossi and Wright, 1979: 103). Thus, such approaches are of little use with respect to distilling the GSIA and reinvention. Moreover, even with respect to assessing the effectiveness of GSIA applications, traditional approaches to evaluation tend to be deficient (for our purposes) in the following respects

1. effectiveness is generally viewed in terms of a simple goal model that ignores a number of important complexities;
2. the provision of information is not necessarily consistent with organizational cycles (McConnel et al., 1982);
3. rigor in data collection and analysis is overly emphasized at the expense of feasibility and meaningfulness to users;
4. the importance of non-analytic inputs into assessing the effectiveness of an intervention is not recognized and the issue of how to integrate non-analytic results is not addressed.

PROCESS MODEL

At the heart of a PLUSS is a site-specific process model of how the GSIA application is expected to work. Such a model is simply a collection of three types of elements--activities, intermediate outcomes (precursors) and terminal outcomes (goals)[6] -- which are arranged into a time-ordered, causal sequence of relationships. Let us consider these three types of elements in some detail.

In general, an outcome is some change in the state or condition of some system. For our purposes, it is useful to distinguish between intermediate outcomes (precursors) and terminal outcomes (goals). A terminal outcome (goal) is a change in state that is intrinsically valuable. A terminal outcome might be intended or unintended. In most GSIA applications, goals will refer to development policy objectives such as productivity, employment, income generation and distribution, health status and literacy. In some cases goals might involve the minimization of some negative consequence such as environmental degradation.

An intermediate outcome (precursor) is a change in state which, if it occurs, affects the observer's confidence that some terminal outcome will occur[7]. That is, a precursor is of interest because it is an instrumental condition with respect to some goal. A precursor may be either external to the GSIA application in the sense of occurring independently of the GSIA application (i.e., it would have occurred even if the GSIA application had not been

undertaken) or internal to the GSIA application in the sense of being caused by the GSIA application. An important class of precursors in the context of GSIA applications is improvements in the "guidance sub-system" of some unit, e.g., achievement of agreement on project objectives or improved "analytical processes."

Certain features of precursors should be emphasized. First, a precursor is restricted to a change in condition in contrast to an activity. Second, a precursor must be expected to occur significantly prior (in time) to its associated goal. Third, a series of precursors may be associated with a single goal as illustrated in Figure 1A. That is, there might be multiple levels of precursors. Some precursors (e.g., precursor I.A in Figure 1A) would be expected to have rather indirect effects on goals, while other precursors (e.g., precursor I.C in Figure 1A) would be expected to be relatively directly related to goals. Similarly, a single precursor may be associated with more than one goal as illustrated in Figure 1B. It should also be noted that a precursor, as the term is used here, does <u>not</u> represent the presence or absence at the start of the GSIA application of a pre-existing facilitative condition necessary for the success of the approach (e.g., top-level support). However, a precursor can represent the change from absence to presence of such a facilitative condition during the life of the GSIA application.

Figure 1A

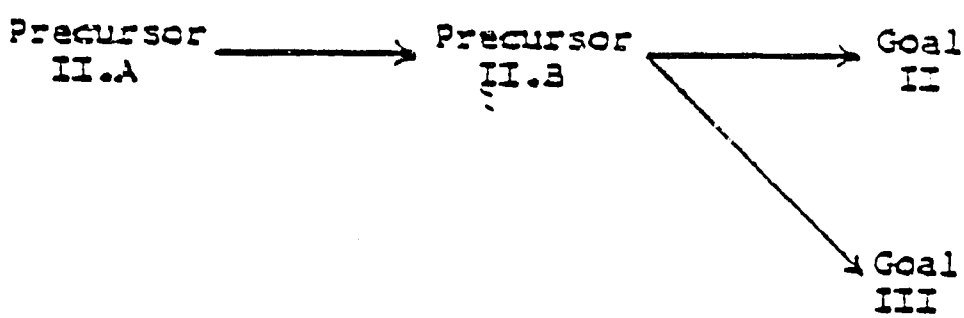

Figure 1B

The final element contained in a process model, <u>an activity</u>, represents some effort undertaken by relevant individuals (e.g., DPMC staff, cadre, field unit personnel) which is intended to change the state (condition) of some system, i.e., attain an intermediate or terminal outcome. Examples of activities might include preparing plans, developing analytic techniques and holding a meeting among certain actors.

As is discussed in detail below, assessing the effectiveness of a GSIA application is a particularly thorny task. However, as the discussion of effectiveness will show, knowledge of the extent to which relevant goals are achieved is at the core of assessing the effectiveness of a GSIA application. Thus, information concerning terminal outcomes has obvious relevance for assessing the effectiveness of a GSIA application.

Somewhat less obviously, information concerning precursors and activities also has important roles to play in assessing the effectiveness of a GSIA application. For one, in many cases a determination of whether a GSIA application results in the achievement of certain goals cannot be made until the latter stages, or even after the completion, of a particular GSIA application. That is, the GSIA application may set in motion a chain of events that eventually results in the achievement of certain goals; this chain, however, may unfold relatively slowly over time. Yet, various organizational cycles, most notably recurrent adoption decisions (see above), often require that the effectiveness of a GSIA

application be assessed at earlier points in time. In such situations, information concerning the attainment of intermediate outcomes can often be relied upon as a surrogate in assessing the effectiveness of a GSIA application. In this sense information concerning precursors can be thought of as a "leading indicator" of the effectiveness (in terms of goals) of a GSIA application. It must be recognized, however, that there may be considerable slippage between precursors and their associated terminal outcomes, especially in cases where the intermediate outcome considered is an early one in a long chain of precursors leading to the particular goal. The key to deciding whether a precursor should be relied upon as a surrogate for a goal in assessing the effectiveness of a GSIA application at a given point in time and, if so, which precursor (if a chain of precursors is believed to lead to the goal in question) is specifying as accurately as possible the time required to achieve the various precursors and goals contained in a process model. As a rule of thumb, if it is realistic to expect that a goal will be achieved by the time the assessment of the effectiveness of a GSIA application will take place[8], then some direct measure of the attainment of that goal should be used in that assessment. However, if achievement of a goal is not expected to occur by the time the assessment of the effectiveness of a GSIA application will take place, then a precursor that <u>is</u> expected to occur by that time should be used as a surrogate for that goal. Of those

precursors which are expected to occur by the point in time at which the assessment of the effectiveness of the GSIA will take place, the one that is closest (in time) in the presumed causal chain to the goal in question should be chosen as the surrogate for that goal.

Information about precursors, along with information on activities, can play another important role in assessments of the effectiveness of a GSIA application. Strictly speaking, to show that relevant goals are achieved is not the same as demonstrating that a GSIA application has been effective. Rather, it must also be demonstrated that achievement of the goals was due to the GSIA application, i.e., that the goals would not have been achieved in the absence of the GSIA. Traditionally, this is accomplished by also observing whether the goals are achieved in a control group in which the innovation of concern (in this case, the GSIA) is not implemented. This tactic, however, is generally not feasible in the current context. A second tactic for determining whether the goals would have been achieved in the absence of the GSIA is to develop and apply a simulation model of the system of concern (see Larkey, 1977). Again, this is not viewed as feasible in the current context. Constructing a plausible causal chain of events which leads from activities to goals and providing as much empirical validation of this chain as possible (through information on activities and intermediate outcomes) is a third way to cope with the issue of whether the GSIA application is the cause of relevant goals being achieved.

That is, to the extent that one can coherently account for goal achievement in terms of activities and intermediate outcomes that are associated with the GSIA application, rival explanations of goal achievement become less credible.[9] (This logic can be extended to attempts to attribute the occurrence of intermediate outcomes to the GSIA application when those intermediate outcomes are used as surrogates for terminal outcomes as discussed above. In such cases, one would attempt to account for achievement of the intermediate outcome serving as a surrogate for a terminal outcome in terms of earlier intermediate outcomes and activities.)

Another important role of information concerning precursors involves revising the set of goals which are viewed as appropriate. In particular, it is being increasingly recognized that "people decide what they want [i.e., define their goals] only in light of what they think they can get" (Johnston and Clark, 1982: 18). Moreover, information on the occurrence or non-occurrence of a precursor, by definition, affects one's perception of the likelihood that some goal will be achieved. Therefore, information concerning precursors is a potentially important aid for identifying appropriate goals. In particular, information indicating the non-occurrence of a precursor that was anticipated to occur might suggest that the goal corresponding to this precursor is unrealistic and should be scaled down or perhaps even eliminated. Conversely, indications that a

previously unanticipated precursor has occurred might suggest that additional goals be considered.

Information on intermediate outcomes can also play an important role in facilitating the process of reinvention. Assessments of the extent to which some precursor is attained provide rapid and timely feedback to site participants concerning the efficacy of their activities. Indications that an important precursor has not been attained by the time it had been expected are important signals of the need for reinvention in time to make the GSIA application succeed. Conversely, the occurrence of a previously unanticipated "positive" precursor, if detected, represents an opportunity which might trigger reactive reinventions. Moreover, simply calling attention to precursors is likely to instigate a search for alternative--and perhaps better--paths to a given goal. Explicitly considering previously overlooked paths is likely to facilitate planned reinventions.

In addition to playing an important role in assessing whether the achievement of certain outcomes was due to the GSIA application, information concerning the activities involved in a GSIA application has obvious importance in helping users to distill the approach.

In summary, then, the primary role of information concerning terminal outcomes is to aid in assessing the effectiveness of a GSIA application. Information on precursors has four primary

roles, namely:

1. it can serve as a surrogate for information on terminal outcomes, which is particularly important in cases where information concerning goals cannot be obtained in a timely fashion;
2. it can help to determine whether the occurrence of a subsequent intermediate outcome or a terminal outcome was caused by the GSIA application;
3. it can facilitate learning about appropriate goals; and
4. it can facilitate the process of reinvention.

Finally, information on activities can help potential users both to distill the GSIA and to determine whether the achievement of certain outcomes was due to the GSIA application.

COMPLEMENTARY SOURCES OF KNOWLEDGE CONCERNING ACTIVITIES, PRECURSORS AND GOALS

It is important to stress that knowledge about the activities, precursors and terminal outcomes associated with a GSIA application can be obtained from sources other than the PLUSS developed. In particular, "ordinary knowledge" processes and "social learning" (Lindblom and Cohen, 1979) as well as standard reporting systems are important mechanisms for developing such knowledge. Thus, a PLUSS should complement and supplement, rather than displace, these other mechanisms. In operational terms, this means that a PLUSS should be selective, rather than

comprehensive, in terms of the activities, precursors and terminal outcomes it focuses upon. That is, although a large number of activities, precursors and goals are relevant in any GSIA application, the PLUSS need not and should not attempt to provide information about all of them. Moreover, these alternative knowledge sources are important inputs into the development of a PLUSS (see Figure 2).

CONCEPTUALIZING THE EFFECTIVENESS OF A GSIA APPLICATION

In our view, the effectiveness of a GSIA application should be conceptualized in terms of its contributions to the performance of some significant real-world system, such as an educational system, an agricultural system or a health system. Recent work in evaluation and organizational effectiveness, however, suggests that basing assessments of effectiveness on this type of conceptualization is likely to make such assessments considerably more complex and less conclusive than assessments of effectiveness are often claimed to be. In particular, a key implication of this work is that it is unrealistic to expect to be able to obtain a single, universal assessment of the effectiveness of a GSIA application.

The first task which must be accomplished if the above approach to assessing the effectiveness of a GSIA application is to be operationalized is the specification of the boundaries of the "system of concern." This system may be a project, a set of projects, an organization, a set of organizations or a

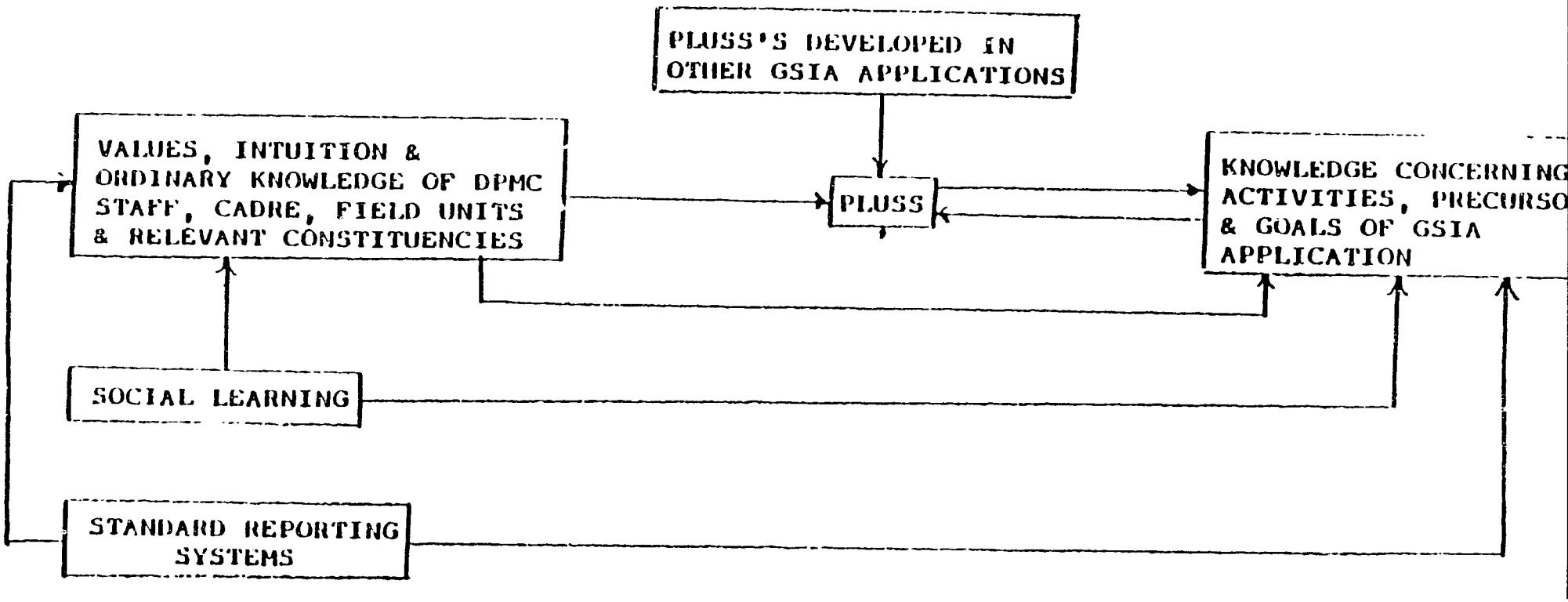
PLUSS'S DEVELOPED IN OTHER GSIA APPLICATIONS
VALUES, INTUITION & ORDINARY KNOWLEDGE OF DPMC STAFF, CADRE, FIELD UNITS & RELEVANT CONSTITUENCIES
PLUSS
KNOWLEDGE CONCERNING ACTIVITIES, PRECURSO & GOALS OF GSIA APPLICATION
SOCIAL LEARNING
STANDARD REPORTING SYSTEMS

Figure 2

part of an organization. Specifying the boundaries of the "system of concern" is analogous to the consideration of externalities in cost-benefit analysis. The question here is how far to extend the boundaries of the system beyond direct, first-order effects. In many cases researchers limit their focus to easily identifiable first-order effects (e.g., organization change within a single organization), but in many instances it is necessary to consider secondary and tertiary effects, especially when there is reason to believe that the GSIA application has potential for significant effects on units beyond the immediate site in which the GSIA is being implemented. The difficulty here is in determining just how far to trace indirect effects and what magnitude of indirect effect is likely to be sufficiently consequential to warrant attention. The appropriate boundaries of the "system of concern" depend in large part on one's purposes (see Katz and Kahn, 1978). Since different actors are likely to have different purposes, no single set of system boundaries is likely to be appropriate.

Even thornier than the specification of the boundaries of the system of concern is the problem of determining whether a given system is effective. Hall (1982) proposes that organizational effectiveness[10] be viewed in terms of a "contradiction model." This model is an attempt to encompass the insights provided by previous models of organizational effectiveness,

such as the goal model, the systems-resource model, participant-satisfaction models and social-function models. In particular, key features of Hall's "contradiction model" are as follows:

1. Organizations have multiple and conflicting constituencies.

   In Hall's usage, a constituency refers to a group of people affected by an organization. A constituency need not be organized or even aware that they are affected by an organization. Moreover, a constituency can be either internal or external to the organization.

2. Organizations have multiple and conflicting goals.

   This multiplicity exists on three levels. First, a single constituency at a given point in time is likely to hold multiple and conflicting goals for the organization. Second, the goal set of a single constituency may shift over time. Third, different constituencies are likely to hold different, though overlapping, goal sets. Thus, if effectiveness is viewed as the attainment of most or all of an organization's goals, an organization cannot be effective, since certain goals can only be attained at the expense of the accomplishment of other goals.

3. Organizations have multiple and conflicting time frames.

   In many instances an action that maximizes the attainment of a certain goal in the short-run will not maximize the attainment of that goal in the long-run, and vice-versa. Different time frames are likely to be important to diff-

erent constituencies. For example, Hannan and Freeman (1977) note that higher levels of an organization often take a longer time perspective than lower levels.

These three features imply that "no organization is effective. Instead organizations can be viewed as effective (or ineffective) to some degree in terms of specific...goals, constituents, and time frames" (Hall, 1982: 302; emphasis in original).

Many of the complexities and contradictions discussed above, it should be noted, are justifications for employing the GSIA. That is, the GSIA can be seen as primarily a means for managing just these kinds of problems. In particular, by relying on project management concepts, the GSIA specifies boundaries of the "system of concern," limits constituencies and identifies a small set of goals which are both agreeable to this limited set of constituencies and mutually compatible. In short, the GSIA artificially reduces the system of concern into a manageable system. While this tactic of reduction is appropriate and necessary for intervening into that system, it is not likely to be appropriate for understanding and assessing the effects of such interventions. In particular, applied research on the effects of a GSIA application ought to aid a variety of constituencies (including, but not limited to, site participants) in judging the effectiveness of that application. The goals, system boundaries and time frames of concern to these constituencies are likely to include, but not be limited to, those which guided the site participants during the GSIA

application. In short, applied research on the effects of a GSIA application cannot avoid the types of complexities and contradictions discussed above. Indeed, some would go as far as to suggest that a (perhaps <u>the</u>) central purpose of evaluation and, more generally, policy analysis is to highlight and provide insight into these complexities and contradictions (see Enthoven, 1975 and Etzioni, 1978).

OVERVIEW OF THE DESIGN AND OPERATION OF A PLUSS

We anticipate that the specific PLUSS which is most appropriate for a given site will not be exactly the same as the PLUSS which is most appropriate for some other site. However, some basic principles for designing and operating a PLUSS can be identified. These principles can be summarized in terms of the process depicted in Figure 3. The first three steps in this process are intended to produce one or more tentative site-specific process models of how the GSIA is expected to work. The next step identifies the information which the PLUSS will provide at various points in time. The final step is the actual collection, processing, interpreting and reporting of the data.

The process represented in Figure 3 is iterative. That is, after proceeding through the process initially, it is necessary to return to certain points in the process from time to time and revise the PLUSS. However, not all steps should necessarily be repeated during all iterations. For example, it might be desirable to return to steps 3 and 4 on a quarterly basis, to step 2 on a semi-annual basis and to step 1 on an annual basis.

Thus, the PLUSS developed in a particular site will constantly evolve in two respects. First, the set of activities, precursors and goals included in the site-specific process model of the GSIA is likely to be revised periodically in response to the identification of previously unforeseen precur-

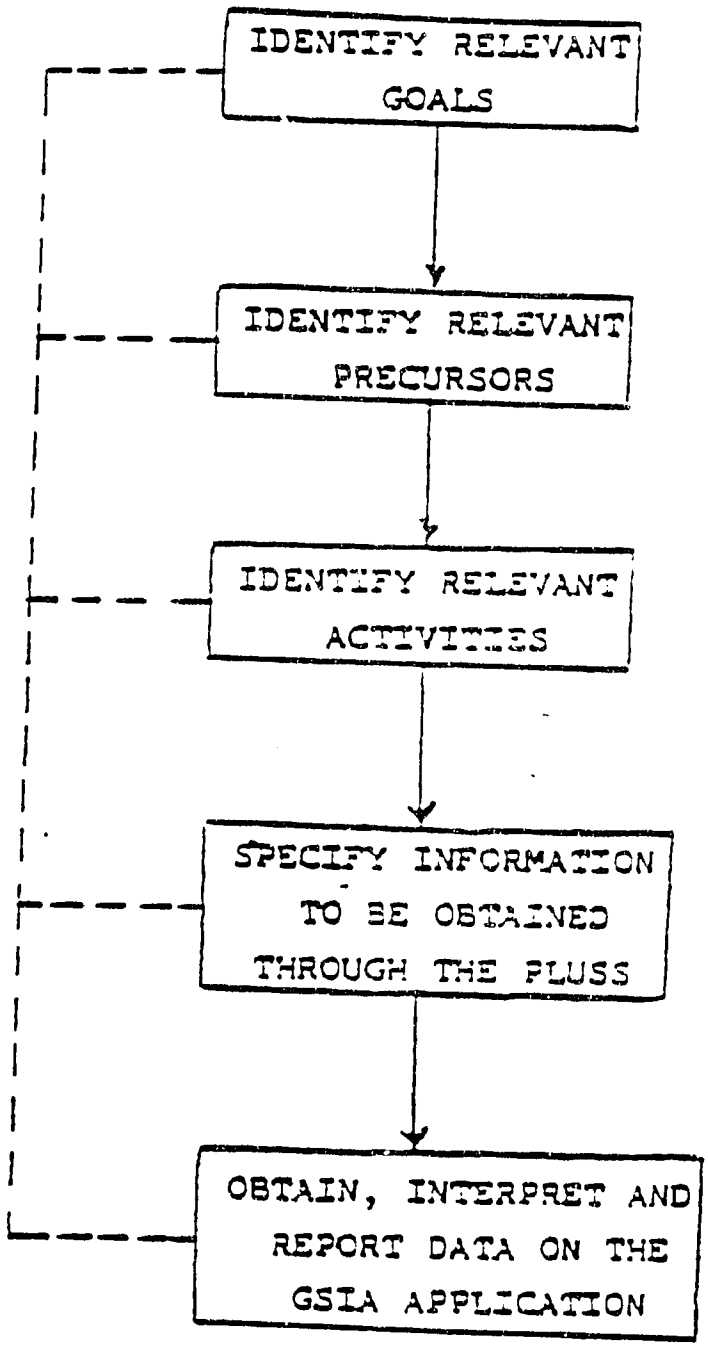
IDENTIFY RELEVANT GOALS
IDENTIFY RELEVANT PRECURSORS
IDENTIFY RELEVANT ACTIVITIES
SPECIFY INFORMATION TO BE OBTAINED THROUGH THE PLUSS
OBTAIN, INTERPRET AND REPORT DATA ON THE GSIA APPLICATION

Figure 3

sors (opportunities), the detection of unachieved precursors, shifts in goals and changes in the "action plan." The information produced by the PLUSS is, of course, expected to play a major role in these revisions, but other forces are likely to play a role as well. Second, even if the process model remained fixed, it might be determined, say, that during the early stages of the GSIA application, the PLUSS should collect information regarding certain activities and precursors, while during later stages a different set of activities should be focused on and efforts to collect information concerning some precursors should be replaced by efforts to collect information regarding selected goals.

A central principle of a PLUSS is that by involving various users (especially site participants and potential target system adopters) at various points in this process, the utility of a PLUSS can be greatly enhanced. Nonetheless, certain obstacles and drawbacks of user involvement should be kept in mind. These obstacles and drawbacks, along with some means of avoiding them or overcoming them, will be discussed in the next section of this report.

IDENTIFYING TERMINAL OUTCOMES

The set of terminal outcomes identified will ideally include

all goals that various constituencies believe to be relevant. The contradiction model of organizational (system) effectiveness discussed above suggests that the list of such goals will be large and diverse. Stakeholder surveys (Lawrence and Cook, 1982) are likely to be a useful tool for identifying relevant goals. The first step in such a survey is to develop a list of stakeholders. A "snowball" approach is likely to be appropriate here. That is, an initial list of stakeholders, perhaps consisting of those with a (known) formal connection to the system of concern and those identified by DPMC staff, cadre and other site participants as being relevant stakeholders, would be generated. Each member of this initial list would then be asked to identify other actors whom they perceive to be stakeholders in the GSIA application. These "second-round" nominees would then be asked, in turn, to identify others whom they perceive to also be stakeholders, and so on. To keep the list of stakeholders manageable, at one or more points in the "snowballing" process, the list of stakeholders generated might be pared down. This might be accomplished by having those identified to date as stakeholders (or a subset of those individuals) rate the "importance" vis-à-vis the particular GSIA application being considered of the other individuals on the list. Alternatively, only those who are identified by at least some minimum number of others as being a

stakeholder might be retained on the list.

After having established a list of stakeholders, the goals which each stakeholder believes to be important components of the effectiveness of the system of concern must be identified. The most straightforward approach to use here would be a simple interview or written questionnaire. A modified Delphi process is a potentially worthwhile variant which should be considered as well. That is, after the initial set of questionnaires/interviews are completed, a feedback report consisting of an edited list of all goals specified would be distributed to each respondent; the respondents might then be given the opportunity to add goals to the list. They might, in addition, be asked to rate (say on a five-point scale) the importance of each goal to them. Other suggestions on identifying stakeholders and their goals are contained in Goodman and Pennings (1980).

IDENTIFYING RELEVANT PRECURSORS

For each goal identified, one or more plausible chains of precursors leading to that goal should be constructed. In addition to the linkages between precursors and goals, linkages among precursors need to be specified. Consensus on each chain is not necessary at this point; indeed, in many respects it is preferable that alternative plausible chains be identified. In particular, identifying alternative plausible chains

of precursors, which is directly analogous to "theoretical triangulation" (Denzin, 1970), can reduce the likelihood of overlooking important precursors, i.e., failing to detect significant opportunities and problems. For one, a PLUSS cannot be "programmed" to collect information concerning the occurrence/non-occurrence of a precursor unless it is explicitly identified. This is not to say that the occurrence/non-occurrence of a precursor can only be sensed if a PLUSS is "programmed" to scan for that event; alternative sources of information (most notably, informal serendipitous observation) might also enable site participants to detect whether that precursor occurs. The point is that the likelihood of accurately detecting the occurrence/non-occurrence of a precursor is increased if a PLUSS is "programmed" to scan for that event. Moreover, the power of serendipitous observation to accurately detect whether a precursor occurs is increased when a possible precursor is a priori explicitly identified, since explicitly identifying a precursor is likely to make it more salient to observers.

Various forms of structured group processes, e.g., nominal group technique and "brainstorming," may be useful for this task. Nominal group technique (Delbecq et al., 1975) is a structured group process intended to result in balanced participation among group members. In the present context it would proceed somewhat as follows. Individuals are brought

together in a conference room. At the beginning of the meeting they do not discuss the GSIA application with each other. Instead, they are asked to write ideas they might have concerning chains of precursors leading to the relevant goals. At the end of a short period of time (say ten minutes), each individual, in round-robin fashion, presents one chain of precursors from his or her list. A recorder writes that chain on a flip-chart that is in view of all participants. Round-robin recording on the flip-chart continues until all private lists are exhausted. During this recording process, there is no discussion of the specified chains of precursors. At the end of the recording process, the plausibility of each chain of precursors is discussed. Additional time might then be devoted to individual, silent generation and revision of chains of precursors. Round-robin recording of these new and revised chains of precursors would then occur, again followed by discussion of the plausibility of each chain. Any number of such three-part cycles (generation, recording, discussion) may, of course, occur. It should be noted that consensus concerning the plausibility of the various chains of precursors should <u>not</u> be sought.

A familiar process for generating ideas--in this case, chains of precursors leading to a goal--brainstorming is really nothing more than an effort to stimulate creative thinking by

explicitly disavowing any censorship of ideas (Osborn, 1957). Brainstorming can be conducted either in groups or by individuals who then pool their ideas. There is no clear relationship between group size or group interaction patterns and the quantity or quality of ideas generated (Lewis et al., 1975). Porter and associates (1980: 7) provide the following guidelines for brainstorming:

1. Criticism of any sort is barred, both of one's own ideas and those of others. "Killer phrases," such as "that's ridiculous," are avoided.
2. Quantity of ideas is the primary objective.
3. Unusual, remote or wild ideas are sought; free-wheeling is welcome.
4. Combinations, modifications and improvements on previously expressed ideas are encouraged. It is therefore important that every idea offered be kept visible to the participants to generate additional ideas.

If brainstorming is to be used in this context, it must be <u>followed</u> by an assessment of the plausibility of the various chains of precursors generated.

Whether or not some sort of formally structured group process is used to construct plausible chains of precursors for a given GSIA application, it is likely that this task can benefit from making participants in this task aware of

plausible chains of precursors constructed in connection with other GSIA applications. In some cases, chains of precursors developed in connection with other GSIA applications will be transferable to the GSIA application at hand with only minor modifications. In other cases, even if a direct transfer is inappropriate, awareness of chains of precursors constructed in connection with other GSIA applications may provide useful "sparks" to participants in this task, even if the relevant goals are somewhat different. On the other hand, however, making participants in this task aware of chains of precursors developed in connection with other GSIA applications might result in barriers to their creativity by creating a "focus" ("tunnel vision") effect. In the present context, this effect refers to participants restricting their efforts to attempt to identify plausible chains of precursors for the GSIA application at hand to those which are directly descended from one which has been developed in connection with some previous GSIA application, even if a radically different chain might be more appropriate for the GSIA application at hand (see van de Ven and Delbecq, 1971). To obtain the possible advantages of making participants in the task of generating plausible chains of precursors for a given GSIA application aware of chains of precursors which have been developed in connection with previous GSIA applications while minimizing the potential dangers of

doing so, participants in this task should be made aware of chains of precursors which have been constructed in connection with other GSIA applications somewhere in the middle of the process (whether it is formal or informal) of generating plausible chains of precursors for the GSIA application at hand.

DEFINING ACTIVITIES

In step 3 the tentative site-specific process model(s) of the GSIA application is completed by identifying relevant activities included in the GSIA application and their linkages to various precursors and goals. A key issue here is whether or not to attempt to identify an exhaustive list of activities included in the GSIA application. Experience in studies of managerial job behavior, which we believe to be directly analogous to the task of determing activities included in a GSIA application, suggests that an exhaustive approach is weak in terms of its ability to distinguish important aspects of managerial behavior (Campbell et al., 1970). Thus, it appears that, in light of the need to help users distill the GSIA, a more purposive, selective approach is more appropriate.

One possibly appropriate technique, which has been used successfully in studies of managerial behavior to pinpoint the more important elements of job behavior, is the Critical

Incidents Method (Flanagan, 1954; Campbell et al., 1970). The initial step in this method is to identify a large number of "critical incidents," which in the current context are activities crucial for achieving a certain outcome (goal or precursor). Traditionally, critical incidents are identified from retrospective behavioral accounts obtained from those close to the system being studied (e.g., managers). In the current context, the body of relevant past experience available at the time the site-specific process model is formulated is not likely to be large enough to allow a sufficiently large number of critical incidents to be identified. Thus, for a PLUSS it will be necessary to replace or at least supplement lists of "critical incidents" obtained through retrospective behavioral accounts with "critical incidents" identified by having site participants specify a priori the activities that they believe will be especially critical for the achievement of outcomes (precursors and goals) specified. Again, "brainstorming" and nominal group technique might be useful tools for this task. After a large number of critical incidents are identified, they are abstracted to form categories of activities which are represented in the process model and also form potential foci of PLUSS-based observations.

SPECIFYING INFORMATION TO BE OBTAINED THROUGH THE PLUSS

The next step shown in Figure 3 entails the determination of the information to be obtained through the PLUSS. This step consists of three related components, namely:

a. determining which goals, precursors and categories of activities the PLUSS will provide information about;

b. specifying the time limits for collecting information on each of these elements;

c. specifying the data collection technique(s) to be used to obtain information about each of these elements (e.g., survey, interview, direct observation; see "Measurement Techniques for a PLUSS," below).

The previous three steps are likely to result in a long and diverse list of goals, precursors and categories of activities about which a PLUSS <u>might</u> be called upon to provide information. However, it will rarely be feasible to rely upon the PLUSS to provide information regarding all of these elements. Fortunately, the existence of information sources other than the PLUSS to provide relevant information generally eliminates the necessity of relying upon the PLUSS to provide information regarding all of the elements included in the process model(s) specified. Thus, it will be necessary to screen the list of goals, precursors and categories of activities

which exists at the end of step 3. In doing so, a number of criteria might be taken into consideration, including:

1. the perceived importance of the element;
2. the extent to which applied research can provide credible information concerning that element and the cost of doing so (see "Errors in Performance-Sensing and Learning" and "Measurement Techniques for a PLUSS," below); and
3. the extent to which alternative sources of information concerning that element are available, and the credibility of those alternative sources.

Researchers and site participants should generally collaborate in performing this task.

It is not necessary for any given goal, all of its associated precursors and all related activity categories to be either included in or excluded from the PLUSS as a class. That is, even if no information concerning a given goal is sought through the PLUSS, information concerning some precursors and/or activity categories associated with that goal might still be included in the PLUSS and vice-versa.

Specifying time limits for collecting information on each element selected for observation through the PLUSS involves more than simply deciding on start-up and termination points. It is also necessary to decide how frequently, and at what times, information concerning a given element is to be obtained.

Of particular concern in this regard is that, since both precursors and goals are defined in terms of changes of state, for each outcome (goal or precursor) considered in the PLUSS, a choice must be made as to whether the PLUSS is to be relied upon to provide information concerning the attribute of the system associated with the outcome prior to the change in state (prior state), after the change in state (ex-post state) or both. Observations of some attribuce of the prior state of a system are often referred to as "baseline" measurements. Our notion of a PLUSS contains important implications concerning baseline measurements which may run counter to conventional wisdom. First, "comprehensive" baseline measurements (i.e., a baseline measurement associated with each and every goal and precursor) are not necessary. Rather, since we see a PLUSS as focusing on only a subset of all activities, precursors and goals of concern, at most baseline measurements for the focal precursors and goals are needed. Moreover, for some focal outcomes, sufficient information concerning the prior state of that aspect of the system may be available from sources other than the PLUSS, thereby eliminating the need for baseline measurements corresponding to that outcome. Finally, it is not necessary for all baseline measurements to be taken at the initiation of a GSIA application nor must all baseline measurements be taken at the same point in time. Rather, it

is only necessary that the baseline observation of the attribute of the system associated with a given outcome be taken prior to the occurrence of the event of interest. The time each event is expected to occur should be specified in the process model of the GSIA application. Indeed, the later the baseline measurement is taken, the lower the likelihood of extraneous events (history, maturation; see Freeman, Rossi and Wright, 1979) confounding assessments of the impacts of the GSIA.

It is also important to emphasize that users (even within a single subset specified above) are likely to have different interests and, hence, place different values on various types of information. This highlights the importance of "stakeholder analysis" (Lawrence and Cook, 1982; Mitroff et al., 1979) within this step. The issue here is the determination of interested parties and, in turn, in most instances, assessments of their interests and information needs. A PLUSS, as opposed to basic research, is intended to have instrumental value, but instrumental to whom? A variety of approaches to stakeholder analysis have been suggested by Mitroff et al. (1979) and by Goodman and Pennings (1980).

OBTAINING, INTERPRETING AND REPORTING DATA ON THE GSIA APPLICATION

The final step in the process sketched in Figure 3 is the

actual production of PLUSS outputs (information). It is important to emphasize that non-PLUSS inputs, such as "ordinary knowledge" (Lindblom and Cohen, 1979) and values, play a key role in this step in two respects. First, in unstructured, complex situations, such as those in which a PLUSS will be employed, applied research can only provide insight into relevant concerns; it cannot yield solutions or "self-evident prescriptions" (Elmore, 1975; also see Blumstein, 1979 and Brill, 1979). That is, PLUSS-based data cannot be interpreted in a straightforward, technical manner since, as Rieker (1980: 225; emphasis in original) noted, "the data never point to one and only one course of action, recommendation, or conclusion." Rather, a number of alternative implications may follow from a given set of PLUSS-based data. In making the leap from such data to valid "action recommendations," such elements as values, other forms of knowledge, assumptions one is willing to make and risks one is willing to take necessarily play a vital role (Rieker, 1980).

Second, as noted above, insights into the occurrence of relevant outcomes (goals and precursors) and activities included in a GSIA application can be obtained from information sources other than the applied research component of the PLUSS. These alternative sources include casual and serendipitous observation and routine reporting systems. Where the

insights obtained from these alternative sources are especially notable--e.g., because they are surprising or indicate the need to revise the process model formulated (which may, in turn, indicate the need or opportunity for reinvention or suggest new foci of data collection)--they should be incorporated in PLUSS reports. However, selectivity in terms of reporting non-PLUSS-based insights in order to avoid information overload should be stressed.

A second point to be emphasized here is that the reporting task encompasses the greatest opportunity to influence the credibility of PLUSS outputs. In particular, as is discussed further in "Determinants of the Credibility of PLUSS Outputs," below, characteristics of the sender of a report and the report's medium are two major elements which are believed to affect the credibility of applied research, in general, and PLUSS outputs, in particular.

KEY ISSUES IN THE DESIGN AND OPERATION OF A PLUSS

USER-INVOLVEMENT IN THE DESIGN AND OPERATION OF A PLUSS

Rationale For User-Involvement

Extensive involvement of users (especially site participants) in the design and operation of a PLUSS is critical to its success. This proposition is based on the importance of user-involvement in applied research, in general (see Conway et al., 1976; Johnson, 1980; Lawrence and Cook, 1982; Pondy, 1975, Reisman and deKluyver, 1975; Rich, 1975; U.S. General Accounting Office, 1977), coupled with unique aspects of the GSIA and the purposes a PLUSS is intended to serve. Specifically, user-involvement provides users with a sense of ownership of the information obtained from a PLUSS, thereby increasing the likelihood that such information will be used. Moreover, user-involvement in the design and operation of a PLUSS is consistent with an orientation to management improvement, such as that embodied in the GSIA, which stresses collaboration and involvement rather than hierarchical control (see Pasmore and Friedlander, 1982).

The unstructured and complex nature of GSIA applications--i.e., multiple and conflicting constituencies, goals and time frames, the "loose-bundled" nature of the GSIA, the complex and poorly understood mechanisms by which the GSIA "works" and

context-specific variations in those mechanisms--provides additional grounds for the importance of user-involvement in the design and operation of a PLUSS. In particular, this unstructuredness and complexity implies that the choices that must be made in designing and operating a PLUSS cannot be made on a purely technical basis. Rather, such choices necessarily depend heavily on non-analytic (subjective and value) considerations which researchers are not uniquely qualified to provide. Thus, user-involvement is important as a means for helping to ensure that the non-analytic inputs relied upon in the design and operation of a PLUSS are reasonable to users; this, in turn, is a necessary (but not sufficient) condition for the information provided by a PLUSS to fit well with users' information needs. To make this point more specific, let us consider again the steps in the process sketched out in Figure 3. With respect to step 1 ("Identify goals"), we have seen that the question of whether a particular goal is relevant cannot be determined in an absolute sense; rather goals can only be defined as relevant on the basis of their saliency to some constituency. Thus, user-involvement in this step is an important safeguard against including irrelevant goals and excluding relevant ones in this step. With respect to step 2, users are likely to possess significant bits of "ordinary knowledge" (Lindblom and Cohen, 1979) concerning the particular system in which the GSIA application is occurring. Consequently, users can identify important precursors

that might otherwise be overlooked while also pointing out precursors which are unrealistic or unnecessary in that context. The criteria suggested in step 4 for selecting those elements on which the PLUSS is to focus are largely subjective. Thus, user-involvement can help ensure that these criteria are applied in a manner that is consistent with the priorities and perceptions of users. In step 5, we have argued, non-analytic inputs, including values, acceptable assumptions and levels of risk, heavily influence the interpretation of PLUSS-based information. Values, acceptable assumptions and levels of risk are properly the domain of the user. Moreover, as noted above, users are often likely to have significant pieces of "ordinary knowledge" that are not available to the researcher, but can influence the interpretation of PLUSS-obtained data.

The final pair of rationales for user-involvement in the design and operation of a PLUSS are logistical in nature and concerned most directly with actually obtaining measurements (step 5). Specifically, user-involvement is a means of "leveraging" limited resources that are available for this task. Secondly, user-involvement can in many cases provide access to data--especially through records and personal observations--that would otherwise be difficult or impossible to obtain.

Practical Considerations Associated With User-Involvement

While the apparent advantages of user-involvement in the design and operation of a PLUSS are impressive, there are certain potential problems associated with user-involvement. These may represent barriers to actually obtaining user-involvement and/or liabilities in the sense that user-involvement, if obtained, may have negative impacts in the larger organizational unit. Some of these problems apply to user-involvement in the overall process depicted in Figure 3; others primarily apply to specific steps within the process. In this section we shall identify some of these liabilities and problems, indicate the conditions in which each is likely to be especially salient, and suggest some means of avoiding or coping with them.

Perhaps the most serious potential problem associated with user-involvement in the design and operation of a PLUSS is that it places a high level of additional demands on the time and effort of users. These additional demands are likely to be a particularly serious problem under the following conditions:

- users resemble the managers described by Mintzberg (1973: 51) in the sense of being called upon "...to perform a great quantity of work at an unrelenting pace;" devoting time and energy to the development of a PLUSS will entail high opportunity costs.

for such users;

-a low level of interest on the part of users in obtaining useful information fro. the PLUSS.

One means of minimizing this potential problem is to involve "user representatives" rather than all users directly in the design and operation of a PLUSS. The notion of a "user representative" is intended to be somewhat similar to "boundary spanners" (Tushman and Romanelli, 1983; Tushman, 1977; Roberts and O'Reilly, 1979). These are individuals who "are strongly linked both within the unit...and to external sources of information" (Tushman and Romanelli, 1983: 13) and act to channel information to the unit from external actors (e.g., researchers). A major difference between our concept of a "user representative" and the more traditional notion of "boundary spanner" is that we envision "user representatives" as serving a two-way, rather than one-way (environment to organization), communication function. It should also be noted that "boundary spanners" are informal roles. Formal designation of "user representatives," however, may be desirable in some situations.

User-involvement in obtaining PLUSS-based information (step 5) is likely to be especially difficult to maintain. For one, user-involvement in this task places additional demands on the time and effort of users over a continuous and extended period of time. In addition, user-involvement in this task may be viewed by users as a threat to their control

over their work environment or as possibly compromising confidentiality of certain information (Cochran et al., 1980; Morell, 1979). Some possible mechanisms for maintaining user-involvement in obtaining PLUSS-based information include:

- make data collection forms easy for the user to complete (e.g., a format similar to one which they commonly use for other purposes; see Mitchell, 1979);
- provide material rewards (e.g., cash, time off, remitted tuition) for cooperation in data gathering (Morell, 1979);
- include data items that are likely to be useful to the user outside of the PLUSS, per se;
- thorough communication to users of the purpose and nature of the PLUSS, how it will benefit them, and how their individual contributions will fit in;
- frequent reports of interim results;
- users are more likely to respond to the request of other users to assist in data gathering than they are to requests coming directly from a researcher; hence, those users who believe the information sought to be important should be enlisted to obtain the cooperation of other users (Rothman, 1980).

Because the choices that must be made in designing and operating a PLUSS depend heavily on non-analytic (subjective

and value) considerations, different users are likely to have differing conceptions of the specific features of an "ideal" PLUSS in a given context. While in some cases it may be possible to develop a PLUSS that satisfies all of these conceptions, in general, we would expect such a system would be too expensive and cumbersome to be within the realm of feasibility. Thus, in most cases user-involvement will require that hard choices be made concerning the relative priority to be given to user A's vs. user B's information needs.

The final potential problem of user-involvement that will be considered here is especially pertinent to step 4. This is the tendency for users to seek an all-encompassing PLUSS, i.e., a PLUSS that provides information about every goal, precursor and activity set identified (see Clark, 1979). This highlights the possible need to "push" users in this step to be selective—perhaps more selective than they might "naturally" be—in screening the list of goals, precursors and activity sets which exists at the end of step 3. In some cases tentative targets for either the absolute number or proportion of goals, precursors and activity sets about which the PLUSS will provide information might be specified in step 4.

MEASUREMENT TECHNIQUES FOR A PLUSS

The purposes and nature of a PLUSS, combined with the con-

texts in which it is to be employed, dictate that an eclectic set of techniques be employed to obtain information concerning an activity category, precursor or goal. The set of data collection techniques employed in a PLUSS should be drawn from a wide-ranging set of tools, including interviews, surveys, direct observations recorded in field logs and unobtrusive measures.

Designing interviews and surveys for use in a PLUSS is a fairly straightforward task which will not be discussed in depth here. The reader is referred to standard works in these areas such as Weiss (1975), Lin (1976), Bouchard (1976A), and Smith (1975).

Field Logs

While field logs will necessarily play a central role in most PLUSS's, we believe that in this context the commonly-practiced version of field logs is inappropriate. In particular, in this commonly-practiced version, field logs are detailed running records of all meetings, interactions, etc. that the fieldworker observes. (For an example of this approach to a field log, see Smith, 1975, Appendix G.). Information pertaining to variables of concern is then sifted out from these logs. While there are situations in which this type of field log is appropriate, we do not believe that field logs incorporated in a PLUSS should generally be of this

form. The primary reason for this is that maintaining detailed running records is generally exceptionally burdensome and unrewarding to fieldworkers--characteristics which are particularly detrimental in cases where site participants are responsible for maintaining field logs. Instead of the traditional running account form, field logs incorporated into a PLUSS should primarily consist of observations that are organized around a small, pre-specified subset of the goals, precursors and categories of activities referenced in the site-specific process model of the GSIA application. Fieldworkers would only be required to enter information pertaining to that subset of goals, precursors and categories of activities into field logs, although they might be permitted, even encouraged, to selectively enter other observations.[11]

Unobtrusive Measures

In selecting data collection techniques to be employed in a PLUSS, unobtrusive measures[12] should not be overlooked. Unobtrusive measures (also termed non-reactive measures) can be defined as means of obtaining data about human behavior without the cooperation, knowledge or awareness (and perhaps even the presence) of those who are being studied (Guba and Lincoln, 1981; Webb et al., 1981). Unobtrusive measures are valuable because they eliminate "reactive measurement errors,"

i.e., the confounding effect on the data obtained of a research subject's knowledge that he is participating in a research effort. A second, and perhaps more significant, virtue of unobtrusive measures is their potential value in "triangulating" other types of observation techniques (e.g., surveys) (Bouchard, 1976B; Guba and Lincoln, 1981).

There are four commonly cited classes of reactive measurement errors.[13] The first of these, which is commonly labelled the "guinea pig effect," refers to the subject behaving in a way in which he does not normally behave in order to make a certain impression on the researcher. Closely related to the "guinea pig effect" is "role selection," in which the respondent selects one of many "true" selves to display to the researcher. The third major type of reactive measurement error is "measurement as change agent," which refers to the introduction of "real changes" in what is being observed due to the measurement process itself. "Destructive testing" is a classic example of this effect in the physical sciences and engineering. The effect of "preambles" in questionnaire or interview items on the knowledge and attitudes of respondents is another classic example of this effect. The final major class of reactive measurement errors is "response sets," i.e., the effect of the position or wording of a question in a questionnaire or interview or the position or wording of possible responses

to a question on the respondent's reply. For example, Webb et al. (1981: 57) note that "respondents will more frequently endorse a statement than disagree with its opposite." Another example of the response set effect, reported by Nisbett and Wilson (1977), is the effect of the position (left-to-right) of an array of consumer goods on respondents' assessment of their quality.

Triangulation

In some cases there may be no single data source which can yield sufficiently valid information concerning some goal, precursor or activity category of concern. This may indicate that the PLUSS should not (cannot) be relied upon to provide information concerning that goal, precursor or activity category. However, in some cases "triangulation," i.e., combining multiple data sources, can be used to provide sufficiently valid information.[14] Two forms of triangulation--between-method and within-method -- are of concern here.[15] The former refers to using multiple methods (e.g., surveys along with unobtrusive measures) to observe a given variable, while the latter refers to multiple variants of the same basic method (e.g., multiple unobtrusive measures, multiple pencil-and-paper scales to measure some attitude, multiple observers) to observe a given variable. The rationale behind triangulation is most clearly stated by Webb

(1970: 450):

> "Every data-gathering class--interviews, questionnaires, observation, performance records, physical evidence--is potentially biased and has specific to it certain validity threats. Ideally, we should like to converge data from several data classes, as well as converge with multiple variants from within a single class."

While triangulation can often increase the validity and credibility of information produced by a PLUSS, certain caveats should be noted. First, triangulation obviously adds to the overall cost of data collection. Second, a key condition for triangulation to significantly increase the validity of information gathered is that the weakness of each data source employed must correspond to the strengths of another (Jick, 1979).[16] Third, while triangulation increases the validity of observations of a variable, it does not counteract other threats to validity in drawing causal inferences such as history and maturation. Finally, few practical guidelines for interpreting data collected from multiple sources exist. That is, the basic rationale underlying triangulation suggests that <u>if</u> the results of the different data sources employed are consistent with each other (converge), confidence in the validity of the findings are presumably strengthened. While on the surface it might seem relatively straightforward to determine whether the results of different data sources employed converge, in practice this task is often a problematic one for which few

guidelines exist. Moreover, the results of multiple data sources may diverge rather than converge. In some cases these contradictions are irresolvable. In other cases, however, divergence "can...turn out to be an opportunity for enriching the explanation (of a phenomenon)" (Jick, 1979: 607). Again, however, there are few guidelines for reconciling divergent results. (See Jick, 1979, and Trend, 1978, for case studies of the process of such reconciliation.)

ERRORS IN PERFORMANCE-SENSING AND LEARNING

In order to make appropriate choices concerning the design and operation of a PLUSS, a clear understanding of the design criteria to be applied in such choices is crucial. The traditional design criterion used in applied social science is whether the data provide accurate answers to the questions posed. In the case of evaluation, this criterion reduces to: do the data provide accurate estimates of the extent to which a specified intervention has influenced a given set of dependent variables? In other cases (e.g., surveys), questions of concern are of the following general form: what are the levels of a given set of variables at a given point in time and what are the relationships among those variables?

This traditional design criterion is, in our view, overly narrow for use in the design and operation of a PLUSS. In this section a broader view, based on the notion of errors

in performance-sensing and learning, is presented as an alternative design criterion. In this view, a "good" PLUSS is one which minimizes the extent to which various errors in performance-sensing and learning are committed. As we shall see, providing accurate answers to questions posed does not necessarily eliminate errors in performance-sensing and learning. Factors such as whether the correct questions are addressed (Mitroff and Bonoma, 1978), timeliness and credibility also affect errors in performance-sensing and learning. In particular, we consider six major types of possible errors in performance-sensing and learning. Two of these types of errors are associated with distilling the GSIA, three are associated with reinvention and the sixth is associated with assessing the effectiveness of a GSIA application. We also consider major sources of these errors and, where possible, suggest means of avoiding these sources of errors.

The two types of possible errors associated with distilling the GSIA are as follows:

I.1: failure to recognize activities which are included in a GSIA application;

I.2: believing that an activity is included in a GSIA application when, in fact, it is not.

Potential errors associated with reinvention can be collectively labelled as problem-sensing (Kiesler and Sproull, 1982) and opportunity-sensing errors. The three types of errors in

this category are:

II.1: failure to detect the non-occurrence of necessary precursors to a goal ("problem stimulus") in time to respond appropriately (take corrective action);

II.2: failure to recognize an "opportunity"--i.e., the availability of precursors which, if achieved, would eliminate the need to achieve some other precursors--in time to take advantage of it;

II.3: erroneously believing that a precursor is necessary for achieving a goal and has not occurred (i.e., needlessly defining a situation as a problem).

The final type of possible error of concern to us (III.1) is the failure to accurately assess the effectiveness of a GSIA application in time to meet various organizational deadlines.

While there are numerous sources of such errors, six are of particular concern in the context of a PLUSS. The relationships between these sources of errors and the types of errors discussed above are summarized in Figure 4.

The first source of performance-sensing and learning errors is inclusion of superfluous elements (goals, precursors and/or activity categories) in a process model. Including superfluous precursors in a process model can lead

Sources of Errors	Error Type I.1	I.2	II.1	II.2	II.3	III.1
INCLUSION OF SUPERFLUOUS ELEMENTS IN A PROCESS MODEL		X			X	X
"TUNNEL VISION"	X		X	X		X
LACK OF TIMELINESS			X	X		X
DATA CONTAMINATION	X	X	X	X	X	X
LACK OF CREDIBILITY	X	X	X	X	X	X
INFORMATION OVERLOAD	X	X	X	X	X	X

Figure 4

to error type II.3 since if a specified precursor has not occurred, it generally will be interpreted as representing a problem; however, if that precursor is, in fact, not crucial for the achievement of some goal, interpreting such a situation as a problem would be erroneous. Including superfluous goals in a process model can also lead to error type III.1. If, for example, data indicate that the goal in question has been attained, the natural result would be an increase in the perceived effectiveness of the GSIA application. If, however, this goal was, in fact, not relevant to any stakeholder, this increase in the perceived effectiveness of the GSIA application would be unjustified. Finally, including superfluous activity categories in a process model can lead to error type I.2, although accurate data on the occurrence of activities in these categories would enable one to avoid this type of error by indicating that, in fact, no activities in these categories have actually been performed.

The second source of performance-sensing and learning errors is labelled "tunnel vision" and is the mirror-image of the previous source of errors. This source of errors, which involves failing to consider relevant elements in a process model, can lead to error types I.1, II.1, II.2 and III.1. A number of suggestions for avoiding these problems have been presented in "Overview of the Design and Operation of a PLUSS," above. These include the use of stake-

holder analysis, nominal group technique and "brainstorming" in steps 1, 2 and 3 of Figure 3. The regularly scheduled opportunities for revising the PLUSS are also expected to guard against this source of errors. "Theoretical triangulation" (see above), in the sense of retaining alternative plausible chains of precursors to various goals, is also primarily intended as a means of avoiding this source of error (see Mitroff and Emshoff, 1979). Finally, encouraging observers to be alert for and report relevant serendipitous findings can help avoid this problem.

Lack of timeliness, the next source of performance-sensing and learning errors, can lead to errors II.1, II.2, and III.1. This is a compelling argument in support of the above-stated prescription to be very selective in terms of the scope of information for which the PLUSS is to be relied upon. Reliance on informal, verbal communication rather than solely on formal, written reports (Cox, 1977; Kiesler and Sproull, 1982) can also contribute to avoiding this problem.

The next two sources of performance sensing and learning errors--data contamination and lack of credibility--are potentially associated with all six types of errors we have identified. "Data contamination" refers to the factors which cause data to be distorted pictures of reality. Included here are the standard threats to validity commonly

considered in evaluation research (see Cook and Campbell, 1979; Freeman, Rossi and Wright, 1979) and other forms of applied social science (see Webb et al., 1981; especially Chapter 3).

As used here, credibility is related to, but distinct from, truth and validity. In particular, by truth we mean an actual state of affairs. In this sense, truth is an objective reality (though we cannot always know that objective reality). Validity has taken on several meanings, but here we use the term as it refers to the logical relationship among propositions. Conclusions are valid if they flow logically from assumptions. Thus, a conclusion may be true and valid, false and valid, true and invalid, or false and invalid. Credibility involves the acceptance or the assessment of the plausibility of an a 'ertion. Truth and validity are neither necessary nor sufficient conditions for the acceptance of research findings. Also, it should be noted that credibility is a variable (unlike truth and validity) that is determined by social and psychological factors and can take different values for different individuals. To put it another way, one man's cup of knowledge is another man's poison. In the next section of this report, literature that sheds light on this source of errors and how it can be guarded against is considered in depth.

The final source of errors in performance-sensing and learning, information overload, is likely to result in users' distorting and ignoring pieces of information. Because the specific bits of information which will be ignored or distorted will vary widely, overload can lead to any or all of the six types of errors identified. This source of errors represents still another rationale for selectivity in terms of the scope of PLUSS-based data.

DETERMINANTS OF THE CREDIBILITY OF PLUSS OUTPUTS

Applied social research is notorious for its inability to promote consensus and, thus, considerations of the social and psychological determinants of credibility are perhaps more important here than in any other realm. To help us use studies on the determinants of research acceptance, we begin with an elementary conceptual model.

In the transmission of a PLUSS-based information we can identify a receiver and a sender of information. The receiver may be a specific client, a broad community of users, a narrow community of specialists or the general public. The sender may be the researcher and his associates, a synthesizer of research, or a "gatekeeper" (Allen, 1977). We can also identify a medium for the information, such as a formal research report, a verbal presentation, a popular account and so forth. In reality these categories are often overlapping and closely inter-related. This is especially

the case in a PLUSS where the users help frame the research and its medium of presentation. Nevertheless, these categories are useful in organizing studies of credibility. Figure 5 presents a basic model, the model used here to organize findings, and Figure 6 presents a more complicated model that is more realistic for a PLUSS.

In Figure 6 it is recognized that the receiver (especially when the receiver is a client) often influences the sender. The receiver not only influences the sender by direct interaction but also, in many cases, by the very choice of sender (especially when the sender is a contract researcher or trainer). By the same token, the receiver often influences the medium either by mandating a specific medium (i.e., technical report, informal verbal presentation) or by influencing the sender's choice of medium. Likewise, the medium influences the choice of sender. Unfortunately, little of the research pertaining to determinants of credibility has focused on those features represented in Figure 6 but deleted from Figure 5. The discussion presented below organizes research on the basis of the three elements common to both models. Table I summarizes the determinants of credibility within each category which are discussed.

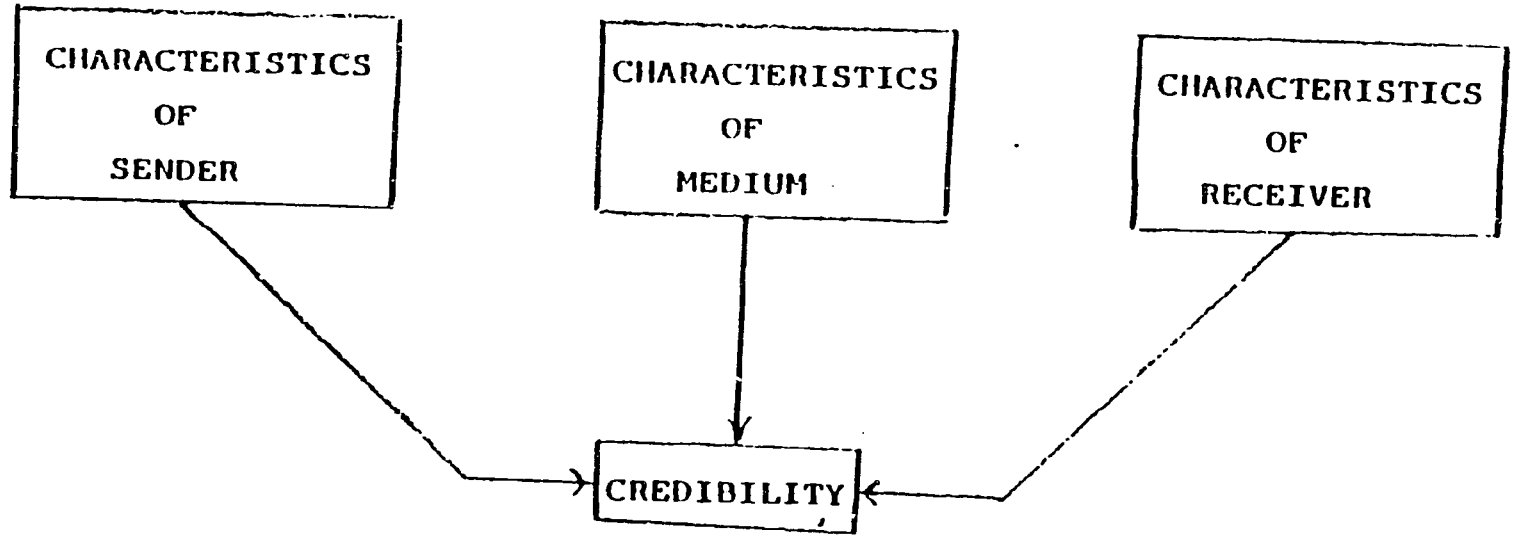
CHARACTERISTICS OF SENDER
CHARACTERISTICS OF MEDIUM
CHARACTERISTICS OF RECEIVER
CREDIBILITY

Figure 5

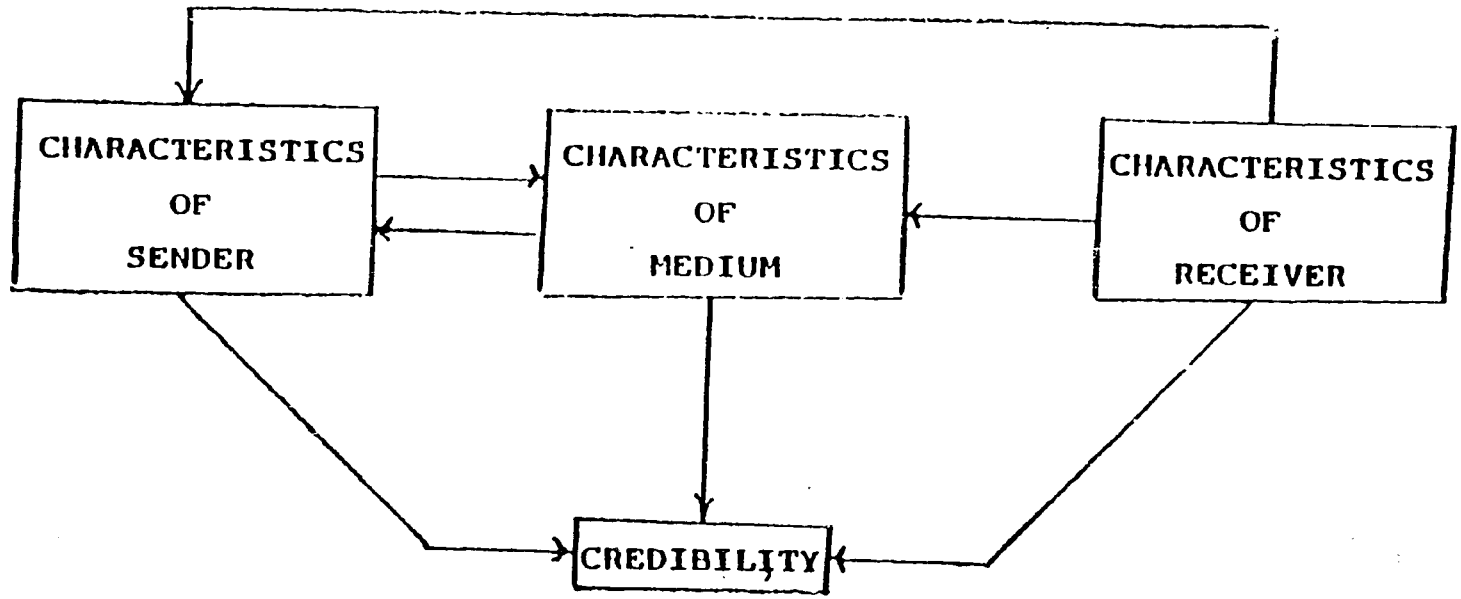
CHARACTERISTICS OF SENDER
CHARACTERISTICS OF MEDIUM
CHARACTERISTICS OF RECEIVER
CREDIBILITY

Figure 6

TABLE 1. SUMMARY OF FACTORS AFFECTING CREDIBILITY OF PLUSS OUTPUTS

CHARACTERISTICS OF MEDIUM

- Jargon
- Amount, Type and Form of Presentation of Data
- Readability
- Negative vs. Positive Information
- Relevance
- Correspondence with Receiver's Day-to-Day Experience
- Proximity and Accessibility of Media

CHARACTERISTICS OF SENDER

- Personal Characteristics
- Professional and Occupational Characteristics

CHARACTERISTICS OF RECEIVER

- Personal Characteristics
- Professional and Occupational Characteristics
- Organizational Position

Credibility And Characteristics Of The Medium

The effects of media characteristics on confidence-inspiringness is of special importance in a PLUSS because it is this set of determinants that is most often manipulable. Applied researchers may have little control over characteristics of the receiver (though they are well advised to take those characteristics into account), but they often can tailor the characteristics of the medium for the needs of the receiver. In this section, a number of studies are reviewed which have examined the effects of media characteristics on credibility or closely related variables.

Jargon, Use of Data and Readability

The effects of jargon, use of data, and readability on credibility are more easily assessed than other determinants in controlled experiments. A number of experimental and laboratory studies dealing with the form and structure of information as it relates to credibility and use of results are reviewed by Braskamp and his associates (1982).

Two studies (Thompson et al., 1981; Brown et al., 1978) address specifically the issue of jargon's effects on the credibility of information. In the initial study (Brown et al., 1978), educators read reports about testing in public schools and were asked to rate the reports in regard

to clarity and agreement with the researchers' recommendations. The reports were identical in content except that one set used more technical jargon and another set used more data. The jargon in the report included technical terms about different types of achievement testing. Brown and his associates found that jargon affected ratings of the technicality and difficulty of the report. This finding is, however, of questionable significance in light of the close relation between the experimental variable (the technical jargon) and the dependent variable (assessment of the technical difficulty of the report). More interesting was the finding that the subjects rated the jargon-loaded report with no data support most difficult to understand (even more difficult than the report that included jargon and data) and the jargon-free, data-supported report was rated the most easily understood. Significantly, the amount of jargon or data had no effect on the subjects' agreement with the recommendations in the report. It should be noted, however, that this latter finding is not consistent with several studies which have found a relationship between the use of data and agreement with information (Gifford et al., 1979; Wright, 1974; Porat and Haas, 1969).

The study conducted by Thompson and associates (1981) is a follow-up study to the Brown and associates study. In the later study educators and business persons read educational

research studies which were variously jargon-loaded or jargon-free and data-free or data-supported. They rated reports void of jargon and data the most believable relative to the other three combinations. They gave consistently poor ratings to reports which contained language inconsistent with the level of data presented; that is, an absence of jargon was preferred for data-free reports, but jargon was actually preferred (even without a change in message content) for the reports that included data. Several writers have suggested that more readable reports are more likely to be positively evaluated and used (Agarwala-Rogers, 1977; Windle and Bates, 1974; Levitor and Hughes, 1981). The findings of Thompson and associates seem to mitigate (though not contradict) the assumption that readability has consistently positive effects. The readability of the report should be appropriate for the level of technical treatment and the sophistication of data analysis.

Not all data are created equal. In a study of the effects of different types of statistical information, Lorenzen and Braskamp (1978) found that cost-benefit information was rated by administrators as more helpful than either other forms of statistical information or political information. This finding obtained for a variety of decision problems and regardless of the position of authority

or tenure of the administrator.

Brown and Newman (forthcoming) examined teachers' and administrators' responses to recommendations as a function of the amount of data and the form of presentation of data in otherwise identical reports. Those who read reports including percentages and graphs were strongest in their agreement with the findings of the report; those who read reports with percentages, graphs, and statistical inference statements were least supportive of the recommendations.

Negative and Positive Information

In a study that has important implications for a PLUSS, Locatis and his associates (1980) found that positive and negative information have different intensities of effects on ratings of usefulness and credibility. It is only a mild oversimplification to say that their study shows that negative information has more potential for harm than positive information has for good.

The study was based on reactions to an educational film. The researchers showed the film to 204 graduate students who were randomly assigned to groups which received no critical review of the film, positive review, negative review, or mixed review. Some sub-groups received reviews before viewing the film, others received them afterward. The researchers found that positive information inflated judgments about the

film but not nearly to the degree that negative information depressed judgments. It made no difference whether the review was received before or after the showing of the film.

The authors conclude that the results of the study suggest "extra care in reporting negative information, particularly when the object of the evaluation has many strong points and the decision makers lack first-hand contact with whatever is evaluated" (Locatis et al., 1980: 822; emphasis added). The authors suggest that researchers may wish to take care to fully describe results and highlight positive information so as to offset the powerful effects of negative information.

The conclusions from the study by Locatis and associates merit close attention but there is some question as to the study's generalizability. The viewing of a film may not be the same as, say, the reading of a research report and graduate students may be more amenable to influence by negative information than persons closer to (and more familiar with) the object of study. Nevertheless, it is worth noting that several studies (Carter, 1971; Floden and Weiner, 1978) have come to similar conclusions about the stronger effects of negative information.

A somewhat different set of findings were obtained in a study mentioned above. Brown and Newman (forthcoming) found when a researcher's recommendations were supportive of

the program (i.e., positive), the researcher was rated as more technically proficient and a better communicator than the researcher making negative recommendations. This does not, however, contradict the findings obtained in studies such as that of Locatis and his associates because the subjects in the Brown and Newman study worked in the program that was evaluated and had a clear stake in the outcome of the evaluation.

Relevance

Several studies relate perceived needs of message receivers (usually clients) to use of information and assessment of credibility. Nielsen (1975) compared program managers' needs for information with the information actually received from applied evaluative research. There was considerable divergence which is somewhat surprising in light of the fact that the sender of the information had, in most instances, contracted with the receivers for a specific package of information.

David (1978) found low use of applied research by school districts and concluded that the low rate of use was a function of the low relevance of the information content. In another study, Resnick and his associates (1979) looked at research reports with varying degrees of perceived relevance and found that perceived relevance was a major deter-

minant of both use and favorable assessment of research. The study also investigated the relevance of research as a function of the position of the recipient and found that there was considerable divergence in the needs and uses for research between policymakers and program staff. Other studies (Pelz, 1979; White and Murnaghan, 1973; Weiss and Bucuvalas, 1977) have come to similar conclusions.

It also appears that the credibility of research is affected by the degree of correspondence between the research findings and receivers' day-to-day experience concerning the topic of the research (Patton et al., 1977; Caplan et al, 1975; Weiss and Bucuvalas, 1977). Results that are counter-intuitive face much stronger resistance than research findings that support existing views. If one views PLUSS-based information as either an improvement on, or a useful alternative to, individual experience, this finding cannot be treated lightly. The implication is that a PLUSS may often be doomed in any attempt to overthrow existing well-entrenched views.

Proximity And Accessibility Of Media

One of the more consistent findings in information and communication research is that the accessibility of an information channel is one of the most important determinants of information use. The finding has been obtained in experi-

mental studies (Gifford et al., 1979; Guetzkow and Simon, 1957) as well as field studies of scientists and engineers (Gertsberger and Allen, 1968; Pelz and Andrews, 1976, and public agencies (Bozeman and Cole, 1981; McGowan and Loveless, 1981)

It is not surprising that the accessibility of an information channel is an important determinant of use, especially among busy managers (Blandin and Brown, 1974). More interesting is O'Reilly's (1982) finding that accessibility of information is even more important than quality (and, relatedly, that quality and use are poorly correlated). In his study based on survey data obtained from 163 workers in four branch locations of a county welfare agency, O'Reilly found that while the relative importance of quality and accessibility varied according to the intended use of the information, accessibility was consistently more important in explaining use. He attributes this finding to the fact that users incur costs in seeking out quality information not readily available and that the users were under pressure to process large numbers of cases under severe time constraints. Bozeman and Cole (1981) account for similar findings in public agencies by noting that information-seeking behavior is likely to go unrewarded and, if it detracts from operations and routine activities, may even result in punishment.

Another consistent finding is that information that is

transmitted by informal communication (especially face-to-face discussion) is more widely preferred, more often used and, sometimes, viewed as more credible (see, for example, Marquis and Allen, 1966; Bacharach and Aiken, 1977; Crane, 1972; Bozeman and Blankenship, 1979). An especially relevant study by Stevens and Tornatzky (1980) examined the effects of formal and informal communication in 37 drug abuse programs and found that for tasks that are uniform and handled by "bureaucratic" structures, formal communication is as useful as informal communication in promoting change. However, in tasks that are non-uniform and informal, face-to-face communication is much more effective in promoting change. These findings are supportive of earlier studies in different settings (Litwak, 1961; Fairweather et al., 1974; Taylor and Utterback, 1975).

Credibility And Sender Characteristics

Research on the effects of sender characteristics on the perceived credibility of information is ubiquitous. Opinion researchers have been drawn to the topic but industrial engineers, information scientists and psychologists have made substantial contributions as well. Unfortunately, when the sender is a researcher (rather than an opinion leader, "gatekeeper," or advertiser) there is less informa-

tion about the effects of sender characteristics on credibility. Most of the more relevant research can be subsumed under two categories: personal characteristics of the researcher and professional characteristics of the researcher.

Personal Characteristics

The credibility of information provided by applied researchers varies as a function of the personal characteristics of the researcher, even if the content of the information is constant. One of the variables that often has an effect on credibility is the sex of the researcher. In one study (Newman et al., 1979) business and education students and business and education professionals were asked to read applied research reports which were identical in every respect but one: for one randomly chosen sub-group a male name and picture were provided for the "researcher" and for the other sub-group a female "researcher" was identified. Both the students and the professionals were more likely to agree with the findings of the male researcher than the female researcher (though the findings were identical). The sex of the students and professionals had no apparent effect on their bias toward the male "researcher." However, sex bias was reduced by familiarity with the topic of the research, implying that secondary considerations, such as sex-based bias, more often come into play when there is

greater uncertainty regarding the information.

A great many studies conducted by psychologists provide evidence that the cognitive style of the message sender affects not only the message itself but also the credibility of the message. However, few of these studies seem generalizable to the question of the credibility of applied research (Weinberg, 1979; Staw and Ross, 1978). One study that is especially relevant argues that the applied researcher's choice (implicit or explicit) of decision-making model is relevant to the likelihood that his or her research findings will be viewed as sufficiently credible to affect organization change. DeYoung and Conner (1982) suggest that applied researchers with an essentially rational view of organization decision-making are in most instances less likely to have research results put into action than those who have an incremental view of organization decision-making.

In a study that adapts many of the findings from studies of engineers and scientists, Bozeman and Cole (1981) found that a group of program managers in a technically-oriented public agency were likely to seek information from colleagues within the organization and especially from those individuals who were more highly trained in technical fields and who were themselves more active in seeking information from a variety of media. A subsequent study (Bozeman and

McGowan, 1981) corroborated these findings for a much larger and more differentiated sample of public managers.

Professional And Occupational Characteristics

The professional and occupational attributes of the researcher often affect perceptions of the credibility of research. In one experiment (Braskamp et al., 1978) teachers and administrators were asked to rate research findings produced by a "researcher" and "evaluator" and a "content specialist." The content of the report was held constant. Teachers and administrators rated the report as more credible when attributed to the researcher than when attributed to the evaluator or content specialist.

Both professional credentials and professional activities are relevant in assessments of credibility. In a pioneering study, Hovland and Weiss (1951) found that the status of professions was a highly significant determinant in evaluations of expertise and credibility. Bozeman and Cole (1981) found that public managers placed a higher value on the information provided by colleagues who were more involved in professional activities such as attendance of professional conferences and holding professional memberships. Others have come to similar conclusions in research conducted in a wide variety of settings (see, for example, Strong, 1978; Glaser and Taylor, 1973; Gertsberger and Allen, 1968).

One of the more important factors mediating the effects of professional and occupational status on credibility is the receivers' perception of the quality of information provided in the past by the same source (O'Reilly, 1982; Zmud, 1978). However, these findings are complicated by variance in individuals' propensity to assume risk.

Credibility and Receiver Characteristics

Receiver characteristics impinging on credibility can be categorized in much the same manner as sender characteristics, but to personal and professional categories we add a third--organizational position.

Organizational Position

In one of the most comprehensive studies of the effects of audience (receiver) characteristics on assessments of credibility, Braskamp and his associates (1978) examined teachers' and school administrators' ratings of information provided by an external applied researcher and an internal applied researcher. The administrators rated information provided by an external researcher as more useful than did the teachers but there was no difference in ratings of the usefulness of information provided by the internal researcher. The professional experience of the recipient was correlated to assessment of written research results. Students in the fields of business and education rated the applied research

more highly and were more in agreement with the research's findings than were the teachers and administrators.

There is considerable research on the information needs of managers and there is some evidence that managers have different views than technicians and non-management professionals of what constitutes useful and credible information. Cox (1977) contends that management theory, and especially Mintzberg's (1973) studies of the nature of managerial work, can facilitate efforts to provide managers with more useful and credible information. Specifically, he suggests: (1) data from applied research will almost never be the sole source of information about a program and, from the managers' point of view, it is unlikely to be the most important; (2) managers' interests are specific and concrete and research directed at managers should generally have these same characteristics; (3) managers have a casual attitude toward validity and technical problems of analysis and if research is to be useful, it should identify but not dwell on technical problems; (4) informal communication is preferred by managers and if information is to be useful, it should be reported succinctly, simply, informally and frequently. Bunker's (1978) analysis complements the suggestions provided by Cox. Bunker seeks to identify the uses of applied evaluative research for managers and contends that five managerial functions are served by such information: clarification of objectives, critical assessment of "inherited" ideas and policies, definition of local operating practices (including identification of unique features),

detection of the fit between actual operations and program design and assessment of the fit of program concepts and operating patterns with key social values. Bunker implies that information addressed to these issues is generally more useful for managers than information provided in traditional summative program evaluations.

Personal Characteristics

A number of personal characteristics of receivers influence assessments of the utility and credibility of research. In general, individuals who are more actively engaged in processing information are more receptive to research-based information and those who are users of data-based information are more likely to accept research findings that are counter-intuitive. Persons who process little information and are not data-oriented are more likely to reject research findings that are not in accord with their personal experiences (Leviton and Hughes, 1981).

In a study focusing on decision-complexity and individuals approaches to decision-making, Weiss (1982) found that older and more experienced public managers were less likely to view externally-generated information as credible, less likely to use it, and as a result, performed more poorly on unstructured decision problems.

Two other studies examined the internal-external orientation of receivers to determine its effects on assessment of in-

formation received (Brown and Newman, 1980; Newman and Brown, 1980). Externally-oriented educators (as measured by a standard locus of control scale) preferred applied research that included specific recommendations for action. By contrast, internally-oriented recipients preferred to draw their own conclusions and make their own recommendations after reviewing the applied research results.

Professional and Occupational Characteristics

In their study of the effects of research recipients' professional experiences on reactions to written applied research reports, Newman and associates (1979) found that students in the fields of business and education rated the findings more highly than did working professionals in those fields. The extent to which the effects of age, experience and knowledge confound those findings is unclear. In another study (Hawkins et al., 1978) of the effects of professional and occupational characteristics on assessments of utility and credibility of information, differences according to occupation were uncovered. Persons in different occupations used different criteria in judging the utility of various types of information. Staff members of human service agencies were more impressed with assessments of staff competence, planners were more favorably disposed to information about managerial efficiency and judicial probation personnel responded more

favorably to information about client impact. Interestingly, none of the groups placed a high priority on information about the cost of programs and agency operation.

LINKING CASES

Thus far, we have treated the PLUSS developed in conjunction with one GSIA application as being independent of PLUSS's developed in conjunction with other GSIA applications. In this section we address issues associated with linking PLUSS's developed in conjunction with various GSIA applications. We start out by considering the traditional goal and techniques of linking case studies. We then discuss some reservations we have concerning the feasibility and appropriateness of this traditional goal. We conclude by suggesting an alternative purpose of linking case studies in the current context and discuss how this alternative purpose might be achieved.

Traditionally, social scientists link various case studies in an effort to develop a generalizable theory to account for some closely related set of phenomena. A theory, as the term is used here, consists of a small set of accurate explanations of the phenomena of concern, along with a specification of the conditions under which each explanation applies. (Mohr, 1982, utilizes the term theory in a similar sense.) Two basic types of approaches for linking case studies in order to develop a generalizable theory of certain phenomena--the case-survey method and the case-comparison approach--have come to the forefront in recent years. While both are somewhat primitive and not yet sufficiently documented to form clearcut and specific guidelines, there have been rapid advances and numerous applications of these two approaches in recent years (see Louis, 1982; Yin, 1981). It

should be noted that in both approaches a series of within-case analyses are typically conducted simultaneously (i.e., in parallel streams) and, when completed, used as inputs into cross-case analysis. Thus, cross-case analysis typically is not begun until all within-case analyses have been completed, thereby limiting the extent to which the various within-case analyses "feed" into each other. Since GSIA applications occur sequentially (both within a given target system and among target systems), this characteristic would require that these two approaches to cross-case analysis be adapted if used to link PLUSS's.

Of the two major approaches to cross-case analysis, the case-survey method (Lucas, 1974; Yin, 1981) is the more highly structured. Lucas (1974) summarizes this approach as follows:

> "[t]he analyst prepares a set of questions to determine the presence and intensity of common characteristics, events, and outcomes contained in each of the case studies. The possible answers to the questions are carefully structured and defined so that the analyst, after reading the case materials, can readily determine the most appropriate response. The answers to these questions are determined in the same manner for each of the cases that have been selected for study. The results can then be put in a machine-readable form and analyzed."

The case-survey method, it should be noted, requires a relatively large number of cases.

Different variants of the case-comparison approach rely, in one form or another, on relatively informal pattern recognition (see Smith and Robbins, 1982; Huberman and Crandall, 1982;

Yin, 1981; Yin, 1982). That is, various single-case explanations (within-case analyses) are arrayed in some fashion and the researcher attempts to identify common patterns among the single-case explanations. The key in the case-comparison approach is to develop procedures, involving both process-type procedures and means of summarizing and displaying single-case results, to stimulate pattern recognition (Louis, 1982).

When the general statement of the traditional goal of linking case studies is phrased in terms associated with the current context, the specific form this traditional goal takes is somewhat as follows: "develop a generalizable theory to account for the success (or lack of success) of the GSIA." Put in these terms, we have serious reservations about both the feasibility and appropriateness of the traditional goal of linking case studies in the current context. Let us first discuss our reservations concerning the feasibility of this traditional goal. We will then turn to the appropriateness of this goal.

There are two major barriers to developing a generalizable theory to account for the success of the GSIA. First, while in abstract terms there appear to be substantial commonalities among GSIA applications, there may be too much diversity in terms of operational measures (in a research sense) involved in various cases. For example, while goals in all site-specific process models might be generically labelled "achievement of development objectives," the operational definitions of these goals will vary substantially among sites. "Cross-site" analysis

would be aimed at uncovering a generalizable theory to account for the achievement (or non-achievement) of these goals, but, in fact, because achievement of the respective goals in various sites are essentially different phenomena, it is unlikely that the same theory applies to all of these cases.

A problem similar to the one involved here has been encountered in at least two other areas of research. In the area of organizational effectiveness, the idiosyncracies present in operational definitions of effectiveness that are appropriate for various organizations (even though each can be legitimately termed effectiveness) are generally recognized as a major impediment to the development of generalizable theories of organizational effectiveness (Hannan and Freeman, 1977). The second area which is instructive in this regard is innovation. Time of adoption and extent of use are two operational variables used to represent the concept innovation. While both can be legitimately included under the same umbrella term, they are too heterogeneous to expect the same factors to determine variations in both of these variables (Mohr, 1982).

The likely non-existence of a single theory to account for multiple phenomena which can be given the same label but are, in operational terms, considerably different can also be illustrated by considering the phenomenon "becoming temporarily disabled." If we are considering two individuals who have become temporarily disabled, we might at first expect to find

a single theory that fits both cases. Suppose, however, that in the first case the disability is due, say, to a broken leg while in the second case it is due to, say, pneumonia. Given this variation in the two phenomena to be explained (even though, at a high enough level of abstraction, they share the same label), we would no longer expect to find a single theory that fits the two cases.

The second major barrier to developing a generalizable theory to account for the success of the GSIA is the extreme context-specific nature of causes of social phenomena (see Cronbach, 1975; Guba and Lincoln, 1981; Mohr, 1982). That is, an explanation of some phenomenon--whether physical or social--is always valid only under certain conditions. The rub is that the validity conditions for explanations of social phenomena are often extremely elusive. Establishing these conditions, at best, requires a large number of cases; conceivably, it is an impossible task to do so. For one, these validity conditions depend on a large number of background variables. Second, an explanation of a social phenomenon often holds only over a narrow range of each background variable. Third, designating the various background variables as Z_1, Z_2, Z_3, etc., the range of Z_1, say, under which an explanation of some social phenonemonen is valid may depend upon the values of Z_2, Z_3, etc. Finally, some of these background variables may be unmeasured, often because their importance is unrecognized.

Our reservations concerning the appropriateness of attempts to develop a generalizable theory to account for the success of the GSIA, aside from the feasibility of such an undertaking, stem from the recognition that, unlike many other social science research endeavors, in the current context theory development is not an end in itself. Rather, theory development should be seen as a possible means of achieving a pragmatic goal, namely, facilitating the effective and efficient application of the GSIA. More specifically, the main purpose of linking cases in the current context is to enable site participants in any one GSIA application to benefit from experiences obtained in other GSIA applications. That is, viewing a GSIA application as a "learning sequence," linking cases might best be viewed as an effort to "use previous experience to facilitate and accelerate this learning" (Johnston and Clark, 1982: p. 220). Accelerating learning regarding data collection techniques and instruments as well as substantive knowledge concerning the GSIA is of concern here.

While a valid theory which accounts for the success of GSIA applications would contribute to the goal of enabling site participants in any one GSIA application to benefit from experiences obtained in other GSIA applications, we wish to propose a simpler, more feasible and quicker, if less elegant, means for doing so. In particular, we believe that the best way to enable site participants in a GSIA application to benefit from experiences obtained in other GSIA applications is to provide each

site participant with access to a concise "catalog" of information regarding previous GSIA applications and the PLUSS's developed in conjunction with them. This catalog might be figuratively (and perhaps literally as well) thought of as consisting of two loose-leaf binders which are continuously updated. The first loose-leaf binder would contain substantive information concerning previous GSIA applications, while the second would contain instruments used in the PLUSS's developed to date.

The basic "unit of entry" in the section of the catalog aimed at conveying substantive information concerning previous GSIA applications would be individual applications. Each entry would consist of two parts. The first would summarize, perhaps in flowchart form, the key activities and outcomes which occurred (or are expected to occur) during, or as a result of, the GSIA application being summarized. This would, in other words, simply consist of the "latest" process model of that GSIA application. This portion of the entry for each GSIA application should contain cross-references to instruments used to obtain information concerning the activities, precursors and goals referred to in the process model which are contained in the "instrument" section of the catalog.

The second portion of the entry for each GSIA application would consist of a brief narrative summary of the setting and background of that GSIA application. The purpose of this is to enable those looking at a process model of a previous GSIA application to gauge the degree of fit between the context in which that previous GSIA application took place and the context in which the reader finds himself. The degree of fit between these two contexts, in turn, provides an indication of whether it is appropriate to extrapolate parts or all of the process model of the GSIA application described in the catalog entry to the GSIA application in which the user is participating (see Guba and Lincoln, 1981; Johnston and Clark, 1982).

Where two or more GSIA applications share a common pattern in terms of the activities and outcomes <u>in operational terms</u> involved in those applications and the relationships among them, the entries for those applications should be clustered. The process of identifying such clusters would be very similar to the "case-comparison" approach described above, although here the comparisons would be conducted sequentially. That is, each time a new GSIA application is added to the catalog (or the process model of a GSIA application already contained in the catalog is modified), its process model would be compared to existing entries in the catalog and clustered with existing entries, if appropriate. The advantage of forming such clusters is to make the catalog more succinct to the user (though probably not in the physical length of the catalog). Given the two

possible barriers discussed above to developing a generalizable theory to account for the success of the GSIA, we would expect there to be a large number of clusters, most containing only one or two cases. It should be noted that this section of the catalog is likely to be useful for transferring knowledge obtained in GSIA applications to other sites almost immediately (although we would expect it to become increasingly useful for this purpose as the number of entries it contains increases). Moreover, if the barriers discussed above to developing a generalizable theory to account for the success of the GSIA are not as severe as we suggest they are, this section of the catalog in the long run will form the basis of such a theory.

The most appropriate basic "unit of entry" in the section of the catalog aimed at transferring data collection instruments among PLUSS's is single techniques or instruments. In most cases a copy of the instrument (questionnaire, interview schedule) will suffice. In other cases (e.g., most unobtrusive measures) it may be necessary to prepare a thumbnail description of the instrument for inclusion in the catalog. The key to success for this section of the catalog is a directory organized by activities and outcomes that indicates which existing instruments include items to provide information about that activity or outcome.

FOOTNOTES

[1]Management, as used here, refers to the process of organizing, directing and coordinating activities and the use of resources to achieve desired ends. See USAID (1981) for a similar definition.

[2]The DPMC is housed within the Technical Assistance Division of the U.S. Department of Agriculture's Office of International Cooperation and Development.

[3]We wish to emphasize that these two characteristics are not necessarily the most salient for applying the approach nor the chief determinants of its success. Rather, these are the two most salient characteristics in considering the design and conduct of applied research on the approach. In other words, our comments on applied research are believed to apply to managerial innovations which share these two attributes.

[4]We will use the term site to refer to that segment of the intended overall target system in which the attempt to implement the GSIA is occurring at a given point in time. Thus, "partialization" entails attempting to implement the GSIA in an increasingly larger site. In many cases the intended overall target system is not (and need not be) explicitly specified. Rather, all that is known is that the intended target system is larger than the current site. Moreover, saying that an intended target system is larger than the current site does not imply that the GSIA will necessarily be extended beyond the current site. Rather, what is implied is that the GSIA will conceivably be implemented in a larger portion of the overall intended target system if it is found to be successful in the current site.

DPMC efforts in Portugal provide a good illustration of "partialization." Initially, the GSIA is being applied within a portion of a single project (PROCALFER); it is conceivable that, if it is found to be successful, attempts will be made to extend the site to encompass the entire PROCALFER project and ultimately, perhaps, to all government interventions into the Portuguese agricultural sector.

[5]This would include responsible DPMC personnel and consultants as well as host-country personnel.

[6]The terminology used here is adapted from Mohr (1973) and Deniston et al. (1968).

[7]In some cases, of course, a given outcome may be both a precursor and a goal.

[8]Of course, in determining whether it is realistic to expect that a goal or precursor will be achieved by the time the assessment of the effectiveness of a GSIA application will take place, a reasonable margin of error should be allowed for.

[9]Downs (1981) presents a similar argument in a different context.

[10]Hall's discussion and the discussion presented here are phrased in terms of organizational effectiveness. However, both discussions apply directly to the more general concept of system effectiveness.

[11]This form of a field log is derived from Yin (1981). It involves observations that are similar to what Lin (1976) terms "structured observations." "Structured observations," as Lin describes them, however, include both pre-specified items (goals, precursors and activity sets in our case) and pre-specified response categories. The types of observations proposed here involve pre-specified items but open-ended responses.

[12]Webb et al. (1981) is the revised edition of the "grandaddy" of work in this area. Bouchard (1976 a) provides a more concise overview with some examples not contained in Webb et al. Other useful references include Webb and Weick (1979), Flaherty et al. (1978) and Sechrest (1979).

[13]The following discussion is based largely on Webb et al. (1981).

[14]Webb et al. (1981), Jick (1979), Trend (1978), Denzin (1970) and Smith (1975) are key references to this topic.

[15]A third form--theoretical triangulation--is discussed elsewhere in this report.

[16]Thus, triangulation might be referred to as a "Jack Sprat" tactic.

REFERENCES

Agarwala-Rogers, R. (1977) "Why is Evaluation Research Not Utilized?" In M. Guttentag (ed.), Evaluation Studies Review Annual, Volume 2. Beverly Hills, CA: Sage.

Allen, T. (1977) Managing the Flow of Technology. Cambridge: MIT Press.

Bacharach, S. and M. Aiken (1977) "Communication in Administrative Bureaucracies." Academy of Management Journal, 20, 365-377.

Baldridge, J.V. and R.A. Burnham (1975) "Organizational Innovation: Individual, Organizational and Environmental Impacts." Administrative Science Quarterly, 20, 165-176.

Berman, P., P.W. Greenwood, and M. McLaughlin (1975) Federal Programs Supporting Educational Change, Vol. 4: A Summary of the Findings in Review. Santa Monica, CA: The Rand Corporation.

Blandin, J. and W. Brown (1977) "Uncertainty and Management's Search for Information." IEEE Transactions on Engineering Management, 24, 114-119.

Blumstein, A. (1979) "Operations Research in the Public Sector: Frustration, Failings and Future." In K.B. Haley (ed.), Operational Research '78. Amsterdam: North Holland.

Bouchard, T.J., Jr. (1976A) "Field Research Methods: Interviewing, Questionnaires, Participant Observations, Unobtrusive Measures." In M. D. Dunnette (ed.), Handbook of Industrial and Organizational Psychology. Chicago: Rand McNally.

Bouchard, T.J., Jr. (1976B) "Unobtrusive Measures: An Inventory of Uses." Sociolocical Methods and Research, 4, 267-300.

Bozeman, B. and L. V. Blankenship (1979) "Science Information and Governmental Decision-Making." Public Administration Review, 39, 27-36.

REFERENCES - continued

Bozeman, B. and E. Cole (1981) "Scientific and Technical Information in Public Management: The Role of 'Gatekeeping' and Channel Preference." Administration and Society, 13, 479-493.

Bozeman, B. and R. McGowan (1981) "Information Acquisition and Political Environments in Technology-Intensive Government Agencies." In D. Sahal (ed.), The Transfer and Utilization of Technical Knowledge. Lexington, MA: Lexington Books.

Braskamp, L.A., R.D. Brown, and D. L. Newman (1982) "Studying Evaluation Utilization Through Simulations." Evaluation Review, 6, 114-126.

Braskamp, L.A., R.D. Brown and D.L. Newman (1978) "Credibility of a Local Education Program Evaluation Report: Author Source and Client Characteristics." American Educational Research Journal, 15, 441-450.

Brill, E.D., Jr. (1979) "The Use of Optimization Models in Public-Sector Planning." Management Science, 25, 413-422.

Brown, R.D., L.A. Braskamp, and D.L. Newman (1978) "Evaluator Credibility as a Function of Report Style." Evaluation Quarterly, 2, 331-342.

Brown, R.D. and D.L. Newman (Forthcoming) "An Investigation of the Effect of Different Data Presentation Formats and Order of Arguments in a Simulated Adversary Evaluation." Educational Evaluation and Policy Analysis.

Brown, R.D. and D.L. Newman (1980) "Decision Making Theory and Utilization of Evaluative Information: Does Sense of Control Affect Information Needs?" Paper presented at the Evaluation Network Meeting, Memphis, Tennessee.

Bunker, D.R. (1978) "Organizing Evaluation to Serve the Needs of Program Planners and Managers." Evaluation and Program Planning, 1, 129-134.

Calsyn, R.J., L.G. Tornatzky, and S. Dittmar (1977) "Incomplete Adoption of an Innovation: The Case of Goal Attainment." Evaluation, 4, 127-130.

REFERENCES - continued

Campbell, J.P., M.D. Dunnette, E.E. Lawler, III, and K.E. Weick, Jr. (1970) Managerial Behavior, Performance, and Effectiveness. New York: McGraw-Hill.

Caplan, N., A. Morrison, and R.L. Stambaugh (1975) The Use of Science Knowledge in Policy Decisions at the National Level: A Report to Respondents. Ann Arbor: Center for Research on the Utilization of Scientific Knowledge, Institute for Social Research, University of Michigan.

Carter, R. (1971) "Client Reaction to Negative Findings and the Latent Conservative Function of Evaluation Studies." American Sociologist, 6, 118-124.

Casley, D.J., and D.A. Lury (1982) Monitoring and Evaluation of Agriculture and Rural Development Projects. Baltimore: The Johns Hopkins University Press.

Clark, R.F. (1979) "The Proverbs of Evaluation: Perspectives from CSA's Experience." Public Administration Review, 39, 562-566.

Cochran, N., A.C. Gordon, and M.S. Krause (1980) "Proactive Records: Reflections on the Village Watchman." Knowledge, 2, 5-18.

Conway, R., S.G. Ballard, J.L. Benson, P.M. Burgess, T. Conrad, and T.E. James (1976) "Promoting Knowledge Utilization Through Clinically Oriented Research: The BENCHMARK Program." Policy Studies Journal, 4, 264-269.

Cook, T.D., and D.T. Campbell (1979) Quasi-Experimentation: Design and Analysis Issues For Field Settings. Chicago Rand McNally.

Cox, G.B. (1977) "Managerial Style: Implications for the Utilization of Program Evaluation Information." Evaluation Quarterly, 1, 499-509.

Crane, D. (1972) Invisible Colleges. Chicago: University of Chicago Press.

Cronbach, L.J. (1975) "Beyond the Two Disciplines of Scientific Psychology." American Psychologist, 30, 116-127.

Cronbach, L.J., S.R. Ambron, S.M. Dornbusch, R.D. Hess, R.C. Hornik, D.C. Phillips, D.F. Walker, and S.S. Weiner (1981) Toward Reform of Program Evaluation. San Francisco: Jossey-Bass.

REFERENCES - continued

David, J.L. (1978) Local Uses of Title I Evaluations. Menlo Park, CA.: SRI International.

Delbecq, A.L., A.H. Van de Ven, and D.H. Gustafson (1975) Group Techniques For Program Planning. Glenview, Il: Scott, Foresman and Company.

Denisson, O.L., I.M. Rosenstock, and V.A. Getting (1968) "Evaluation of Program Effectiveness." Public Health Reports, 83, 323-335.

Denzin, N.K. (1970) The Research Act: A Theoretical Introduction to Sociological Methods. Chicago: Aldine.

DeYoung, D.J., and R.F. Conner (1982) "Evaluator Preconceptions About Organizational Decision Making." Evaluation Review, 6, 431-440.

Downs, G.W. (1981) "Monitoring the Health Planning System: Data, Measurement and Inference Problems." In Committee on Health Planning Goals and Standards (ed.), Health Planning in the United States, Volume 2. Washington, D.C.: National Academy Press.

Elmore, R.F. (1975) "Design of the Follow Through Experiment." In A.M. Rivlin and P.M. Timpane (eds.), Planned Variation in Education. Washington, D.C.: The Brookings Institution.

Enthoven, A.C. (1975) "Ten Practical Principles for Policy and Program Analysis." In R. Zeckhauser, A.C. Harberger, R.H. Haveman, L.E. Lynn, Jr., W.A. Miskanen, and A. Williams (eds.), Benefit-Cost and Policy Analysis, 1974. Chicago: Aldine Publishing Company.

Etzioni, A. (1978) "Introduction: Policy Research." In A. Etzioni (ed.), Policy Research. Leiden, The Netherlands: E.J. Brill

Fairweather, G.W., D.H. Sanders, and L.G. Tornatzky (1974) Creating Change in Mental Health Organizations. New York: Pergamon Press.

Flaherty, E., E. Barry, and M. Swift (1978) "Use of an Unobtrusive Measure for the Evaluation of Interagency Coordination." Evaluation Quarterly, 2, 261-273.

Flanagan, J.C. (1954) "The Critical Incident Technique." Psychological Bulletin, 51, 327-358.

REFERENCES - continued

Huberman, A.M. and D.P. Crandall (1982) "Fitting Words to Numbers: Multisite/Multimethod Research in Educational Dissemination." American Behavioral Scientist, 26, 62-83.

Jick, T.D. (1979) "Mixing Qualitative and Quantitative Methods: Triangulation in Action," Administrative Science Quarterly, 24, 602-611.

Johnson, K.W. (1980) "Stimulating Evaluation Use by Integrating Academia and Practice." Knowledge, 2, 237-262.

Johnston, B.F., and W.C. Clark (1982) Redesigning Rural Development: A Strategic Perspective. Baltimore: The Johns Hopkins University Press.

Katz, D. and R.L. Kahn (1978) The Social Psychology of Organizations, 2nd ed. New York: Wiley.

Kiesler, S. and L. Sproull (1982) "Managerial Response to Changing Environments: Perspectives on Problem Sensing From Social Cognition." Administrative Science Quarterly, 27, 548-570.

Kimberly, J.R. (1981) "Managerial Innovation." In P.C. Nystrom and W.H. Starbuck (eds.), Handbook of Organizational Design, Volume 1. New York: Oxford University Press.

Larkey, P.D. (1977) "Process Models of Governmental Resource Allocation and Program Evaluation," Policy Sciences, 8, 269-301.

Larsen, J.K. and R. Agarwala-Rogers (1977) "Reinvention of Innovative Ideas: Modified? Adopted? None of the Above? Evaluation, 4, 136-140.

Lawrence, J.E.S. and T.J. Cook (1982) "Designing Program Evaluations With the Help of Stakeholders." Journal of Policy Analysis and Management, 2, 120-123.

Leviton, L.C. and F.X. Hughes (1981) "Research on the Utilization of Evaluations." Evaluation Review, 5, 525-549.

Lewis, A.C., T.L. Sadosky, and T. Connolly (1975) "The Effectiveness of Group Brainstorming in Engineering Problem Solving." IEEE Transactions in Engineering Management, EM-22, 119-124.

Lin, N. (1976) Foundations of Social Research. New York: McGraw Hill.

REFERENCES - continued

Freeman, H.E., P.H. Rossi, and S.R. Wright (1979) Evaluating Social Projects in Developing Countries. Paris: OECD.

Gertsberger, P.G., and T.J. Allen (1968) "Criteria Used by R.& D Engineers in Selection of Information Sources." Journal of Applied Psychology, 52, 110-127.

Gifford, W.E., H.R. Bobbitt, and J.W. Slocum, Jr., (1979) "Message Characteristics and Perceptions of Uncertainty by Organizational Decision Makers." Academy of Management Journal, 22, 458-481.

Glaser, E.M., and T.E. Backer (1977) "Innovation Redefined: Durability and Local Adaptation," Evaluation, 4, 131-135.

Glaser, E.M., and S.H. Taylor (1973) "Factors Influencing the Success of Applied Research." American Psychologist, 28, 140-146.

Goodman, P., and J. Pennings (1980) "Critical Issues in Assessing Organizational Effectiveness." In E.E. Lawler, D. Nadler, and C. Comman (eds.), Organizational Assessment. New York: John Wiley and Sons.

Guba, E.G., and Y.S. Lincoln (1981) Effective Evaluation. San Francisco: Jossey-Bass.

Guetzkow, H., and H. Simon (1957) "The Impact of Certain Communication Nets Upon Organization Performance in Task-Oriented Groups." Management Science, 1, 233-250.

Hall, R.H. (1982) Organizations: Structure and Process, 3rd ed. Englewood Cliffs, N.J.: Prentice-Hall.

Hannan, M.T., and J. Freeman (1977). "Obstacles to Comparative Studies." P.S. Goodman and J.M. Pennings (eds.), New Perspectives on Organizational Effectiveness. San Francisco: Jossey-Bass.

Hawkins, J.D., R.A. Roffman, and P. Osborne (1978) "Decision Makers' Judgments: The Influence of Role, Evaluative Criteria, and Information Access." Evaluation Quarterly, 2, 435-455.

Hovland, C.I., and W. Weiss (1951) "The Influence of Source Credibility on Communication Effectiveness." Public Opinion Quarterly, 15, 635-650.

REFERENCES - continued

Lindblom, C.E. and D.K. Cohen (1979) Usable Knowledge. New Haven: Yale University Press.

Litwak, E. (1961) "Models of Bureaucracy That Permit Conflict." American Journal of Sociology, 57, 173-181.

Locatis, C.N., J.K. Smith, and V.L. Blake (1980) "Effects of Evaluation Information on Decisions." Evaluation Review, 4, 809-823.

Lorenzen, G.L. and L.A. Braskamp (1978) "Comparative Influences of Political, Cost/Benefit, and Statistical Evaluation Information on Administrative Decision Making." Evaluation and Program Planning, 1, 235-239.

Louis, K.S. (1982) "Multisite/Multimethod Studies: An Introduction. American Behavioral Scientist, 26, 6-22.

Lucas, W.A. (1974) The Case Survey Method: Aggregating Case Experience. Rand Report R-1515-RC. Santa Monica, CA: The Rand Corporation.

Marquis, D.B. and T.J. Allen (1966) "Communication Patterns in Applied Technology." American Psychologist, 52, 574-579.

McConnel, W.A., T.D. Nguyen, and S. Barto (1982) "Enhancing Program Review for Evaluation and Management." Evaluation and Program Planning, 5, 37-44..

McGowan, R.P. and S. Loveless (1981) "Strategies for Information Management: The Administrator's Perspective." Public Administration Review, 41, 331-339.

Mintzberg, H. (1973) The Nature of Managerial Work. New York: Harper and Row.

Mitchell, W.D. (1979) "Finding Out What Happens in the Field: The Use of Field Logs." Evaluation Quarterly, 3, 427-445.

Mitroff, I. and T.V. Bonoma (1978) "Psychological Assumptions, Experimentation and Real World Problems: A Critique and an Alternative Approach to Evaluation." Evaluation Quarterly, 2, 235-260.

Mitroff, I.I. and J.R. Emshoff (1979) "On Strategic Assumption-Making: A Dialectical Approach to Policy and Planning." Academy of Management Review, 4, 1-12.

Mitroff, I.I., J.R. Emshoff, and R.H. Kilmann (1979) "Assumptional Analysis: A Methodology for Strategic Problem Solving." In L.-E. Datta and R. Perloff (eds.), Improving Evaluations. Beverly Hills, CA: Sage.

REFERENCES - continued

Mohr, L.B. (1973) "The Concept of Organizational Goal." American Political Science Review, 67, 470-481.

Mohr, L.B. (1982) Explaining Organizational Behavior San Francisco: Jossey-Bass.

Morell, J.A. (1979) Program Evaluation in Social Research. New York: Pergamon Press.

Newman, D.L. and R.D. Brown (1980) "An Exploratory Investigation of Decision Making Style and Locus of Control as Factors Related to Choice of School Evaluation Strategies." Paper presented at the meeting of the American Educational Research Association, Boston.

Newman, D.L., R.D. Brown, and M. Littman (1979) "Evaluator Report and Audience Characteristics Which Influence the Impact of Evaluation Reports: Does Who Says What to Whom Make a Difference?" CEDR Quarterly, 12, 14-18.

Nielsen, V.G. (1975) "Why Evaluation Does Not Improve Program Effectiveness." Policy Studies Journal, 3, 385-389.

Nisbett, R.E. and T.D. Wilson (1977) "Telling More Than We Can Know: Verbal Reports of Mental Processes." Psychological Review, 84, 231-259.

O'Reilly, C.A. (1982) "Variations in Decision Makers' Use of Information Sources: The Impact of Quality and Accessibility of Information." Academy of Management Journal, 25, 756-772.

Osborn, A.F. (1957) Applied Imagination. New York: Scribners.

Pasmore, W. and F. Friedlander (1982) "An Action-Research Program for Increasing Employee Involvement in Problem Solving." Administrative Science Quarterly, 27, 343-362.

Patton, M.Q., P.S. Grimes, K.M. Guthrie, N.J. Brennan, B.D. French, and D.A. Blyth (1977) "In Search of Impact: An Analysis of the Utilization of Federal Health Evaluation Research." In C.H. Weiss (ed.) Using Social Research in Public Policy Making. Lexington, MA.: Lexington Books.

Pelz, D. (1979) Knowledge Utilization in Criminal Justice. Ann Arbor: Institute for Social Research

Pelz, D. and F.M. Andrews (1976) Scientists in Organizations. Ann Arbor, MI.: University of Michigan Press.

REFERENCES - continued

Pondy, L.R. (1975) "Behavioral Perspectives." In L.J. Shuman, R.D. Speas, Jr., and J.P. Young (eds.), Operations Research in Health Care: A Critical Analysis. Baltimore: The Johns Hopkins University Press.

Porat, A. and J. Haas (1969) "Information Effects on Decision Making," Behavioral Science, 14, 48-104.

Porter, A.L., F.A. Rossini, S.R. Carpenter, and A.T. Roper (1980) A Guidbook for Technology Assessment and Impact Analysis. New York: North Holland.

PROPLAN (The Planning and Project Management Division, Inter-American Institute for Cooperation on Agriculture)/DPMC (The Development Project Management Center, U.S. Department of Agriculture) (1983) Development Management for Agriculture and the Rural Sector: An Emerging Conceptual Model and Promising Improvement Approach. Unpublished paper.

Reisman, A. and C.A. deKluyver (1975) "Strategies for Implementing Systems Studies." In R.L. Schultz and D.P. Slevin (eds.), Implementing Operations Research/Management Science. New York: American Elsevier.

Resnick, S.M., C. O'Reilly, and A.Majchrzak (1979) "Utilization: Different Users, Different Uses." Paper presented at the Annual Meeting of the Evaluation Research Society, Minneapolis, Minnesota.

Rice, R.E. and E.M. Rogers. (1980) "Reinvention in the Innovation Process," Knowledge, 1, 499-514.

Rich, R.F. (1975) "Selective Utilization of Social Science Related Information by Federal Policymakers." Inquiry, 13, 239-245.

Rieker, P.P. (1980) "Evaluation Research: The Design-to-Use Process." Knowledge, 2, 215-235.

Roberts, K. and C. O'Reilly (1979) "Some Correlates of Communication Roles in Organizations." Academy of Management Journal, 22, 42-57.

Rothman, J. (1980) Using Research in Organizations: A Guide to Successful Application." Beverly Hills, CA: Sage.

Rothman, J., J.L. Erlich, and J.G. Teresa (1981) Changing Organizations and Community Programs. Beverly Hills, CA: Sage.

REFERENCES - continued

Sechrest, L. (ed.) (1979) New Directions for Methodology of Behavioral Science: Unobtrusive Measurement Today. San Francisco: Jossey-Bass.

Smith, A.G. and A.E. Robbins (1982) "Structured Ethnography: The Study of Parental Involvement." American Behavioral Scientist, 26, 45-61.

Smith, H.W. (1975) Strategies of Social Research: The Methodological Imagination. Englewood Cliffs, N.J.: Prentice-Hall.

Staw, B.M. and J. Ross (1978) "Commitment to a Policy Decision: A Multitheoretical Perspective." Administrative Science Quarterly, 23, 40-64.

Stevens, W.F. and L.G. Tornatzky (1980) "The Dissemination of Evaluation." Evaluation Review, 4, 339-354.

Strong, S.R. (1978) "Social Psychological Approaches to Psychotherapy Research." In A. Bergin and S. Garfield (eds.), Handbook of Psychotherapy and Behavior Change. New York: John Wiley.

Taylor, R. and J. Utterback (1975) "A Longitudinal Study of Communications in Research: Technical and Managerial Influence." IEEE Transactions on Engineering Management, EM22, 80-87.

Thompson, P., R.D. Brown, and J. Furgason (1981) "The Impact of Evaluation Report Styles on Evaluation Audiences." Evaluation Review, 5, 269-279.

Trend, M. (1978) "On the Reconciliation of Quantitative and Qualitative Analysis." Human Organization, 37, 345-354.

Tushman, M. (1977) "Communication Across Organizational Boundaries: Special Boundary Roles in the Innovation Process." Administrative Science Quarterly, 22, 587-605.

Tushman, M.L. and E. Romanelli (1983) "Uncertainty, Social Location and Influence in Decision Making: A Sociometric Analysis." Management Science, 29, 12-23.

USAID (U.S. Agency for International Development) (1981) Management Development Strategy Paper: AID's Response to the Implementation Needs of the 1980's. Working Paper of the Management Development Working Group.

REFERENCES - continued

U.S. General Accounting Office (1977) Social Research and Development of Limited Use to National Policymakers. Washington, D.C.: U.S. General Accounting Office.

Van de Ven, A.H. and A.L. Delbecq (1971) "Nominal Versus Interacting Group Processes for Committee Decision-Making Effectiveness." Academy of Management Journal, 14, 203-212.

Walton, R.E. (1975) "The Diffusion of New York Structures: Explaining Why Success Didn't Take." Organizational Dynamics, 3,(3), 3-22.

Webb, E.J. (1970) "Unconventionality, Triangulation and Inference." In N.K. Denzin (ed.), Sociological Methods: A Sourcebook. Chicago: Aldine.

Webb, E.J., D.T. Campbell, R.D. Schwartz, L. Sechrest, and J.B. Grove. (1981) Nonreactive Measures in the Social Sciences, 2nd Edition. Boston: Houghton Mifflin.

Webb, E. and K.E. Weick (1979) "Unobtrusive Measures in Organizational Theory: A Reminder," Administrative Science Quarterly, 24, 650-659.

Weinberg, H. (1979) "The Role of Cognitive Biases in the Interpretation of Program Evaluations." Dissertation in progress. Yale University.

Weiss, C.H. (1975) "Interviewing in Evaluation Research." In E.L. Struening and M. Guttentag (eds.) Handbook of Evaluation Research, Volume 1. Beverly Hills, CA: Sage.

Weiss, J. (1982) "Coping With Complexity: An Experimental Study of Public Policy." Journal of Policy Analysis and Management, 2, 66-87.

White, K.L. and J.H. Murnaghan (1973) "Health Care Policy Formation: Analysis, Information and Research." International Journal of Health Services, 3, 81-91.

Windle, C. and P. Bates (1974) "Evaluating Program Evaluation: A Suggested Approach." In P.O. Davidson et al. (eds.) Evaluation of Behavioral Programs. Champaign, IL: Research Press.

REFERENCES - continued

Wright, P. (1974) "The Harassed Decision Maker: Time Pressures, Distractions, and the Use of Evidence." Journal of Applied Psychology, 59, 555-561.

Yin, R.K. (1982) "Studying Phenomenon and Context Across Sites." American Behavioral Scientist, 26, 84-100.

Yin, R.K. (1981) "The Case Study Crisis: Some Answers." Administrative Science Quarterly, 26, 58-65.

Zaltman, G., R. Duncan, and J. Holbeck (1973) Innovations and Organizations. New York: Wiley.

Zmud, R. (1978) "An Empirical Investigation of the Dimensionality of the Concept of Information." Decision Sciences, 9, 187-195.