

PRE- AND POST-PLANNED EVALUATION: WHICH IS PREFERABLE?

STEPHEN STRASSER

O. LYNN DENISTON

School of Public Health
University of Michigan

ABSTRACT

This paper compares the advantages and disadvantages of the "pre-planned" and "post-planned" approaches to evaluating program effectiveness. These evaluative approaches are compared along a number of dimensions which include: (a) Reliability of data and the cost of collecting it; (b) Internal validity; (c) External validity; (d) Evaluation obtrusiveness and threat; and (e) Program goal displacement and program direction. A model designed to help program managers decide when and under what conditions either of these two evaluative approaches should be employed is presented. One major theme throughout this discussion is that despite the growing interest in and use of pre-planned evaluation, the post-planned method has many advantages which often go unnoticed. This paper will help program administrators, planners and evaluators in selecting and implementing these two methods in light of their respective strengths and limitations.

INTRODUCTION

Over the past decade, the idea of developing systematic, formal evaluation plans to assess program effectiveness has gained greater acceptability. Clearly, the federal requirement that all government-sponsored programs allocate 1% of their funds toward evaluation activities substantiates this point. In certain cases, there is a requirement for an evaluation plan prior to initiation of the program; at other times there is simply a requirement that evaluation be done subsequent to program operation. Usually, the former, a *pre-planned* evaluation is deemed preferable.

This pre-planned approach to evaluation is not an unmixed blessing: problems can be created by the pre-planned methodology in addition to those it may solve. The purpose of this paper, therefore, is to identify the potential benefits and adverse effects of designing evaluation before or after a program is conducted. Following this we will present a decision-making model which should help program planners, administrators and evaluators identify those conditions under which it is best to employ either the pre- or post-planned methodology. In all, this discussion should help responsible program personnel make better decisions about which evaluation methodology to implement and minimize the problems inherent in the use of the model selected.

DEFINITION OF TERMS

Evaluation of Effectiveness Defined

We define evaluation of program effectiveness as having three parts (Deniston, Rosenstock, & Getting, 1968). The process begins with measuring the attainment of objective(s) by a given program. (Synonymous with the concept of "objective" are goal, aim, mission, purpose, outcome, etc.). The second step is an attempt to establish the degree to which the level of objective(s) attained was in fact a result of program activities. The last stage is to compare the

measured amount of objective(s) attained to what responsible program personnel intended to attain. Thus, evaluation involves measurement, the establishment of causality and comparison, and can be depicted by the following equation:

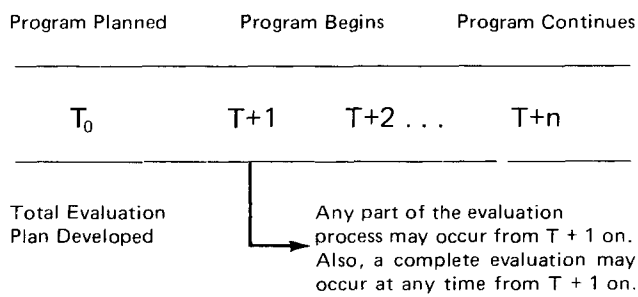
Evaluation of Program Effectiveness =
Measurement + Analysis of Causality + Comparison

At this point we should carefully distinguish between planning a program and planning an evaluation. Planning a program, which involves specifying objectives and selecting activities to attain them from the many alternatives available, must be done prior to implementation. It requires making assumptions about the future results that are uncertain; evaluation aims to confirm or deny the assumptions. Similarly, planning the evaluation must occur before the evaluation is implemented, but the actual evaluation will occur after program implementation; there is no such concept as prospective evaluation (though it has been implied [Meredith, 1976]).

Pre-Planned Evaluation Defined

Pre-planned evaluation of program effectiveness is defined as the establishment of the total evaluation plan *before* program activities begin. The plan includes specification of criteria to be measured, who gets measured and when, how measurement will occur and how program effects will be estimated. Figure 1 depicts this concept.

Figure 1. Pre-planned evaluation.



Post-Planned Evaluation Defined

Post-planned evaluation of program effectiveness is defined as the establishment of the total evaluation plan after the program has either fully completed its activities or performed a portion of them. Figure 2 depicts this con-

ceptualization of post-planned evaluation which occurs after the program has ended. Figure 3 depicts how pre- and post-planned evaluations can sequentially occur, within the context of a single program.

It is important to point out that Figure 3 suggests that while the evaluation conducted at T + 4 is *post-planned*, a subsequent evaluation repeated at, say, T + 7 which was similar in nature to that conducted earlier becomes a *pre-planned evaluation*. In essence, the process that was a post-planned evaluation of program effectiveness from T + 1 to T + 4, becomes more accurately described as pre-planned evaluation of program effectiveness when repeated from T + 5 onward.

Figure 2. Post-planned Evaluation.

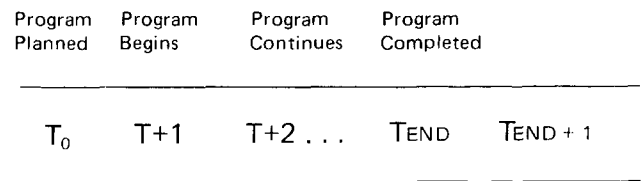
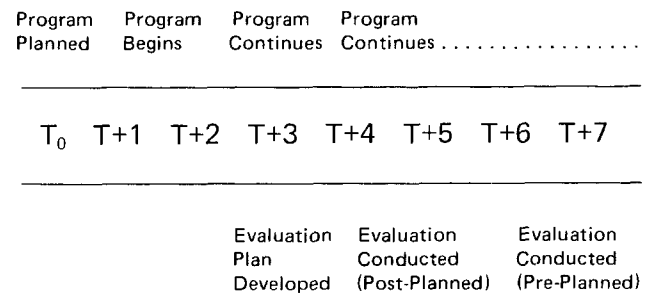


Figure 3. Combination of post- and pre-planned evaluation.



COMPARING PRE AND POST-PLANNED EVALUATION

Reliability and Cost of Data

Compared with post-planned evaluation, pre-planned evaluation facilitates the collection of more reliable data. The fact that the pre-planned method tends to be ongoing in nature allows the evaluator to collect the necessary data soon after they have been generated. As a result, the evaluator does not need to search cumbersome records or dredge inaccurate or distorted memories. In short, the evaluator — by knowing what kinds of data is needed prior to the time of its generation — is able to collect data while it is easily accessible and undistorted by time. By contrast, the post-planned method, due to its “after-the-fact” design, must frequently rely on these less reliable data collection methodologies.

While the pre-planned method may facilitate the collection of more reliable data, we suggest this method is often the more costly one to employ. This is largely the result of the continued demands that the pre-planned method

frequently makes over the life of a program in terms of repeated or ongoing instrument development, data collection, analysis and feedback. In essence, the pre-planned method may frequently take on some of the characteristics of any longitudinal study and on balance would incur larger costs than a “one shot” post-planned evaluation.

Establishing a Program’s Effect

For two reasons, the pre-planned method allows the evaluator to estimate a program’s impact with greater “internal validity” than the post-planned method. (Campbell & Stanley, 1963, p. 5). First, the pre-planned method’s “before-the-fact” and ongoing design easily generates either time-series or baseline data. With comparative data (e.g., before-after) the evaluator is able to make at least tentative statements about the causal relationships that may exist between program activities and outcomes. Although the post-planned method may still go back in time and

collect baseline data (e.g., through the use of a “retrospective pre-test” [Campbell & Stanley, 1963, p. 66]) it is, nonetheless, more difficult to accomplish and still have reliable data. Second, the pre-planned method gives the evaluator and program personnel the opportunity to randomly assign program subjects to control and treatment groups. Subsequent causal inferences can then be made more confidently than in the post-planned case, where the evaluator at best can use comparison groups to which subjects have not been randomly assigned.

These advantages of the pre-planned method are important. By knowing the strength of the causal relationships that exist between program activities and outcomes the program administrator can decide which activities should be modified because they are not generating the desired outcomes and which should be left alone because they are performing as planned. Furthermore, by having information on the causal impact of the program, the likelihood of the evaluator convincing others of the relative efficacy of the program increases dramatically.

Generalizing Program Findings

We suggest that the findings of programs which use the pre-planned approach can theoretically be as generalizable (i.e., have the same amount of external validity [Campbell & Stanley, 1963, p. 5]) as programs that rely on post-planned evaluation. However, while this is theoretically valid, two modifying conditions must be addressed.

First, findings from programs which use the pre-planned method *are not* generalizable to similar programs which use the post-planned method and vice versa. Our reasoning here, is that in pre-planned evaluations, for example, the evaluation becomes not only a part of the program’s activities but also a part of the complete fabric of program life (Weiss & Rein, 1970) since program personnel are often aware of and influenced by its presence. As a result, the organization’s performance becomes generalizable only to similar programs which also have similar pre-planned evaluation components. Clearly, generalizing the findings of a cervical cancer early detection program which used pre-planned evaluation to a similar program which will use a post-planned evaluation method would be incorrect since the nature of each program’s activities differs dramatically by virtue of the presence or absence of ongoing pre-planned evaluation. If the pre-planned evaluation can be designed in a completely unobtrusive and non-influential way (a theoretical possibility), then the findings of a program using this evaluative mode are generalizable to programs which will use the post-planned method. We argue, however, that such an occurrence is unlikely.

Second, if it is true that most *current* programs are *not* employing pre-planned evaluation and if this trend continues (though we suspect it is changing somewhat), programs using post-planned evaluations will have greater generalizability. This phenomenon, however, is a function of today’s style of evaluation and has nothing to do with the inherent characteristics of either evaluative approach.

Evaluation Obtrusiveness and Threat

A major advantage of the post-planned method is its almost guaranteed low interference with program functioning. Here, personnel are not being constantly harassed to

fill out questionnaires, consent to interviews or observation, or to be assessed in other ways while attempting to run the program. Furthermore, the post-planned design may minimize the threatening feelings program personnel experience when always under the gun of ongoing evaluation. Although one could argue that pre-planned evaluation, being ongoing, tends to socialize and desensitize program personnel to it, this may only apply to immediately successful programs.

Naturally, the potential is there for the opposite to occur in the pre-planned case. If this mode is less than thoughtfully implemented the ongoing nature of the evaluation may not only interfere with the daily routine of program functioning, but may also be seen as threatening the security of program personnel. If the threat of evaluation looms large, then we speculate that the pre-planned approach may also act to undermine the confidence of program personnel in their conduct of program activities. The fear of failure and constant second-guessing program personnel may experience can alter the performance of program activities from the way initially intended, and the result would be a program that operationally has only minimal similarity to the program that was originally planned.

Goal Displacement and Program Direction

A further problem may occur with the pre-planned approach; it has the potential of becoming the “master” rather than the “servant” of the program (Etzioni, 1964), especially when this approach is poorly implemented. For reasons related to the obtrusiveness issues discussed, it is not difficult to imagine the evaluation activities of a social service program superceding or even displacing the actual goal-directed activities of that program. Furthermore, the spontaneity and flexibility which a program administrator frequently feels must be exercised to be effective, can be stifled if the pre-planned evaluation scheme is too strictly held. However, not all effects of pre-planned evaluation are negative in the above sense. At times the pre-planned evaluation design gives the program administrator consistent direction and guidance in the operation of his program.

Though post-planned evaluation may not cause goal displacement and rigidity, it does not guide the administrator in program operation. Furthermore, this after-the-fact approach may cause the program administrators to either consciously or unconsciously encourage the evaluator to (a) select evaluation criteria on which they know they have performed well, rather than criteria which more closely reflect the objectives that the program was initially intended to attain; or (b) select comparative standards of program performance that they are certain they have already attained, rather than the *true* performance expectations program planners had established. Both occurrences are likely if there are impressions of failure of goal attainment, and in both areas would result in a spurious evaluation.

In closing this part of our discussion, we would like to point out that the advantages and disadvantages of each evaluative approach can be positively or negatively influenced by the competence of the evaluators and the situational constraints under which they are operating. For

TABLE 1
PRE-AND POST-PLANNED EVALUATION COMPARED

Area of Evaluation	Method	
	Pre-Planned	Post-Planned
1. Reliability and cost of data	More reliable, generally More costly	Less reliable, generally Less costly
2. Establishing internal validity	Causality often easier to establish	Causality often harder to establish
3. Likelihood the evaluation findings will convince others	More likely	Less likely
4. Generalizing program findings	In theory, equal external validity ^a	Greater external validity ^a
5. Evaluation obtrusiveness and threat	Generally more obtrusive and threatening	Generally less obtrusive and threatening
6. Goal displacement and program direction	Greater probability of goal displacement and program inflexibility; facilitates program direction	Less probability of goal displacement and program inflexibility; does not facilitate program direction and may encourage the use of spurious evaluation criteria (i.e., goal replacement)

^a Both models theoretically have the capacity of generating equal amounts of external validity. However, since most programs use the post-planned method, the findings of programs using this method will have greater external validity.

example, a biased or devious evaluator can take a pre-planned evaluation which has only the *potential* of threatening employee security or becoming the “master” of the program and ensure these negative outcomes occur; the incompetent or insensitive evaluator may cause or allow this; the skilled evaluator can minimize this possibility.

Therefore, we urge evaluators, organizational theorists and program administrators to consider the strengths and weaknesses of these evaluation modes in light of these kinds of mediating factors.

Table 1 summarizes the major advantages and disadvantages of both approaches.

DECISION MAKING MODEL FOR THE SELECTION OF A PRE- OR POST-PLANNED APPROACH

Having identified some of the strengths and limitations of pre- and post-planned evaluation, we now can suggest an initial decision-making model which provides planners a formal mechanism for making a choice between these two approaches.

Dimensions of the Model

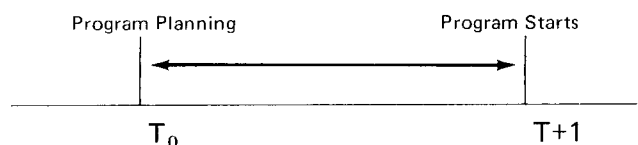
Our decision-making model is comprised of three *major sequential decisional questions* that program planners, administrators, and evaluators must consider when deciding which evaluation method to employ. These are:

1. What will the nature or composition of the program being planned look like when it becomes operational?
2. What resources will be available for evaluation?
3. What methodological procedures are necessary to generate convincing program findings (as determined through the establishment of valid measures and causal inferences)?

Implicit in these decisional questions is a time frame which suggests that the *initial* decision about which evaluation mode to use often gets made *before* the program begins. This, however, does not preclude the possibility of

initially opting for a post-planned approach and then subsequently switching gears in the middle of the program and implementing a pre-planned evaluation of subsequent program operations. Figure 4 summarizes these sequential decisional questions and the implicit time frame of our model.

Figure 4. Decisional questions and the model's time frame.



1. Decide which evaluation method to use (post or pre-planned)
2. This decision will be based on the planner's answers to the following questions:
 - a. What will the program look like when operational?
 - b. What resources are available for evaluation?
 - c. What methodological procedures are necessary to generate convincing program findings?

Decisional Question No. 1

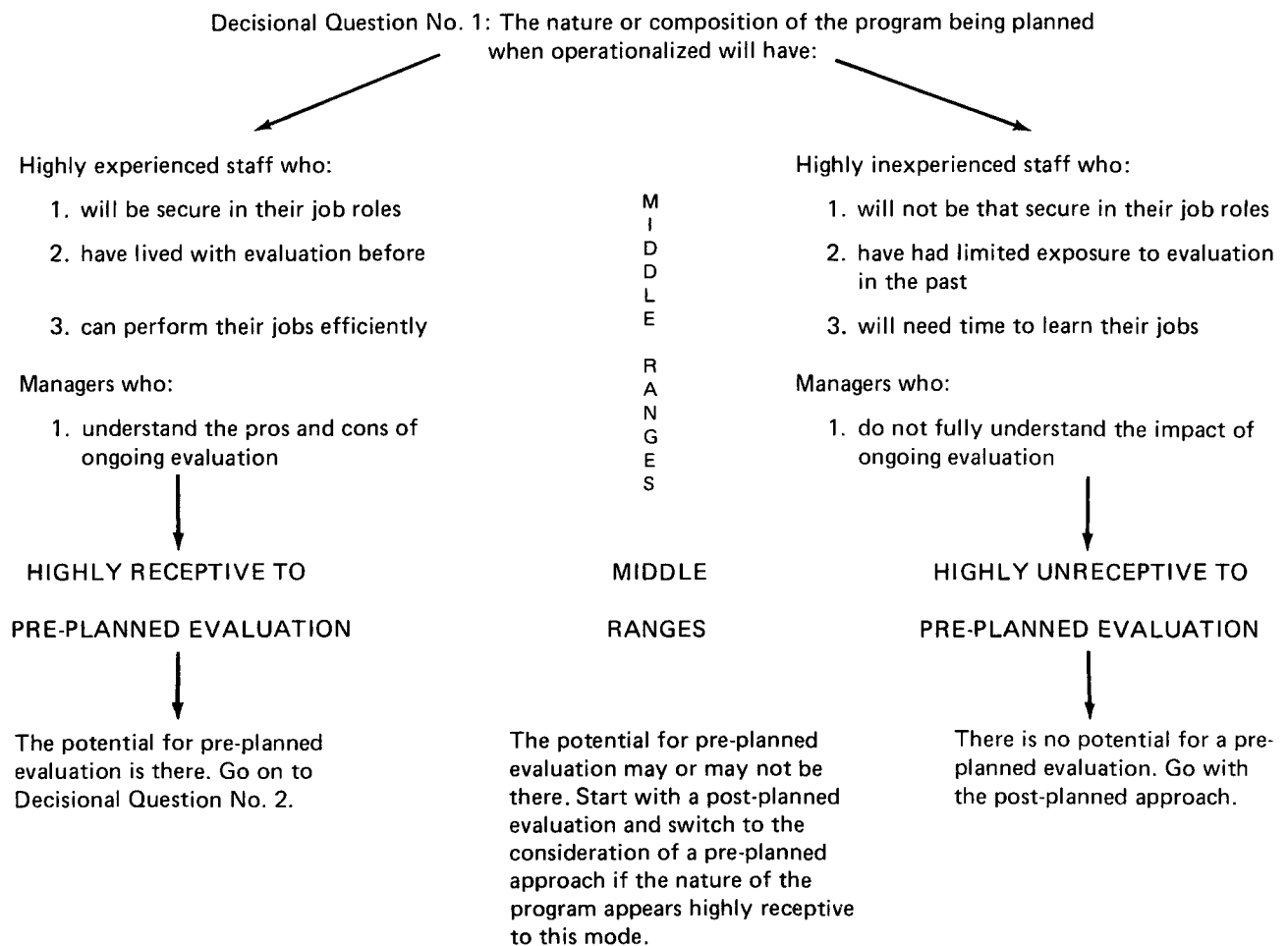
The first decisional question requires program planners to make predictions about what the actual program will look like. The planner can do this by thinking of a continuum that is bounded on one end by a program which would be *highly receptive* to a pre-planned evaluation strategy and on the other end by a program that would be *highly unreceptive* to this approach. A program characterized as being highly receptive is likely to be one that has: (a) highly experienced staff who would be relatively secure in their jobs and not threatened by ongoing evaluation; (b) staff who because of their experience are likely to efficiently perform job-related functions and thus not find ongoing evaluation obtrusive; and (c) experienced program management who understands the costs and benefits of ongoing evaluation and utilizes its feedback to facilitate rather than autocratically *enforce* changed organizational performance. A program characterized as being highly unreceptive to a pre-planned evaluation would, naturally, look just the opposite. If these highly receptive conditions to the pre-planned approach are likely to occur, the program planners still should not begin plans for a pre-planned evaluation. Instead, the planners must try to answer the second decisional question and eventually the third before making this choice. If the program is likely to be highly unreceptive to pre-planned evaluation, program planners

must decide initially to go with a post-planned assessment. If a pre-planned approach were attempted under highly unreceptive conditions, we suggest that it would result in a less than optimal evaluation. Here, staff insecurity, threat of ongoing evaluation and potential evaluation obtrusiveness are likely to compromise the reliability and validity of the pre-planned assessment. Finally, if the planners predict the future program to fall somewhere in the middle of the continuum, we urge planners to initially decide on a post-planned evaluation leaving open the option of going with a pre-planned approach if the program, when operationalized, develops highly receptive characteristics. Figure 3, presented earlier, shows how post- and pre-planned evaluation can be combined within the context of a single program. Figure 5 graphically describes the "Receptivity Continuum" pertaining to the first decisional question of our model.

Decisional Question No. 2

The second sequential decisional question requires program planners to ask what resources will be made available for the program's evaluation. As in the first question, this decisional question requires program planners to make future estimates. Here, planners must identify what the probabilities are of having (a) the money available to conduct a reliable and valid ongoing evaluation, and (b) the staff

Figure 5. Receptivity continuum.



expertise needed (i.e., evaluation skills in instrument development and data analysis) to properly conduct an ongoing, pre-planned assessment. If the program budget permits the above two conditions to be met *and* the previous condition of having a “highly receptive” program is a likelihood, then the groundwork for conducting a meaningful pre-planned evaluation has been laid. If, however, the money and/or personnel are not available, a post-planned approach should be selected, even if the first decisional question suggests a “highly receptive” program will occur. The issue here is really one of properly allocating resources. To conduct an ongoing pre-planned assessment without sufficient resources would run the risk of generating an incomplete, unreliable and invalid evaluation. It would make more sense to allocate what resources there are to the conduct of a post-planned assessment. Where the number of evaluation dollars is quite limited, the post-planned method will offer program planners an alternative to purely impressionistic evaluations.

Decisional Question No. 3

The third decisional question (What methodological steps are necessary to generate convincing findings?) is the last issue program planners need to confront in implementing our model. This is a complex issue and we shall divide it into two sub-issues.

Sub-issue No. 1 is that of (a) validity of measures used, and (b) validity of assertion about causality, linking evaluation “scores” to existence of the program. The first of these is typically labeled measurement validity, the second internal validity. Sub-Issue No. 2 is who is to be convinced? These issues are related in that there are several interested parties to any program evaluation — they may differ in the validity they ascribe to a particular measurement technique or design for estimating causal relationships.

Sub-Issue No. 1. Whether or not the selection of a measurement procedure is critical in choice of evaluation approach depends upon the nature of the phenomena of concern. Should it be a condition such as distance, mass, time, temperature, or other dimension for which measurement systems with wide acceptance exist, the choice of either method is probably of no consequence. On the other hand, if it is a condition requiring *adaptation* or *invention* of new measurement technology, such as social demeanor or quality of teaching, pre-planned evaluation is apt to be indicated. The rationale for this opinion is that time will be needed to identify, test, adapt, retest, train, etc. and the before-the-fact nature of the pre-planned approach provides the time necessary to do this adequately.

If a measurement technique is discovered, and a “score” obtained, it will usually not be interpretable. It is usually necessary to formulate an estimate of what that score would have been had there been no program — or some alternative program. The difference between the obtained “score” and the “estimate” is desired in order to judge the program. There are several ways to obtain this estimate; the degree to which the estimate is trusted will vary depending on the method by which it was obtained. The “standard of excellence” for obtaining this estimate is the randomized double blind trial. If implemented successfully, few will challenge the estimate obtained from

the “control group.” However, utilization of this method requires a pre-planned evaluation.

Other methods of obtaining this estimate include “time based” and “found groups.” The most frequently used method, the score at the time of program initiation (commonly termed a “baseline” or “before” measure) may require a pre-planned evaluation. Whether or not depends upon if there are routine measures obtained independently of the program’s operation or if they must be specially collected. If the latter, collecting the “before” data for the purpose of evaluation requires pre-planning. If the measures are made independently, this approach can be utilized just as well with a post-planned approach if the before measures are “retrievable.” The same reasoning holds for the more elaborate time based approaches, typically termed time series or “trend analysis,” which utilize a series of “before” measures. These approaches are usually applicable to post-planned use since most program operators, once a program is conceived, are not willing to wait an extended period just to amass a series of “before” measures for the purpose of evaluation.

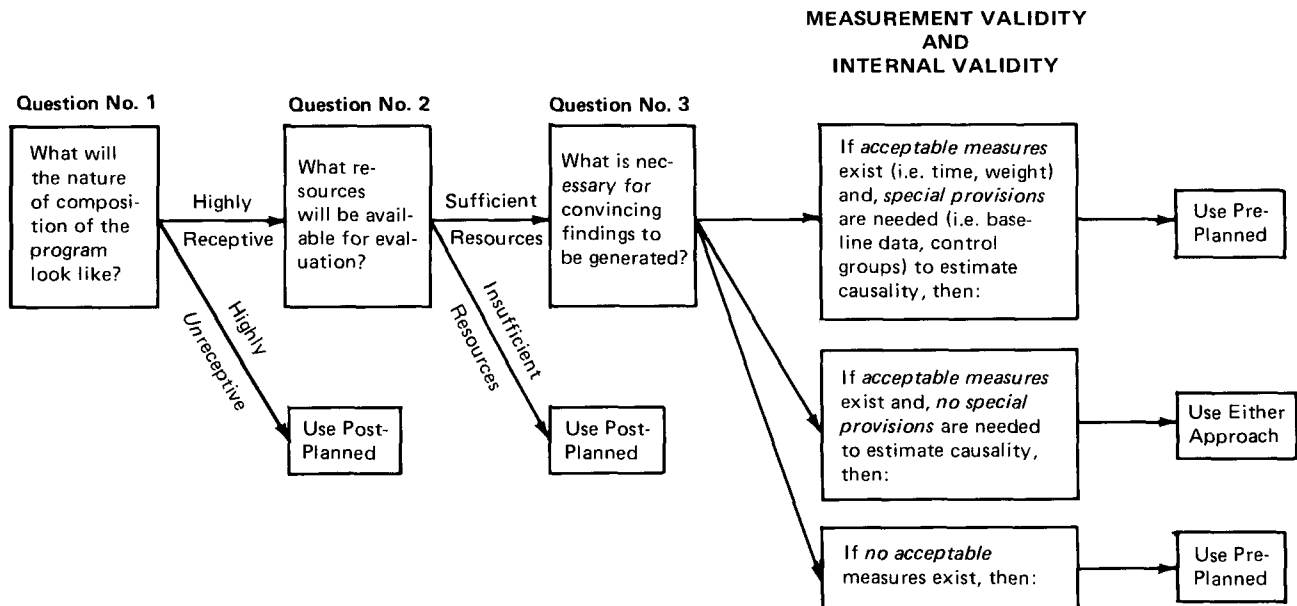
Finally, the “found” groups may be further separated into those “found by design,” such as an attempt made to find a comparison group which is similar to the program’s target population and those “found accidentally,” such as self-selected refusers of service within the intended target population.

If this general approach is selected, both evaluative methods may be applicable. However, if an attempt is to be made to find a group, the search, administrative arrangements for data collection and active data collection are going to be more timely if there is pre-planning. Even with the self-selected refuser group, data collection is apt to be more timely if pre-planned. In the rush of designing post-planned evaluation, and then implementing it, the need for collection of data from an unserved group is often overlooked until analysis starts; by then the time frame for data collection from the served and unserved groups can be quite different.

Sub-Issue No. 2. For any measurement system selected, there are apt to be doubters of the validity of the measures and for any method of estimating internal validity there are apt to be those who argue for plausible alternative explanations (Campbell, 1969). The need then, is for an evaluation technology that will yield data that are convincing to those that “matter.” For any given evaluation, this may be only internal program managers who want data to help them decide whether to alter operations or continue, only clients who want data to help decide whether to remain clients, or only “third parties,” typically bureaucratic bodies with resources to allocate to operating agencies.

Any given evaluation is undertaken to gather data that will be convincing to one or more of these groups. Since at times a pre-planned evaluation is necessary to generate data convincing to a particular audience, it is the only approach worth doing. Realization that an evaluation is desired after the start-up of a program may be too late; the best possible post-planned evaluation may be useless. On the other hand, a post-planned evaluation may very adequately provide convincing data. In essence the choice of

Figure 6. Decisional Model.



approach will depend upon the nature of the program, for this will determine the measureability of the phenomena of concern and the potential to obtain data that will be convincing to those who care.

In sum, our model offers program planners a basis on which to choose one evaluation method over another. While the decisional model works on a conceptual level,

its ultimate utility will depend on the program planner's ability to make accurate estimates or responses to the three sequential decisional questions presented. Figure 6 gives a full description of the decision model program planners can follow when choosing between a pre- and post-planned evaluation.

POST-PLANNED EVALUATION AND THE REALITIES OF EVALUATION

It should be evident by now that if a pre-planned evaluation is going to be employed, the conditions required for its effective implementation are often not going to be present. From our experience, we can think of few examples where program planners could state with confidence and honesty that the program to be developed would be both highly receptive to a pre-planned approach and have the resources necessary to properly implement it.

Those in social service organizations have a curious psyche. On an intellectual level they uniformly argue that evaluation is the greatest management tool to be developed in the last two decades. On an emotional level they are often frightened of it. Their fears are not entirely irrational, for many have experienced evaluation in its most negative form. They have been badgered by evaluators to cooperate, have watched evaluators attempt to take control of the programs in which they work (Weiss & Rein, 1970) and have seen program managers misuse the evaluator's findings. In short, people are generally not highly receptive to a process that if misimplemented may hurt them.

The realities of evaluation go beyond the extent to which an organization is receptive to it. Conduct of systematic evaluation requires resources. The amount of resources required will often vary depending upon who collects and analyzes the data. Who will be considered ac-

ceptable in turn, will depend upon who must be convinced by the findings. If program managers desire evaluation to guide their decision making, they will often trust data collected by staff although at times they may suspect that objectivity is impossible and employ others for this specific task. If "others" (funding sources, organized client groups, etc.) are seeking evaluation, they will often not trust the in-house program staff to evaluate themselves and seek external evaluation by outsiders.

We believe the costs of collecting identical data will be greater if done by outsiders than if done by program staff. If, as argued earlier, pre-planned evaluation tends to use more resources, the combination of pre-planned evaluation, done by external evaluators will often require more resources than can be devoted to evaluation. Thus, until internal evaluations earn more trust, and/or more resources are allocated to systematic evaluation, post-planned evaluations may be the only feasible systematic evaluation to conduct.

The realities of evaluation settings at this point in time suggest a general situation which is depicted by programs which have low receptivity to and inadequate resources for the *proper conduct* of pre-planned evaluation. The alternative should be a greater willingness of planners and evaluators to opt for the post-planned method. Though evaluation purists begin to frown in despair at this, the post-planned method is simply not by definition a second

best alternative, but rather another evaluation approach which can at times have value and frequently be preferable. While it may not be able to help program administrators direct or easily establish the internal validity of findings about their programs, it can offer meaningful information which is cheaper to gather, and causes less disruption of program activities. Furthermore, a skilled evaluator should be able to employ research methodologies which allow the post-planned approach to approximate some of the benefits (i.e., ability to establish internal validity or convince others of findings) which are typically attributed to the pre-planned approach. In addition, most programs today are not designed in a way which allows

the pre-planned method to capitalize on its potential strengths. For example, while pre-planned methods theoretically permit the development of experimental designs to assess a program's effect, the selection biases inherent in most social service programs (Campbell & Stanley, 1963) take away the evaluator's control over randomization and force the use of less powerful quasi-experimental designs. With this in mind, the question now becomes: Is there a distinct advantage of pre- over post-planned evaluation when the former must often rely on the same less powerful experimental designs as the latter? The answer we suggest is that the advantages are often simply not there!

DISCUSSION

This paper proposes several strengths and weaknesses of pre- and post-planned evaluation. In addition, a model for deciding which evaluation approach to use under a given set of conditions has been suggested. One of the major themes throughout this discussion was that while the pre-planned evaluation method has gained greater acceptability

over the years and does offer some unique advantages over the post-planned approach, the pre-planned method is still a less than perfect evaluative tool. Given this and the *realities* of the evaluation setting, we strongly urge program planners, evaluators, and administrators to seriously consider the post-planned method.

REFERENCES

- CAMPBELL, D. T. Reforms as experiments. *American Psychologist*, 1969, 24, 409-429.
- CAMPBELL, D. T., & STANLEY, J. D. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally, 1963.
- DENISTON, O. L., ROSENSTOCK, I. M., & GETTING, V. A. Evaluation of program effectiveness. *Public Health Reports*, 1968, 83, 323-335.
- ETZIONI, A. *Modern organizations*. Englewood Cliffs, N.J.: Prentice-Hall, 1964.
- MEREDITH, J. Program evaluation techniques in the health services. *American Journal of Public Health*, 1976, 66, 1069-1073.
- WEISS, R. S., & REIN, M. The evaluation of broad aim programs: Experimental design, its difficulties, and an alternative. *Administrative Science Quarterly*, 1970, 15(1), 97-109.