PROGRAM EVALUATION:
A Conceptual Tool Kit For Human Service Delivery Managers

(Revised July 15, 2000)
William B. McCurdy

PREFACE

This monograph was first published by the Family Service Association of America (FSAA) in 1979. (The original is now out of print.) It attempts to bring to bear on the problems of human service delivery organizations some basic systems analysis/management science concepts. It grew out of a unique effort in the late 1970s by 60 family agencies to "boot strap" themselves into what is now referred to as the "Information Age." They were the sponsors and sole funders, along with FSAA, of an organization known as IMPAC Systems Inc. Unfortunately, IMPAC came on the scene about 12 years too soon to solve the field's information systems problems.

In its 2 ½ year life, however, IMPAC made a fortuitous connection with a very unusual, very talented individual, and great teacher, named Stephen M. Drezner (formerly--and subsequently--from the RAND Corporation). Through Steve, the field had its first solid exposure to systems analysis concepts, especially in the area of program planning and evaluation. This monograph was one IMPAC product as was Stephen M. Drezner and William B. McCurdy, A Planning Guide for Voluntary Human Service Delivery Agencies (New York, NY: Family Service Association of America, 1979). The Planning Guide is also out of print.

I have revised and extended the original Program Evaluation monograph. It is available for copying, if anyone wishes, at: programevaluation.homestead.com. If you would like a fancier version of this document I would be glad to send you a zipped copy of the document in MS Word (for 97 or 2000) as an attachment. This was scanned by an updated Norton’s Anti-Virus before and after it was zipped, so I think it is safe. It is about 366K.

The monograph is available almost without cost. You are asked only to feedback your observations on, your challenges to, even your agreement with any or all of the conceptual tools covered and their presentation. If you are a manager, a board member or a staff member of a human service delivery organization we believe your most powerful tool in
program planning and evaluation is a conceptual grasp of basic planning and evaluation issues. Help us to sharpen this effort to provide you and others with these tools. Please direct your comments to me, Bill McCurdy, wbmccur@yahoo.com

Enjoy!
CHAPTER I
INTRODUCTION

For most managers in a human service delivery organization program evaluation is near the top of the list of tough, frustrating problems. Our aim in this brief monograph is to describe and illustrate some conceptual tools we think can help you grapple more effectively with program evaluation. We do not promise to simplify the problem, but we will try to strengthen your understanding of it as well as your ability to deal with it.

Here, in brief, is an outline of what follows:

CHAPTER II: AN INITIAL PERSPECTIVE ON PROGRAM EVALUATION
As the title indicates the aim of this brief chapter is to provide a beginning perspective on program evaluation. In less formal terms, we try here to “set up” program evaluation for the attacks that follow.

CHAPTER III: A USEFUL MODEL AND A CASE ILLUSTRATION
A valuable conceptual tool is the first focus of this chapter, namely a systems analysis view of a decision extended to include implementation and evaluation. The second focus is a case example that will help to illustrate the model as well as provide some “tangible” examples for use in subsequent chapters.

CHAPTER IV: THREE BASIC PROGRAM EVALUATION QUESTIONS
Program evaluation, like saintliness, is far too broad a topic to grasp in its entirety. The theme of this chapter is to partialize or, as the unsaintly might say, to divide and conquer. This chapter digs into the “how-to-do-it” of program evaluation in a conceptual sense.

CHAPTER V: ONE VIEW OF PERFORMANCE MEASUREMENT
In the field of human service delivery, many a writer has lost his readers to reverie or to anger with the very first mention of performance measurement. It is a risk we have to take. If you will bear with us, you might actually find a kinship with the concepts discussed. Our aim, as you will see, is to persuade you that performance measurement, properly viewed, can help in our field as well as in others.

CHAPTER VI: A CONTEXT FOR PROGRAM EVALUATION
We firmly believe that a conceptual understanding of program evaluation can be a very valuable tool for human service delivery managers. Our aim in this chapter is to describe a fairly general model, or view of the world that helps to put program evaluation into a larger
context. This model also gives us a way to review some of the concepts covered in previous chapters, that is, a chance to restate them in slightly different terms. We might just as well say this up front—we hope to entice you to go well beyond the bounds of this monograph to keep adding both conceptual and procedural evaluation tools to your managerial kit.

CHAPTER VII: FROM MISSION TO PROGRAM
Organizations are often exhorted to decide on their mission. Much discussion and “word smithing” may well yield a well crafted mission statement. What next, though? How do you connect the mission to specific programs?

CHAPTER VIII: CHOOSING PLANNING AND EVALUATION TECHNOLOGY
If you are persuaded by this point in the monograph to at least consider moving from conceptual tool to more tangible procedures, how do you decide what to do? We use this decision as a final example of the application of some of the concepts in earlier chapters.

SOME WORDS ON STYLE
The style of this monograph is deliberately informal. This is a conscious attempt to follow a style the writer and a colleague used in a book that approached another difficult problem, that of program planning. [See Note 1 on last page] The reasons are the same here as there: (1) the subject matter is tough enough without adding more difficulties through somber verbiage, and (2) what is covered is not intended as gospel but rather, we hope, as a set of useful tools for the reader.

The “you” in the text is the reader who, as the title indicates, we assume to be a manager in a human service delivery organization. The “we” is the writer. It is plural because, to the extent that what is presented here has utility, it comes from the input of many individuals—including many of your colleagues. Where the text falls short in this respect you are asked to substitute “I” for “we.”

Finally, if you have observed the size of this monograph, the following qualification is probably unnecessary. This document is not a complete survey of conceptual tools or conceptual issues in program evaluation. We cover what seem to us to be especially useful ideas for managers in human service delivery organizations. Moreover, we present this selection not as asserted truths, but as ideas and concepts to be tested by you for their utility to you and to your organization.
CHAPTER II
AN INITIAL PERSPECTIVE ON PROGRAM EVALUATION
(An Invitation To Test The Water)

Our prime concern is the “how” of program evaluation with an emphasis on conceptual tools. Before we address the how question, however, we need to at least touch base on some other important questions. This first tour through program evaluation is short and we urge you to join us.

What do we mean when we use the term “evaluate”? Let's start with a dictionary definition. Webster's Seventh New Collegiate Dictionary says “To determine or fix the value” of something. By this definition, then, program evaluation is the result of a process that determines the “value” of programs. This, in turn, can--and should--quickly lead to a few basic questions: When do you evaluate? Why evaluate? What do you evaluate? Value to whom? And how do you evaluate?

WHEN DO YOU EVALUATE?
We tackle the “when” question first to make what seems to us to be a useful distinction. At first consideration, program evaluation may appear to be wholly retrospective: “We have had this program for some time. What value has it had?” Yet, if you look a bit closer at your own experience, you are likely to spot how often you consider the value of a program prospectively. “If we were to offer this program in the future, what value would it have?” This prospective or future-oriented view of program evaluation is an integral part of planning and more will be said about this later.

WHY EVALUATE?
Given this view of the “when” question, we will now deal with what is clearly the most important question of all: why evaluate programs? Here it helps to separate the retrospective view from the prospective view.

Retrospective View
There are two major observable themes or rationales in retrospective program evaluation. One is accountability and the other, knowledge building. We tackle accountability first.
**Accountability:** The literature abounds with material on accountability; often the presentation is exhortative. This is especially true for external accountability where the major issue is testing past agency action against past expectations or promises: “Your agency has offered this program for a year. Did it do what you said it would? Did it achieve what we think is important?”

In recent years program evaluation to meet external accountability demands has taken on much greater significance in most human service delivery agencies. More and more what is at stake is agency credibility and, through that, continued funding and continued community and client acceptance of programs.

There are, of course, some important internal benefits from program evaluation in relation to accountability. Evaluation results, for example, can contribute to the stewardship function of a board of trustees by helping to make explicit the extent to which the implementation of plans has, in fact, conformed with the original intent of those plans. The ability to meet these and other internal accountability needs takes on greater importance as the size of an agency increases, as its programs become more complex and technical, and as the pace of change in its program and environment accelerates.

**Knowledge Building:** Although knowledge building has tended to be somewhat overshadowed by recent attention to accountability, it is a vitally important potential benefit of retrospective program evaluation. This is the contribution evaluation can make to knowledge and to improving the effectiveness of future decisions and actions.

In a systems view of the world, evaluation aimed at knowledge and skill building is identified as “feedback.” The results of some action (for example, a program) are monitored and compared to some expectation or standard. If an “error” is detected, the action or program is modified to correct the problem.

Retrospective evaluation aimed at knowledge building is, in reality, directly concerned with the future. Its intent is to identify ways to improve (or control) future actions and decisions. In contrast, evaluation which emphasizes accountability is directly concerned with the past and has only an indirect--although very important--tie to the future in the sense of establishing credibility for future support.

**Prospective Evaluation**

Prospective program evaluation, as we indicated earlier, is concerned with projecting the future value of some program or set of programs. The potential benefits of prospective evaluation lie in the chance for better choices among possible programs, which is the essence of program planning. “This agency could offer a number of different programs or variations
of programs next year. Which should we choose?” At least a part of the answer depends upon the relative “value” of the alternatives. With severely limited resources available to it, the human service delivery organization has a clear mandate to squeeze greater value from these resources.

**WHAT DO YOU EVALUATE?**

The “what” here is what it is that you determine the value of. In a general sense, the sky is the limit--actions, objects, people, ideas, decisions, even the absence of something. We speculate, however, that the major evaluation concerns of managers cluster around the following two “whats”: agency programs and agency personnel. Of the two, our concern here is with programs. We hasten to add, however, that we are using a broad definition of program, one which includes:

**Service programs** (for example, family counseling, day care, and so forth).

**Components of service programs** (curriculum development in family life education, studying prospective adoptive parents in adoption service, even services to a specific family in family counseling).

**Support programs** (for example, the enigmatic “Management and General”)
**Components of support programs** (agency planning, fund raising, information system design or operation, and so forth).

You may differentiate “projects” from “programs” if you wish, but note that we include projects as programs in our working definition. And this working definition even embraces what some would set off--not inappropriately--as “activities” such as leading a group treatment session, conducting an interview, or holding a meeting of an agency's development committee.

To say this another way, we define a “program” as “**any set of actions organized for the purpose of achieving organizational aims and, in the process, consuming organizational resources--dollars, facilities or personnel time and energies.**”

Our focus on the evaluation of programs is an attempt to narrow the problem, to make it a bit more manageable within the limits of these pages. We believe, nevertheless, that many of the conceptual tools described can be carried over to other “Whats” to evaluate, such as agency personnel. In fact, you might like to test this out as you read on.
VALUE TO WHOM?
Some very tough issues are involved in this question, whether program evaluation is retrospective or prospective. Although it is very tempting to ignore or shortcut these issues--and we often do--they really will not go away. Take this hypothetical confrontation as an example:

Agency: “Program X is far and away the most valuable program this agency offers.”

Funder: “Absolute nonsense. You're simply wasting our money on X. Program Y is clearly the most valuable program you have, and it is the only reason we give you one thin cent.”

We will be generous and assume that both the agency and the funder have complete information on both programs and that each has taken great pains to analyze them. Why, then, do they disagree?

Let us further assume that we somehow manage to get inside the heads of these two antagonists and we find something like this:

Agency: “In our head, it is vital for this agency to emphasize prevention in our services. Program X really zeros in on prevention and that is why it is very valuable to us, even more than Program Y.”

Funder: “We are not against prevention but when we think about the problems in this community, it is just crystal clear in our head that services must emphasize direct, problem solving assistance and that is why Program Y is so much more valuable than Program X.”

Of these two quite different evaluations of Program X, which one is valid? Given the assumptions we have made, we think the answer is that both evaluations could be valid. We suggest that you cannot evaluate programs without reference to what value means and to whom. This is a point that will be made in a number of different ways in this monograph. When the consequences of evaluation are important, it is crucial to get values out of people's heads and out in the open--and preferably on paper. The pragmatic result of overlooking values is an essentially futile situation for those involved, whether it be in debates such as the one just illustrated or simply in an agency's own internal efforts to “evaluate” its own programs.
HOW DO YOU EVALUATE?
The “how” question, of course, is our primary concern in this monograph. We will emphasize conceptual tools for program evaluation, that is, tools that may be most useful in gaining an understanding of the problem of program evaluation and in keeping that problem in proper perspective among other problems. Conceptual tools are especially critical for managers who will, in any event, manage program evaluation. In some instances, however, managers will also be involved in detailed procedures related to program evaluation. If so, there is a very large and rapidly growing volume of literature on specific techniques for evaluation. Skills with conceptual tools can be enormously helpful as you use your scarce time to grapple with what may appear to be a bewildering array of detailed, often conflicting, approaches to evaluation.

With this brief perspective in mind, we tackle in the next chapter a model and a case example.
We have two aims in this chapter. First, we will introduce a model that is our candidate for one of the most useful conceptual tools available for human service delivery managers and one that is useful for a lot more than program evaluation. The model is a systems analysis view of decision making extended to include implementation and evaluation.

Our second aim is to introduce a case example that we will use not only to help breathe some life into the model but also to provide some specific, “real life” examples to carry along with us as we proceed through the remaining chapters of this monograph.

You will find that this chapter is concerned almost entirely with program planning. However, don't think we have lost sight of our primary concern with program evaluation. In our view, nearly every question involved in program evaluation you face requires a look back into program planning for at least some key ingredients of the answer, and these ingredients relate principally to what value means, as we discussed earlier. In the terms we will use in this chapter and later, these ingredients relate principally to criteria for evaluation. The close linkage of program planning and program evaluation is, in fact, one of the major conceptual tools we offer for your consideration and we will try to give you opportunities to test its usefulness.

A MODEL OF DECISION MAKING EXTENDED TO INCLUDE EVALUATION

Here is one definition of decision making:

*Decision making is a process that results in choosing from a collection of known alternatives limited in number by the time and money available for the decision process and the lack of knowledge of the decision makers. Decision making means selecting a single alternative that is most likely to achieve the desired objectives.*

The decision maker seeks to make rational decisions, that is, he or she is taking into account the possible consequences of available courses of action and is considering them in the light
of what he or she considers his or her purposes. The elements of the decision-making process can be identified as follows:

1. Problem Analysis
You can view a decision as an effort to solve a problem. Decisions do not exist for their own sake or out of context. A thorough understanding of a problem is invaluable in solving it, although the tendency is often to proceed quickly to possible solutions before the problem and its various components are really understood.

2. Identification of Objectives (Decision Criteria)
Desired objectives form a basis for choosing among alternative solutions. They are what you are trying to achieve by making this decision, that is, the characteristics of a good solution to the problem. The degree to which they are satisfied should indicate how good a decision you made. Each objective should have a value (preference, weighting) placed on it that describes its importance relative to the other objectives. It is not likely that the achievement of every objective would have the same relative importance to you. As noted in the title, “decision criterion” is for us another term for objective.

3. Identification of Alternative Solutions
Sometimes only one potential solution is examined for lack of time or inclination to consider more. Many times there will be more than one alternative that needs to be considered. Sometimes the identification of alternative solutions requires a good deal of creativity. In a few instances, all possible alternative solutions can be identified. Most frequently, however, the search has to be cut off. Many decision makers fail to consider any alternatives and thus severely limit their ability to achieve their objectives.

4. Evaluation of Alternatives
This step can be the most highly analytical; it is also the most amenable to exploitation of management sciences for evaluation techniques. Here, you will note, we are referring to prospective, or “future oriented,” evaluation. One problem in evaluating alternatives is in predicting their consequences, that is: “What will happen if we do this?” The answer usually depends on many factors, many of them not under the control of the decision maker. He is dealing with the future consequences of current events and must guess or use some guessing techniques, a simulation model, for instance. Another problem is placing a value on the possible outcome of an alternative. This can be done by evaluating how well, or to what degree, the predicted result would achieve the objectives of the decision maker.

5. Choice
The decision-making process ends in a choice, a selection of one alternative solution over others. This choice could be made according to the following principle: The best alternative
is the one that yields the highest degree of achievement of objectives. There are other principles equally rational, such as: The best alternative is the one that provides the least possibility for any bad consequences to occur. In any case, it is most important that the decision maker make public whatever principle he uses to select a solution. We will say more about choice in the next chapter.

6. Implementation
With this step, we start to extend the model of decision making. The choice has been made and the alternative solution picked is now carried out.

7. Evaluation
The final step is evaluation and here, you will recognize, we are referring to retrospective evaluation. This is the step often prefaced informally by the question: “How did it go?” In the next chapter, in fact, we will cover some strategies to help answer this question.

Before we get into the “how” of retrospective evaluation, however, we want to review the seven steps above using a specific case example. The example covers a program planning and evaluation problem. We suspect you have probably already reached the conclusion that program planning is a type of decision, or more accurately, a series of decisions, that lead to the choice of which possible program will, in fact, be implemented.

A CASE EXAMPLE
To add some substance to the conceptual model of decision making just described, we turn to a service delivery agency called HFS (Hypothetical Family Service) and run through a program planning decision made by this agency.

1. Problem Analysis
For this case example, we will assume that HFS staff were examining a part of a general problem of inadequate funding for agency services. So that we can use the problem and what they did about it as a working example, we need to set some boundaries. We will, therefore, ask you to assume the following as “givens.”

Before the time chosen for this case example, HFS decision makers had determined the agency had to increase its funding base. As part of general planning for the year, they had decided to make a concerted effort to obtain more funds, principally through purchase of service arrangements. Being realistic, they allocated some resources to develop purchase contracts. Some general guidelines were established, but no detailed plans were devised. The immediate problem in our case example is how to go about the development effort. What approach should HFS follow?
In preparation for this specific problem, planners in HFS spent a good deal of time thinking about the problem (that is, they used a basic systems analysis technique sometimes called the “pondering process”). We will not go into details but we will assume that the planners dug into the problem as much as time and anxiety level permitted. The principal aim was to ferret out the characteristics of a good solution.

2. Identification of Objectives (Decision Criteria)
Using the analysis of the problem, HFS staff went on to identify the following four objectives or decision criteria:

**Criterion #1. Increase the Amount of Funds Available:** Very clearly near the heart of the problem was the amount of funds available. They concluded that if they could acquire $300,000 in new money by the end of the year it would represent about as happy an outcome as they could reasonably expect to get.

**Criterion #2. Diversify Funding Sources:** Past problems with agency dependence on a limited number of funding sources led them to this criterion. As they thought about it, a reasonable performance measure seemed to be the number of new funders acquired. Realistically, five new funders would be the most they could expect to get.

**Criterion #3. Acquire Effective and Relevant Services:** HFS was understandably concerned that the services made possible by new contract funds be consistent with what the agency viewed as effective and relevant services. In their general program planning, the decision makers had thought a great deal about this. For the problem at hand in this example they decided to seek new services that conformed as closely as possible to the agency's current view of what was effective and relevant service.

**Criterion #4. Protect Agency Autonomy:** Here planners expressed the agency's desire to acquire funders who, in the next year or so, would not lessen the agency's ability to make its own decisions. As they thought about this criterion, they concluded it would be reasonable to believe that they might conceivably achieve substantial improvement in agency autonomy.

To HFS planners, these four criteria expressed what a reasonably good solution to the problem at hand would look like. The four, however, were not judged to be equally important. The first (“Increase income”) and the third (“Effective and relevant services”) were viewed as the most important, and they decided to treat these two as equally important. The second criterion (“Diversify”) and the fourth (“Protect autonomy”) were both of about the same importance and they decided they would treat each of these as about a third as important as either #1 or #3. For convenience, they expressed these judgments in weights as follows:
1. Increase Income--Weight of 3
2. Diversify --Weight of 1
3. Effective and Relevant Services--Weight of 3
4. Protect Autonomy--Weight of 1

Please note that the weights, or ratings, that were attached to the four criteria had a special meaning. The achievement of either criterion #1 or #3 was judged to have three times the value of achieving either criterion #2 or #4.

3. Identification of Alternative Solutions
We will assume that HFS planners identified three alternative development programs, labeled Plans A, B, and C. We will take it for granted that a reasonably detailed plan was developed for each. We will not go into the details of these plans, but we ask you to assume that each plan required the same amount of resources and that this amount had already been allocated to the development program. (We make these assumptions to keep the case example from getting too elaborate for this monograph.)

4. Evaluation of Alternatives
How good were the alternative plans? To gain insights on this question, planners took each of the four decision criteria in turn and speculated how well (relatively) each of the three plans were likely to satisfy the particular criterion. A chart was prepared to keep track of this step, as illustrated in Figure 1.

As Figure 1 indicates, all three plans were judged to be close in the degree to which they would meet the “effective and relevant service” criterion (#3). Plans A and C, however, had a small edge over B. In relation to the “increase income” criterion (#1), it was clear that Plan C was out of the running. Plan B did best on money (#1) but Plan A had an edge on autonomy (#4). The income criterion had more weight than autonomy but they concluded that they would still prefer to trade off the $30,000 margin of Plan B for the gains afforded by Plan A on autonomy and its very slight edge on effectiveness and relevance. It was a close decision, but it was concluded that Plan A made the best overall contribution. 5. Choice
After evaluating the alternatives, HFS staff picked Plan A. More will be said about this choice in the next chapter.

6. Implementation
For one year, everyone in HFS worked very hard to carry out Plan A.

7. Evaluation
Now we come to the general, if informal, retrospective evaluation question: “How did it go?” We will carry this question, and our friends at HFS, into the next chapter.
Figure 1
How The Three Plans Were Judged To Contribute To The Criteria

<table>
<thead>
<tr>
<th>PLANS</th>
<th>CRITERIA</th>
<th>WEIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#1 Increase Income [Performance Measure = &quot;Amount of New Income&quot;]</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>#2 Diversify Funding Sources [Performance Measure = &quot;Number of New Sources&quot;]</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>#3 Acquire Effective and Relevant Services</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>#4 Protect Agency Autonomy</td>
<td>1</td>
</tr>
<tr>
<td>PLAN A</td>
<td>$240,000</td>
<td>Poor New Funding Sources</td>
</tr>
<tr>
<td>PLAN B</td>
<td>$270,000</td>
<td>Poor New Funding Sources</td>
</tr>
<tr>
<td>PLAN C</td>
<td>$150,000</td>
<td>Three New Funding Sources</td>
</tr>
</tbody>
</table>
CHAPTER IV
THREE BASIC PROGRAM EVALUATION QUESTIONS
(“How Did It Go?” He Asked, And I Wonder What He Meant By That.)

The premise of this chapter is that it helps to pin down the meaning of an evaluation question before you attempt to answer it. This is not a revolutionary premise, and certainly it is not one unfamiliar to human service delivery managers. Our aim, therefore, is not to sell it but rather to build on it in ways that may help you identify how you can answer program evaluation questions.

The chapter is constructed around three basic program evaluation questions and, for each, we will suggest some ways to approach an answer. We think most (perhaps all) “how did it go” questions that you will meet in program evaluation can be broken down into one or more of these three basic questions:

I. Did the program achieve what was expected?

II. Was the program carried out as planned?

III. Was it the “right” program?

We do not propose these three as a fundamental trinity in a doctrine on evaluation. However, if they provide you with a useful organizing tool, all the better. If not, we hope you can use them as a starting point for developing your own framework.

PROGRAM EVALUATION QUESTION I: DID THE PROGRAM ACHIEVE WHAT WAS EXPECTED?

This question may be the most important of all for a human service delivery organization to be able to answer for both accountability and knowledge building purposes. In this section, we will explore some approaches to an answer and some of the very tough problems involved. As a beginning point, however, we will turn to HFS staff for what might have been their answer.

HFS’s director of development put it this way:
“We had four criteria in mind when we picked Plan A. Here is how Plan A did by the end of the first year:
Criterion #1. Increase the Amount of Funds Available: We had projected that Plan A would yield $240,000 of new money. We succeeded in getting $234,000 or very close to the target amount.

Criterion #2. Diversify Funding Sources: Our projection for Plan A was four new funders. We missed out here because we were only able to get three new funders.

Criterion #3. Acquire Effective and Relevant Services: We appear to be right on target here. As we had projected, Plan A has given us contract funding for services that are in very good agreement with the agency’s view of effective and relevant services.

Criterion #4. Protect Agency Autonomy: We really lucked out on this criterion. We had thought Plan A might give us some improvement in relation to agency autonomy. It is clear now that we have actually achieved substantial improvement.

On balance, putting these results together, we feel we came very close to meeting the expectations for Plan A.”

The Critical Role of Criteria
We think you will agree that this answer would appear to be reasonable to reasonable people. But, we hope you note that the answer was totally dependent on the fact that: (1) HFS had specified the criteria it used in selecting Plan A; and (2) the agency had projected in sufficient detail how well Plan A would do in relation to each criterion.

We cannot stress this point too strongly. As you consider how your organization can get itself geared up to answer the question, did a program achieve what was expected, it is absolutely essential that you start at the planning end and make decision criteria explicit. We urge you to use the HFS example to test this, even if you find yourself in intellectual agreement with our assertion. You might, for example, assume that planners in HFS had followed the much more usual planning practice in our field and skipped directly from a review of alternatives to the choice of Plan A, with no explicit identification of criteria and no specific evaluation of alternatives in relation to these criteria. How would they then have answered this question?

The Issue of Causality
We just suggested that the answer HFS gave to the question was a reasonable one. The question, however, is complex, so this answer was very likely not a complete one. To illustrate this, let us look at another “version” of Program Evaluation Question I: “Did the implementation of the program cause the expected results?”
HFS staff carried out Plan A and when they looked at the results it was clear that these results were close to what they had projected would happen when they picked Plan A, but were these results really caused by Plan A? For example, did the activities carried out in Plan A persuade funders to contract for new services or did some other factor predominate, such as a strong, personal recommendation made by someone unbeknown to the agency?

The problem of establishing causality in human service delivery programs is an extremely complex and frustrating one. We do not intend (nor, frankly, do we have the capacity) to analyze this problem in depth here. We will, however, present a viewpoint that has helped us—and it may help you—to live with the problem.

First, we believe that for major parts of the human service delivery field we do not yet have the capacity to establish causality to the degree that the physical sciences have managed to achieve. This is a “belief statement” and we are not going to defend it further.

Our second point, however, follows closely on the heels of the first. Human service delivery managers who use gaps in technology as an excuse to avoid any systematic effort to explore causality are missing some very rich learning opportunities. For example, there is probably no way HFS staff could “prove” that Plan A caused the results, but if they were to try to explore the question of causality they might very well come up with some useful insights—practice wisdom, if you will. Very often this exploration can be done without a major investment of time and energy. For example, HFS staff could take a bit of time to discuss with the new funders what it was that seemed to influence the funding decisions. The answers might give the agency some indication of ways to design the next development effort it undertakes. To use another example, it may not “prove” that family counseling helps if an agency does a carefully structured follow-up study with clients to get their appraisal of outcome. Yet these results could be (and we think are) enormously helpful to agencies in gaining insights and suggestions for improvement of services.

Our third point is to suggest tolerance for and, whenever possible, help to those who pursue knowledge development with no “compromise” on the issue of causality. To be reminded frequently that there is no proof that a service program helps people is frustrating, perhaps even threatening. Yet, if the technological gaps in evaluation are ever to be filled, even in part, it will come from the relentless pursuit of that technology. We think tolerance is in order for some of these individuals who appear to believe that this knowledge can (and must) be gained in some absolute, objective sense. If they were to turn their demand for proof back on to their own efforts, we suspect they would eventually have to rest their own evaluation on subjective values and beliefs.
Whose Criteria?
When asked by someone outside the organization, such as a funder, there is still another
version of Program Evaluation Question I. It is often asked imprecisely, so let us try here to
state it very clearly: “Did the program achieve what I expected?”

The criteria used by the agency in planning will help answer this modified question only to
the degree that they conform to the criteria held by the questioner. More than this, answering
the question requires clarity about the weights attached to criteria and, clearly, projected
performance of the program in relation to those criteria. The questioner may or may not have
the same weights in mind or the same projected performance as the agency.

When you are asked this question you can, of course, try to evaluate the program using new
criteria, new weights, or new projections. This, however, requires the questioner to be clear
about these essential ingredients, and you will find you often have to help them be explicit.
An alternative course is to try to persuade the questioner of the merits of your agency's
criteria and weights and of the accuracy of your projections. If you do this with an open
mind, you might even end up with some desirable changes in your agency's view of criteria
or relative value.

There is, alas, one outcome that we urge you to avoid like the plague, that is, the situation
where neither you nor the other party have the vaguest idea of what anybody's criteria are or
of what weights have been attached. This is a totally “no win” situation where advantage, if
any, is gained solely by raw yelling power. Just the benefits of avoiding this one situation
may alone justify an investment by your agency in establishing the capacity to do structured
planning and evaluation.

PROGRAM EVALUATION QUESTION II: WAS THE PROGRAM
CARRIED OUT AS PLANNED?
This second program evaluation question may appear at first glance to be fairly
straightforward. There was a plan for the program. Steps have been taken to implement that
plan. Were these steps carried out as planned? There are, alas, some complications and what
we will try to do in this section is to identify them and put them into perspective and,
perhaps, even under control.

Elements of a Plan for a Program
In the HFS example we decided to simplify the case by omitting the details of the three
development plans considered by that agency. Since we cannot use HFS for this question,
we will instead suggest a general way to group the elements that should be present in a
program plan. We think these can be clustered in three categories:
1. **Input to the Program:** A program plan should specify the resources that will go into the program including staff time, money, facilities, and clients. For example: “This program should involve 200 families in the year, at least 50% of which should have a preschool age child.” Or: “This program will use 20% of the time of the CEO and 30% of the time of the director of professional services.”

2. **Actions to the Carried Out:** A plan should indicate something about the actions involved in the implementation of a program, perhaps including volume of actions, frequency of actions, scheduling considerations, and so forth. For example: “Families should receive at least one in-person interview per week.” Or: “There will be a series of perhaps ten contacts made with potential funders.” Or: “A preliminary progress report will be prepared after three months of project activity.”

3. **Decisions to be Made:** A plan should also say something about the decisions—the choices—that will be made in carrying out the program. There may be general policy guidelines or very explicit standard operating procedures. Decision making responsibility needs to be identified. For example: “Client referrals are made by practitioners who should consider the aims of the program, the preferences of the client, and current workload.” Or: “The director of the project will determine which potential funding sources will be contacted and how.”

Given this suggested grouping of elements of a plan into input, actions and decisions, let us go back to Program Evaluation Question II. Conformance to a plan would appear to be dependent on whether the specifications for input, actions, and decisions have been met.

**Criteria for Conformance to a Plan**

In the last paragraph we used the term “specifications” and we suggest that it is analogous to the term “criteria” used earlier. To judge conformance to a plan requires criteria for conformance. Here we can draw on some points made earlier in this and the previous chapter.

First, you will find it difficult to answer Program Evaluation Question II if, in fact, you have not invested some effort in spelling out a plan for the program, that is, input, actions, and decisions. This is directly analogous to trying to evaluate program achievement without the criteria for achievement.

Second, it is extremely unlikely that all of the criteria (all of the specifications) that you do develop for conformance to a plan will be equally important. For example, assume that a plan for some service project specifies these two things: (1) “A progress report should be prepared for the board every three months” and (2) “Services provided to clients should
conform to accepted professional standards.” We would find it very hard to conceive of a project in which the first would be more important than the second.

There is a great deal of practical value in sorting out criteria for conformance to a plan by their relative value. There is a tendency in some organizations to try to measure and monitor everything about program implementation in order to be able to assess how well that program is going. Often, this is compounded by measuring everything, without considering the criteria to which these measures might relate. Perhaps just as problematic is another tendency, which is to measure and monitor only what seems easiest or traditional and let it go at that. All too often the result is a considerable investment of resources in collecting information which, when it comes time to answer Program Evaluation Question II, seems to fall short of providing reasonable answers.

We believe it is preferable to first sort out the elements of a plan by their relative importance to the eventual success of that plan. With this established, there is a much greater likelihood that an investment made in assembling information about performance will yield useful answers to this question. In short, do not judge your agency's capacity to answer Program Evaluation Question II solely by the amount of information you collect. Base your judgment principally on how well you have thought through criteria for conformance.

**Accountability vs. Knowledge Building**

Earlier we suggested that retrospective evaluation could be aimed at either the knowledge building theme or the accountability theme. As you have probably observed, Program Evaluation Question II has relevance for both themes. When an organization assesses conformance to a plan with a knowledge building theme in mind, it is seeking to improve future actions in that or similar programs. This provides feedback, and spotting nonconformance leads to steps to get a program back on the track or, equally important, to doing a better program design the next time the program is delivered.

The accountability theme in Program Evaluation Question II is a familiar one to human service delivery organizations with external funders. A common example is the financial report required to demonstrate adherence to an approved budget (that is, adherence to some criteria on input to the program), and here an organization can run into trouble when a funder does not recognize that some specifications for a program may be more important than others. For example, if a funder attaches the same importance to conformance to a budgeted line item of expense (for example, supplies) as it does to effective utilization of staff time in service delivery, both the funder and the agency are losers.
PROGRAM EVALUATION QUESTION III: WAS IT THE “RIGHT” PROGRAM?
The third program evaluation question may be the most difficult of all. A program could have achieved all that was expected of it. It could have been implemented precisely as planned. And still that program could leave you—and others—with questions about its “rightness.” Should it have been implemented at all?

We suggest that Program Evaluation Question III can be boiled down to the issue of evaluating the planning process that led to the choice of the program. If the question is recast in terms of the planning process, it is then possible to lean on the model of a decision in Chapter III to partialize the question into more manageable components. We propose to tackle this question using four, more specific questions based on this model.

Program Evaluation Question III-A: How Adequate Was the Identification of Criteria and the Weights Attached?
If you think about Program Evaluation Question I, covered earlier in this chapter, you will see that the validity of an answer by our case example agency HFS would depend heavily on the adequacy of the criteria used. Component question III-A, therefore, is very important although it is, alas, extremely hard to answer.

As you will remember, criteria are derived from an understanding of (or point of view on) a problem. Was HFS's understanding “correct?” Criteria, and especially the weightings, are rooted in the values and informed opinions of decision makers, and we suggest you cannot “prove” that criteria, or their weightings, are either true or false. They are opinions that exist in the minds of decision makers: they are subject to analysis and argument, but not to “scientific proof.”

Although tough to answer, Program Evaluation Question III-A is still very important. We will give an example from HFS's experience, an example that illustrates one useful strategy for tackling the question. At the end of the year, HFS took another careful look at the problem and its experience with Plan A and reported this finding:

“What we discovered, to our dismay, is that one of the new funding sources—the one giving us the most money—is actually paying us something less than the comprehensive cost of the service we provide them. We are not forced to provide an out-of-pocket subsidy for these services but we should have had a criterion that called for achieving close to comprehensive cost. In fact, if we had identified such a criterion and if we had given it adequate weight, we probably would have picked Plan B rather than Plan A.”
The strategy HFS used to answer Program Evaluation Question III-A was a conscious search for unanticipated outcomes. The result was the identification of one important aspect of the problem that had not been reflected in decision criteria. Needless to say, we can be sure it would be included in future HFS planning!

First of all, the answer to Question III-A requires a judgment on the completeness of the set of decision criteria identified. Do these criteria reflect the important considerations that should go into the evaluation of alternatives? The answer also requires a judgment on the appropriateness of the weights attached to these criteria. This argument, in turn, requires a great deal of information about the problem and the organization. Both judgments are useful to make as long as you are prepared to accept that they are just that and no more.

**Program Evaluation Question III-B. How Adequate was the Set of Alternative Solutions Identified?**

This is the question “answered” by the person who might challenge HFS with a statement like this: “None of the three plans you considered was any good! Why didn't you think of Plan D?” The challenge is that HFS did not have an adequate view of alternative solutions. If this challenge is examined it might be found that HFS staff did not have enough time to invest in a search for more alternatives. They might not have had access to pertinent technological information (to know, for example, that some agency had devised and tested another approach). They might have failed to use available technology. Finally, they might not have been creative enough to think of another alternative.

Here, then, are clues to judgments that can help analyze Program Evaluation Question III--B. Was enough time spent in devising alternative solutions? Was pertinent knowledge available? Was available knowledge utilized? Was planning sufficiently creative? These, of course, are difficult judgments to make, but there is a good deal of value in trying. The potential payoff lies in “better” future decisions.

There is another aspect to this question that relates to the description of alternative solutions. All too often alternatives are rejected (when they are good ones) or accepted (when they are not) because of an inadequate, usually incomplete description. If you examine the model of a decision you can get some insights into appraising the adequacy of a description of an alternative. That description should be sufficient to enable alternative solutions to be evaluated. In other words, the description of an alternative solution should contain enough information to enable informed decision makers to speculate on how well that alternative would meet decision criteria. To go back to HFS briefly, one key criterion dealt with increasing the amount of funds available. To judge Plans A, B, and C on this criterion, HFS needed to understand these three plans well enough to be able to speculate on how successful each might be. If one plan had simply said: “We will seek contract service funding from a
public agency” it would be very difficult to speculate on its probably success. However, if a plan of action had been identified (including ideas on specific public agency programs, information on how other agencies had fared, what steps would be taken, and the like) the task of predicting results would be easier.

Program Evaluation Question III-C. Were Alternative Solutions Adequately Evaluated?

This particular component question is targeted on the step in the planning process in which alternative solutions were evaluated against decision criteria. This was future-oriented evaluation, involving judgments (estimates, speculations, guesses, and so forth) about how well the alternatives would meet the criteria. The issue here is the effectiveness of that future-oriented evaluation.

HFS evaluated three plans. Plan A came in first, with Plan B a close second. At least, that was the result of the evaluation step. The edge that Plan A had over B might have been due to errors in projecting the extent to which these two alternatives would meet the criteria. Perhaps Plan B would have resulted in substantial improvement in agency autonomy. Perhaps even Plan C would have yielded substantially more than $150,000 of new income.

It is very difficult to appraise the future oriented evaluation step. It is, in essence, a double crystal ball question. The original projection involved a high degree of judgment, and now another judgment is required on the adequacy of that original judgment.

One approach to an answer, however, is to recognize that several specific judgments are involved in future-oriented evaluation. One is speculating about the effects of a well-implemented program. Another is speculating on how well the alternative would, in fact, be implemented. Still a third is speculating on what other factors could influence the results one way or another. For these judgments it is critical that the alternative be clearly described and that criteria be clear and understood, but it is also necessary to draw heavily on experience, practice wisdom, research results (if available), and a broad understanding of the problem and the environment.

Hindsight will be of some help in answering this component of Program Evaluation Question III-C. if you took time to answer Program Evaluation Questions I and II. HFS knows how Plan A did; still in doubt, however, is how well Plans B and C would have done had they been selected.

Because you will be hard-pressed to go beyond some very rough judgments, why bother to tackle this component? We suggest again that the potential payoff lies in the insights you may get into ways to make better, future-oriented evaluations the next time around.
Program Evaluation Question III-D. Was the “Right” Choice of Alternatives Made?
When we described the planning process HFS used, we passed rather hastily over the “choice” step but with a promise to say more later. We will now try to fulfill that promise. The choice step is usually a very significant one.

We again turn to HFS for help. Agency staff now tell us they made a crucial assumption when the three plans were evaluated (the step preceding the choice). They assumed that their very able director of development would stay with the agency throughout the year. In effect, they judged the risk of losing this person low enough to ignore. We have it on good authority, however, that if staff had gone through the future-oriented evaluation step and assumed that this person would not stay the year, Plan C would have had a very strong edge because it was not as dependent on this person's expertise.

To generalize a bit, a projection on how well alternative solutions will meet decision criteria typically involves assumptions about factors outside the control of the agency that could have a bearing on the outcome. These are sometimes termed assumptions on “states of the world.”

Going back to HFS, we now ask you to assume that there was, in fact, a fairly high risk that the director of development might leave by mid-year—not a certainty, but a risk. With this assumption, should staff have selected Plan A?

We suggest the answer would depend not only on how big the risk was but also on how they chose to treat that risk. They could have decided to ignore it, or they could have opted to “gamble,” that is, to go for the higher payoff of Plan A despite the risk that the director might leave. Conversely, they could have decided to “play it safe,” to avoid risk consciously and pick Plan C. Or, they could have decided to “play the odds,” that is, to be guided by how probable it was that this person might leave. These possible bases for choice are illustrations of what are technically termed, “decision rules.” The term may not be familiar, but we think you will recognize the concept as an all-too-familiar part of your daily life as administrators (as well as the agenda for many sleepless nights).

With this somewhat enlarged view of the choice step, we return to Program Evaluation Question III-D. A “right” choice would be one that forecast significant risks (that is, states of the world) and one that applied the appropriate decision rule. On the former it is grossly unfair to make a judgment based solely on hindsight. The point here for HFS is not whether the director of development did, in fact, resign but whether the risk of his leaving was properly gauged when the choice of alternatives was made.
To judge the appropriateness of a decision rule requires an evaluator to do some very heavy “second guessing” of the decision maker. If a gamble was taken, was it justified? If risk was avoided, was it warranted given the general situation the agency faced at the time of the decision? Again, it is grossly unfair to answer this part of Program Evaluation Question III-D on hindsight only (that is, an answer based on whether the risk avoided did or did not materialize or whether the gamble taken did or did not succeed). Unfortunately the unfairness of such hindsight judgments does not stop everybody from making them, and we are sure you do not need to be reminded of this painful reality.

As was true for the other three components of Program Evaluation Question III, firm answers are not in the cards. Partial answers, however, can help. In fact, if an agency simply gets insights into ways to deal with the uncertainty of its environment we think it is ahead of the game.

A SUMMARY COMMENT
We used a fairly narrowly defined illustrative program in this chapter--a financial development program--to help delineate three basic program evaluation questions. What we will do now to end the chapter is to put that example into a broader context.

You will remember that HFS had gone through an agency-wide planning process before tackling the specific problem in our case example. That planning process yielded a decision to allocate some resources to relieve the financial plight of the agency. In the case example, the task faced was selecting and implementing a specific development program, and it was to this program that the three program evaluation questions were addressed, questions derived from the more general question of: “How did it go?”

Now, assume someone steps back and, thinking of HFS's total program for the year and its total planning process, asks: “How did it go?” We suggest you could identify the three basic program evaluation questions in this more general question. The difference is that these questions would now center on total agency program.

How would you go about answering these more general questions? If we have been reasonably clear in this chapter we will have made this point: You cannot deal effectively with program evaluation unless you have already dealt with program planning. If someone promises you a painless shortcut, please be very skeptical.

The other side of this assertion is that there is a lot you can do to help answer agency evaluation questions, even with a modest investment, if you have, in fact, dealt with agency planning issues. In our view some voluntary agencies will not survive the “age of accountability.” Ironically, for many of these agencies their demise will not be attributable to
inadequate evaluation technology, but rather to their failure to utilize readily available planning technology.
CHAPTER V
ONE VIEW OF PERFORMANCE MEASUREMENT
(Tail-Wagging, And The Viability, If Not Total Desirability, Of Tailless Dogs)

Performance measurement plays an important role in evaluation, perhaps even a leading role. It has a tendency, however, to be a scene-stealer, to so dominate the stage that other key performers are badly slighted or alienated. This tendency is especially unfortunate for the human service delivery organization because measurement per se is not part of the daily professional practice of the great majority of personnel in the field. Our aim in this chapter is to put performance measurement into perspective, and our hope is to persuade you that performance measurement, properly viewed, can help in our field as in others.

SOME EXTREMES
To get into the topic of performance measurement, we want to start by examining what we observe as two extreme viewpoints on the role of performance measurement in evaluation. With some exaggeration to make distinctions sharper, here is how we see it:

The “Hard-Facts” School
At one end of the continuum are those who view performance measurement using “hard” objective measures as the only solution to the evaluation problem. Objectivity, in fact, is the name of the game and the target is an evaluation conclusion uncontaminated by human judgment, human biases and self-interest. Often, their advice is: “If you can't measure goal achievement, drop the goal! It has no meaning.”

In fairness, some spokesmen for this school seem to be trying to counteract what they perceive as blind opposition to measurement: “The only way you can get those people to measure anything is to give them no choice.”

We see several problems with the hard-facts school. First, many real-life problems in human service delivery agencies (and even, perhaps, in the corporate world), involve important goals (or criteria) that cannot be measured using “objective measuring scales. The obstacle to measurement may be insufficient time or money to design, build, and collect performance measures, or there may be gaps in present technology that make measurement infeasible (we
suspect that some of these gaps may never be filled). Given these obstacles, should the criteria involved be ignored? In the case example in Chapter III, HFS was concerned about further erosion of its autonomy. Should the agency ignore that concern in planning simply because they could not identify a hard performance measure? The plight of voluntary agencies who have done so speaks all too clearly to the risks.

A second problem in the hard-facts view lies in the issue of objectivity. The “hardest” of measures, the one easiest to replicate, may itself be objective. The tough issue, however, is not the objectivity of the measure but the meaning of the measure. The temperature of a room can be measured with great precision and objectivity. But where is the objective basis for saying 75 degrees (or 80 or 85) is “too hot”?

In the case example, HFS found a very objective measure for its first criterion, “increase income.” We can be quite sure new money received would be measured to the exact penny, with no significant subjectivity. But where is the objective guarantee that $300,000 of new money would be a great result for HFS to achieve? And what is the objective basis for saying that the $240,000 predicted for Plan A was “satisfactory”?

Finally, a third problem is the high risk of “tail-wagging” when you move close to the extreme hard-facts school. We will develop this point at the end of the chapter but, briefly, we mean the risk of letting hard performance measures crowd out other concerns, even the criteria that give these measures their meaning.

The “Total Avoidance” School
At the other end of the continuum are those who see performance measures as providing no solution whatsoever to the evaluation problem, at least not in human service delivery agencies. Often this school views measurement as impotent (“You can't measure anything of significance in our field”). Sometimes, measurement is seen as antithetical to the values of human services, perhaps even dangerous to the complex process of individualizing the client in services.

Again, in fairness, we think some members of this school are rebelling against what they perceive to be stubborn narrow-mindedness in the hard-facts school. “If we give them an inch, they'll take a mile.” (We hope members of the total avoidance school will forgive this analogy.)

The major problem with the total avoidance view is one of lost opportunity. We suggest that performance measurement has a potential to help the human service delivery agency not only in evaluation but also in planning. In the instance of HFS we think the fact that two of the four criteria lent themselves easily to relatively objective performance measurement helped
that agency deal with a complex planning decision and with the evaluation of results achieved. The agency's attitude helped, of course, because it refrained from downgrading the other two criteria to second-class status simply because there were no hard performance measures for them.

Our view, as you have no doubt gathered, is middle of the road, somewhere between the two extremes. To present this view, we will begin by exploring what a performance measure does.

WHAT DOES A PERFORMANCE MEASURE ACTUALLY DO?
As one writer put it, a performance measure serves as “...a bridge.. linking the subjective inner mind...of decision makers to the objective outer world of physical alternatives.” [See Note 2 on last page] In the case example, the decision makers in HFS identified a decision criterion concerned with increasing agency income. The performance measure--new dollars of income--helped the agency translate what that criterion meant to it into an attribute of specific, alternative financial development plans. It was an attribute it could speculate about when considering the choice among Plans A, B, and C, and it was an attribute it could measure when the plan actually implemented was evaluated.

HFS had another criterion lending itself fairly readily to measurement, the one labeled “diversify funding sources” and we ask you to assume that HFS planners decided to be especially clear in their thinking about a performance measure for this criterion.

Their first step was to pick a performance measure. An attribute of a development plan that seems to have some reasonable bearing on the issue of diversification is the number of new funding sources that would result if that plan were implemented. Hence, this appears to be a useful performance measure. We are not sure it is the only performance measure and, if there are others, the choice would be a judgment and not foreordained.

Given a performance measure, the next step is to pin down the boundaries of that measure. HFS concluded that the most new funding sources that it could reasonably hope to get was five. Actual performance of a specific financial development plan, then, might result in some number of new funders ranging from zero to five. Given the nature of the criterion it is fair to say that no new funding sources would be a very unhappy result. But how about the possible results of one, two, three, or four new funders? What value should be attached to those results?

This leads to an important third step, which is to pin down the meaning or value of all possible results. This relationship is not obvious. For HFS it would depend on the way the
agency viewed the criterion and that, in turn, would depend on how the agency viewed the problem that led to a development effort.

Assume that HFS staff came up with the following judgments. They said five new funders would mean full satisfaction of this criterion. (Please remember we are relating only to the “diversify” criterion.) Four new funders, while not the best, would still not be too bad and would represent 80% achievement. Three new funders, however, would really fall short of expectations and would represent only 50% achievement. If a development plan yielded only one or two new funders, it would make only a very limited contribution, say 10%. Zero new funders, as noted, would be absolute disaster.

We think these judgments gave HFS staff a good handle on the performance measure and the criterion. They could readily communicate what they meant by “diversify.” If they could project the number of new funders that various alternative plans would yield) they could do a reasonably good job of future-oriented evaluation of those plans in relation to this criterion. And when the plan they selected had been implemented, they could add up the number of new funders and evaluate performance against the “diversify” criterion.

But, please be sure to note that the meaning of a result using this very objective measure (number of new funders) depends on the subjective judgments of the decision makers when they set up the relationship between measurement results and the criterion. That relationship is not self-evident.

To reinforce this last idea, here is a slightly different scenario. HFS might have concluded that five new funders would represent “too much of a good thing” and that, instead, four new funders would represent full achievement. Under these conditions getting five new funders would mean a lower contribution to the criterion than would four new funders.

We want to alert you to the fact that what HFS just did for the performance measure of number of new funders is often described as “setting standards” for the measure. You may run into this label in the literature.

DIRECT WORTH ASSESSMENT--AN ALTERNATIVE TO PERFORMANCE MEASUREMENT

What do you do if you cannot readily identify a performance measure? Our view (and not everyone will agree) is that all is not lost. There is an alternative, which we will label “direct worth assessment.”
To describe direct worth assessment it helps to have a view of what happens when you do, in fact, have a performance measure, which is why we took some time in the previous section to analyze the use of performance measures. Direct worth assessment involves the same judgments, only now without the assistance of a readily measurable attribute of alternative solutions.

A first key judgment rests in the decision criterion and, through it, in the underlying problem. In our case example, HFS staff were concerned about possible loss in agency autonomy. As they sought to enlarge the agency's funding base, they wanted to avoid, if possible, further erosion of autonomy. Examples would include funders who require conformance to certain practices as a condition for funding. In some instances, these requirements may extend beyond the service involved as, for example, a funder who might require the agency hold all of its salaries to community established maximums or a funder who retained veto power over other services the agency might want to add or drop.

Given this understanding of the problem underlying the criterion on autonomy, HFS staff would then need to clarify what reasonably complete satisfaction of the criterion would look like. From what they have said, it seems they felt it conceivable that the agency's overall situation could be improved substantially. It was with this outcome in mind that they established the relative weight they attached to this criterion vis-a-vis the other three criteria (see Chapter III).

As HFS looked at the three plans, a third judgment would be to speculate about what each plan might do to improve or diminish autonomy. As Figure 1 in Chapter III notes, staff concluded that Plan C, if implemented, would likely result in fairly substantial improvement in autonomy. Plan A, if implemented, might leave the agency a bit better off. They guessed that Plan B, however, would probably result in essentially no change. Finally, we know that the implementation of Plan A actually yielded substantial improvement in the degree of decision making authority retained by the board of HFS.

As you will see, these are the same judgments that would have to be made if HFS had a performance measure for autonomy although, without that measure, the judgments are tougher. Although everybody is likely to agree if “four new funders” are actually acquired or if “$240,000 of new income” is obtained, there could well be disagreement about a result described as “we held our own on autonomy” or “we actually improved autonomy substantially.”

Here, then, is one of the problems of direct worth assessment. It will usually be harder to get agreement on how performance relates to the decision criterion. To offset this problem HFS staff could take time to discuss what is involved. What are examples of threats to agency
autonomy? Which illustrations are really serious? How prevalent are these threats now? What are some examples of reasonable improvement?

Discussions like these might help decision makers get a shared view of what they meant by the criterion. However, the agency could go one step further and set up a simple rating scale ranging from “no achievement” of the criterion to “full achievement.” Important points on that scale could be pinned down (or anchored) with some examples. Here is one sample:

Full Achievement: (1.0) “We would find ourselves with fewer significant constraints than we have now. For example, we could set salaries for some key positions without outside approval.”
Fifty Percent Achievement: (0.5): “We would not be able to identify a significant example of improvement or deterioration or we would find any changes would really seem to offset themselves.”
No Achievement : (0.0) “We would find ourselves saddled with more constraints than we have now, for example, more mandated procedures or restricted actions.”

Decision makers using such a scale could project the likely results of a plan and perhaps reach agreement that these results would be “about 0.6 on the scale.” Or they might take a look at what happened when a plan was implemented and agree that “we got very close to 1.0 on the scale.”

Is such a scale a performance measure? It suffers from some problems of reliability (that is, two people might not come up with the same number given the same evidence). Yet, it is probably more reliable than having people, on their own and with no guides, come up with conclusions about criterion satisfaction. If you accept that measures can differ in reliability then this scale does seem to qualify as a performance measure. There is, in fact, good reason to view all performance measures as lying on a continuum ranging from very objective, highly reliable measures to very subjective, hard-to-replicate measures. If you take this view, then direct worth assessment can be seen conceptually as a form of performance measurement, albeit with less objective measures involved.

**HOW HFS COULD HAVE STRUCTURED ITS ANSWER TO PROGRAM EVALUATION QUESTION I**

Having dealt with performance measurement and direct worth assessment as conceptual tools, we would now like to put these together in a specific illustration. In Chapter III we described how HFS approached the problem of deciding on a financial development program. We said that HFS staff identified and weighted four criteria, evaluated (future-
oriented) three plans (A, B, and C), selected Plan A, and carried out Plan A with certain results. As you will remember, HFS had readily available performance measures for two criteria and earlier in this chapter we suggested steps that need to be taken to attach meaning to the results of measurement (that is, to set standards). The remaining two criteria did not have readily available performance measures but, as we just suggested, relatively simple rating scales could help achieve more consistent evaluation.

Here is an example of how HFS staff could have structured the planning decision and how they could have used that structure to help them answer Program Evaluation Question I: *Did Plan A achieve what was expected?* We believe that these steps could have taken an insignificant investment of time on their part.

Figures 2 through 5 give a reasonably complete view of how HFS staff could have set standards for the measures used for each of the four criteria. The last two involve rating scales.

![Figure 2: Criterion 1: Increase the Amount of Funds Available](image1)

![Figure 3: Criterion 2: Diversity Funding Sources](image2)
Figure 4
Criterion 3: Acquire Effective and Relevant Services
To be judged by the extent to which services acquired conform with the agency's current view of what is effective and relevant.

**Extent to Which Criterion Met**

Plan A and Plan C
Judged to be 0.9

Plan B
Judged
To be 0.8

- **1.0**
  - Excellent agreement with our view. Couldn’t be better. We would give priority to these services using our own resources.

- **0.5**
  - Fair agreement with our view. We would still offer these services using our own resources but probably not much.

- **0.0**
  - Poor agreement with our view. We would probably not offer these services using our own resources.
In Figure 6 we illustrate how they could have dealt with the fact that the criteria were not of equal value to them when they evaluated the three alternative plans. As you will see, a relative priority was computed for the three plans. To do this, HFS multiplied the predicted contribution times the weight for the criteria and added across to get the priority for each plan (the “weighted sum of contributions”). The results, in the far right column, give a slight edge to Plan A over Plan B (6.6 compared to 6.4). This result seemed to make sense to HFS. If the numerical results had not felt right to the decision makers, they should have backtracked to find where the problem lay. It could have been in the judged contribution of alternatives to the criteria or in the weight assigned to the criteria or both. To assume that a number “has to be right” because it is hard data is to ignore the judgments and subjectivity behind that number.
At the end of the year, HFS staff asked Program Evaluation Question II: *Did Plan A achieve what was expected?* Using the groundwork laid in Figures 2 through 6, HFS developed the answer shown in Figure 7. The weighted sum of contributions helped to deal with the fact that Plan A did poorly on criterion 3 but unexpectedly well on criterion 4. In other words, it gave the agency a handle on a summary evaluation of Plan A.
[If you are interested in pursuing this approach to establishing program priority, see Note 1 on the last page of the monograph.]

**ON TAIL-WAGGING**

We hope this chapter has conveyed our view that performance measurement can be a helpful tool for you in both planning and evaluation. If we end it by describing one of the maladies encountered in some applications, you will know it is not to discourage your growing use of these tools. As you will soon see, you very likely already know the cure.

The malady we have in mind has been named the “tail-wagging syndrome,” after a conceptually similar and painful syndrome noted by vets in dogs confused about what wags what. It tends to thrive in situations where some criteria lend themselves readily to measurement and others do not. It could have afflicted HFS if the agency had not been careful. To show how, let us change a few facts in the case example.

A very early symptom is the view that a decision criterion (or objective) without a performance measure is somehow marred or imperfect. Sometimes there is even an
underlying sense of guilt that a measure cannot be found. Whatever the trigger mechanism, criteria with hard measures start to grow in stature independent of their place in the underlying problem. For HFS, this symptom would have manifested itself first in the relative weight attached to the criteria “increase income” and “diversify.” Both had hard performance measures, so both would have acquired more and more weight relative to the hapless criteria on “maintain relevant and effective services” and “preserve autonomy.” Plan B, therefore, would have easily won out over Plan A.

The second stage in the syndrome begins when the agency implements the alternative selected. By now, those responsible for implementation have gathered that “measurable is better.” So, with appropriate concern about their own evaluation, they pay very special attention to those criteria with hard measures appended. In HFS, implementers would have stressed, even more than planners did, “getting new money” and “getting more funding sources.”

Please note the last sentence. The implementers have already started to label the criteria with the names of the performance measures involved. Just a convenient shorthand, you protest, but there is a danger. Taken together, the criteria express what the agency saw as a good solution to a complex funding problem, a problem much richer than simply dollars of new money or number of new funding sources. The third stage in the syndrome has started. Not only have the people involved shelved “relevance and effectiveness” and “protect autonomy” as criteria, they have also simplified the two remaining criteria that do have performance measures. More importantly, they make detailed, day-to-day decisions as though these two simplified criteria were the only ones that mattered.

The fourth and final stage of tail-wagging emerges when it comes time to evaluate results. Very proudly, the implementers in HFS point out that they have succeeded far beyond expectations. Instead of $240,000 of new money, they have $285,000. Instead of the predicted four new funders, they have actually acquired six. Granted, they had to make radical compromises on the relevance and effectiveness of services offered by the agency to achieve these fantastic results, but these compromises are nothing the agency and the community won't get used to in time and really a very small price to pay for such evident success. Autonomy? Well, isn't this the age of interdependence? And, again, look at those impressive, tangible results!

In advanced cases, the fourth stage of the syndrome finds agency management and the board not only agreeing but quickly starting new projects to exploit this exciting new management technique. No one questions the success of an effort with such impressive measurable results, so no one discovers that the agency has not only failed to get comprehensive cost
from its new funders, it is, in truth, providing a direct, out-of-pocket subsidy for some of the
new services.

We are exaggerating, of course. Tail-wagging is readily detected and controlled with the
application of good management skills. However, add some basic conceptual understanding
of the role of performance measures and there is no need to miss the many benefits these
measures can offer. And that is really the point of this whole chapter.
CHAPTER VI
A CONTEXT FOR PROGRAM EVALUATION
(Another Trail Through The Semantic Jungle)

This chapter starts with one view of the context in which human service delivery agencies operate. We find this view--or model--useful in clarifying some of the issues in evaluation and planning and you might find it a helpful addition to your conceptual tool kit. The model also helps to clarify the meaning of some jargon terms associated with evaluation and planning (specifically: input goals, output goals, and outcome goals) and to provide some perspective on one of the “in” slogans of some of the field's current critics, that is: “Why can't we get those people to be results oriented?” Finally, the model provides us with a way to review this monograph that may help you tie together some of the conceptual tools we have discussed.

THE MODEL
If you will glance at the upper portion of Figure 8, you will see that the model is a very general one. It suggests that actions, which consume resources, produce results. These results, in turn, have consequences, and it is hoped that consequences will help achieve ends or ideals.

Let us hasten to be more specific. In the middle part of Figure 8 we suggest an example. The action is a family life education program in a family agency, one that happens to be directed at single-parent families with young children. Input to this program consists of plans, staff time, money, and most importantly, participants. One illustrative result (and we stress that this is just one of many) is that participants acquire some new skills in coping with problems they are likely to meet as single parents raising young children. One illustrative consequence that might follow from this result is some reduction in the future incidence of family problems in the community, especially in the target population of single-parent families. Moving forward in time, and upward in generality, this consequence (it is hoped) will help achieve an end or ideal valued by the agency, that is, some improvement in the quality of family life in the community.
This model suggests that there is a continuum from the immediate results of a program to some future, much broader ends or ideals. The labeling of the extreme right as “ends or ideals” suggests that there is a purpose behind the action; the model, therefore, is goal oriented.

INPUT, OUTPUT AND OUTCOME GOALS

At the bottom of Figure 8, the goal orientation of the model is labeled with some terms often used in the literature. Input goals refer to expectations on the input to an action (resources to be invested, participants to be involved, and the like). Evaluation of input goals would deal with the extent to which plans had been carried out, and that recalls a friend from Chapter IV, Program Evaluation Question II: Was the program carried out as planned? Also, because plans are an essential input to an action, Program Evaluation Question III is applicable here, that is, the issue of how well planning was carried out.

Output goals cover expectations on the immediate results of a program. From a planning perspective, these are the characteristics of a good solution for the problem the program addresses. As we suggested in Chapter III, another useful label is “decision criteria,” and still another, “objectives.” From an evaluation perspective, Program Evaluation Question I comes to the forefront: Did the program achieve what was expected?
Outcome goals reach out from the action and from immediate results toward consequences and toward ends or ideals. From a planning perspective, there is an interesting relationship between outcome goals and output goals. Output goals (that is, those related to immediate results) can be viewed as the means by which outcome goals may be achieved. From an evaluation perspective, the extent to which outcome goals are achieved is also the subject of Program Evaluation Question I.

From a practical point of view, however, something important happens when you shift from output goals to outcome goals. Evaluation becomes much more difficult. In the family life education program in Figure 8 it would be hard enough to assess the extent to which participants had acquired new skills, but consider how difficult it would be to go to the outcome goal and assess the extent to which there has been some reduction in the incidence of family problems in the target population of single-parent families.

Much of this added difficulty lies in the generality that will characterize most outcome goals: They tend to be global. One way to deal with global goals is to break them up into significant, but more manageable, parts. In the example in Figure 8, one part of the outcome goal would be the output goal of: Improve skills of single-parent families to cope with problems of preschool children. Another output goal might be: Increase the availability of community support services, such as day care. In other words, one way to deal with outcome goals is to partialize them into smaller, bite-sized goals, which, as the examples just given indicate, can be viewed as output goals. We will say more about this strategy at the end of the chapter.

SOME THOUGHTS ON THE BATTLE CRY: “GET THOSE PEOPLE RESULTS ORIENTED!”
Some critics of human service delivery organizations emphasize what they perceive to be the failure of these organizations to be concerned with results. In terms of the model in Figure 8, their complaint is an undue emphasis on input goals and little, if any, attention given to output or outcome goals. How valid is this criticism?

There is a lot of evidence that input goals do, in fact, receive a relatively large amount of agency attention in both planning and evaluation. Consider, for example, the extent to which input is emphasized in assessing counseling services. Agencies measure the number of clients seen, the characteristics of those clients, the number of interviews produced, the number of hours of contact, the percentage of counselor time spent in direct contacts, case
turnover, cost per interview, cost per family, and the like. These are performance measures largely relevant to input goals.

We suggest that attention to input goals is not, in itself, a valid reason for criticism. As we argued in Chapter IV, answers to Program Evaluation Question II (“Was the program carried out as planned?”) can be extremely valuable for both internal and external users of evaluation. If agencies totally ignored input goals, they would justifiably be targets for criticism.

Where the critics have a point, however, is the very sparse attention paid by many human service delivery agencies to output and outcome goals. For what it's worth, we suggest the principal reason is the failure of many agencies to identify these goals in planning. To restate this observation in the terms used in Chapter III, relatively few agencies make decision criteria explicit when they make program choices. Without clear criteria and without the projected performance of a program in relation to these criteria, you have little hope of evaluating output goals, of being “results oriented.”

Some organizations might protest that they do, in fact, pay a lot of attention to their purpose or their mission. When you look closely, however, you often find the agency stopped at the point of making its ends or ideals explicit. A well-worded statement of the mission of an organization is a useful product, but it is a long way from being a practical tool in making tough program choices. We have more to say about this in Chapter VII below.

More attention to output goals in planning and evaluation would probably appease outside critics. Would it have any internal value for agencies? In our view it would, and here is one supporting argument. When an organization focuses almost exclusively on input goals, especially in program planning, it is making an implicit assumption. It is assuming that the justification for the program is a “given.” The choice involved in program planning is not whether to offer the program, but rather how. In contrast, if an agency at least attempts to examine output goals, it opens up the opportunity to examine whether a particular program is, in fact, the preferable way for it to try to achieve those goals. And even if the process confirms the value of continuing that program, the agency may get insights into how to modify that program to better achieve the desired results. Without attention to output goals, this tailoring process tends to find its rationale solely in input goals, such as “reduce unit cost” or “increase the number of family life education sessions.”

A BRIEF RECAP
We have suggested so far in this chapter that it helps to put both evaluation and planning into the perspective suggested by the model in Figure 8. A human service delivery agency is a purposeful entity; it seeks ends or ideals. To guide itself toward these ends, an agency can
work backward from them to identify the immediate results that it should seek from its programs, and it can also attend to the input required if those programs are to achieve those results.

Some alternative terminology was suggested in Figure 8, namely: input goals, output goals, and outcome goals. For the evaluation of goal achievement (the “how-did-it-go” question), the three program evaluation questions in Chapter IV can be tied to this model. Program Evaluation Question I (on whether expected results were achieved) relates to the appraisal of output and outcome goals. Both Program Evaluation Question II (on whether programs were carried out as planned) and Program Evaluation Question III (on whether it was the “right” program) tie into the appraisal of input goals.

“Interesting theoretical stuff,” you may be saying, “but is it possible, or practical, to apply it in real life?” We think it is both possible and practical if you set reasonable expectations. By way of support for this view we will describe in the next chapter how our friends at Hypothetical Family Service might have approached the problem.
CHAPTER VII
FROM MISSION TO PROGRAM
(A Leap Of Faith Or A Path To Follow/)

Have you heard any of the following? “What is your organizations mission?” or “What business are you in?” or “What is the real purpose of your agency?” Such questions are often be followed with the comment “Oh, by the way, keep your answer very short and to the point.” We use the label “mission statement” for a response to such questions.

A mission statement can have a variety of uses for an organization:

- It can inspire and motivate people in the organization.
- It can focus and shape organizational behavior (e.g. “The customer is always right”).
- It can help create a positive public image for the organization.
- It can help differentiate the organization from its competitors (e.g. “As the original maker of widgets, we can best meet your needs.”)
- It may help in program planning and evaluation.

All of these potentials can serve an organization’s legitimate needs. The effort to get agreement on a sharp, clear, well crafted mission statement can, we think, be a good investment. Our concern here, however, is with the last potential noted. How, in fact, can an organization move from a broad mission statement to detailed program plans to carry out that mission? We have observed three answers to this question;

1. Make no connection whatsoever. In fact, consider them two as totally separate activities. (This was actually the position of one consultant we met!)
2. Accept a connection but consider it to be “by inspiration only.” Immerse people in the task of crafting a mission statement then move them to design specific action where they will find themselves “inspired” by their work on the mission statement.
3. Move from mission statement to action by first working down from the mission by asking the question “What do we mean by that?” Keep at it until you reach objectives that can serve as useful decision criteria in program planning decisions.

Is the third answer feasible? Consider how our friends at HFS might have tried this option to connect the mission to program planning.
HOW HFS TACKLED THE PROBLEM
Deciding to follow the systems analysis view of a decision (Chapter III), HFS staff and board began with the problem. For them, the problem was the prevalence of many serious obstacles to the achievement of a high quality of family life in the community. They believed the resolution of this problem was their ultimate mission as a family agency. At this level of generality, however, they found it virtually impossible to grapple with the problem, so they went on to describe some more specific characteristics of a good solution. (They embarked, as you will recognize, on a search for decision criteria.)

Conscious of the enormity of the problem and the fact that the agency was not omnipotent, HFS planners deliberately sought out those characteristics of a solution to which they might reasonably contribute. They identified three major characteristics:

I. Families should have greater internal strengths to achieve a higher quality of family life.
II. The environment in which families function--especially significant institutions--should, on balance, have a much more positive impact on family life than it does now.
III. HFS, as an entity, should have a greater capacity to contribute to the attainment of an improved quality of family life in the community.

Although still at too general a level to work with, these three broad characteristics of a good solution seemed to HFS decision makers to be reasonably complete from the agency's perspective. If you glance at the top of Figure 9, you will see they went on to restate them as criteria or goals. They appear to be outcome goals in the sense of the model covered in the previous chapter. Because these three high-level outcome goals still provided a very inadequate tool for either evaluation or planning, HFS proceeded to partialize or parse them into more specific subgoals. The aim was to work toward identifying output goals.

As the left-hand portion of Figure 9 indicates, the first major outcome goal was split into two parts. The first covered the capacity of families (that is, skills and knowledge) to live their lives so that they could achieve their full potential. The second zeroed in on the availability of direct assistance to families to help them reduce stresses that adversely affect the quality of their life. To get output goals for these two outcome goals, HFS would have to keep going. We do not show it in Figure 9, but the agency could go on to be more specific about priority target groups, and more specific about priority areas of skill and knowledge.
We will illustrate this detailing process for Goal III but let us pause here to make a point. HFS was tackling a problem. They were concerned with both planning and evaluation. For planning, the agency wanted criteria specific enough to help them make choices, especially choices among programs. For evaluation the agency needed criteria specific enough to assess the extent to which they had been met. For Goal I this meant criteria that would, at a minimum, discriminate between various possible target populations of families in the community, various areas of skill and knowledge, and various areas of stress.

Shifting now to the third outcome goal, on the far right of Figure 9, we show an example of how HFS could have moved from outcome level goals to output level goals. A part of the goal of strengthening the capacity of the agency was strengthening agency resources. A part
of this was acquiring needed resources. A part of this, in turn, was strengthening the financial base. And when this last goal was dissected, HFS ended up with:

[] Increase the amount of funds
[] Increase the diversity of sources of funding
[] Increase the reliability of sources that do fund the agency

The first two criteria (or “output goals”) lend themselves readily to performance measures as discussed in Chapter V. And, as we are sure you have already spotted, these two criteria played a prominent role in the case example in Chapter III. (Parenthetically, there were two other criteria in that case example. The “protect autonomy” criterion would tie into the goal on the far right, “strengthen the agency's capacity to direct its resources to achieve goals.” The other criterion, “acquire effective and relevant services,” would tie into the detailed, output-level goals under goal I.)

If you assume HFS went on to detail all of the goals in Figure 9 as we have just illustrated, the agency would still be left with a very crucial step before it could use the result effectively in either planning or evaluation. Output goals would have to be weighted to express the relative importance of their achievement to the agency and its sense of the role it should play in the community. Without clarity about relative importance, these output goals (or decision criteria) would have very limited utility. Even in the narrowly defined problem we used in the case example in Chapter III, the weights attached to criteria played an important role in the choice among Plans A, B, and C and a key role in the appraisal of Plan A when it had been implemented. For major program planning and program evaluation, these weights are even more crucial.

Assume that HFS did, in fact, express the relative importance of the goals in Figure 9. Would the result have helped them in planning and evaluation? We leave the answer to your judgment. There are many approaches to planning and evaluation. In fact, the next and last chapter discusses the decision to pick one. (The particular approach described in Figure 9 may be one you would like to explore further. If so, please see Note 1 on the last page.)
CHAPTER VIII
CHOOSING PLANNING AND EVALUATION TECHNOLOGY
(OK, I’m Sold. So, What’s Next?)

If you think you need to invest in planning and evaluation, how do you go about choosing from among the many competing technologies? We will use this choice—this decision—to try once more some of the conceptual tools described earlier. Specifically, here are some thoughts on the first four steps in a decision, working up to the choice.

STEPS #1: PROBLEM ANALYSIS AND STEP #2: IDENTIFICATION OF OBJECTIVES
While each situation will have unique characteristics there are some themes that often arise. For example:

A. What really motivates the desire for more planning and evaluation? It may not be obvious. Some possibilities: “We know what our programs should be, we just need a way to persuade others we are right.” Or “We—the board, management and staff-- do not agree on directions and we should.” Or “They keep bugging us to plan so what’s the cheapest way to get them off our back?” Or “There must be a better way for us to use our resources.”

B. How can do effective planning and evaluation and still keep the demands on the time and patience of board and staff to tolerable levels?

C. How can we get sufficient comprehension of the process by board, staff and management to enable us all to gain ownership of the results and to retain that understanding past the initial effort?

D. How can we get results that will be perceived as credible by our significant publics (funders, clients, etc.)?

Again, these are illustrative. From an understanding of the problems should come criteria for a solution, the objectives by which to judge possible solutions. From this same understanding should come the weights that should be attached to these objectives.

STEP #3: IDENTIFY ALTERNATIVE SOLUTIONS
Someone once said that there are as many different approaches to planning and evaluation as there are consultants willing to sell their services. To locate appropriate planning and evaluation technology an organization should tap all of its contacts and affiliations for
recommendations. Referrals and recommendations from those you trust are especially useful.

As you assemble information on alternative technologies be very wary of the semantics. Terms such as goal, objective, mission and strategy can have widely varying meaning. It can help, we believe, to have a conceptual framework into which you can try to fit the components of a specific technology regardless of the terms they use. One such framework has been described in this monograph.

As you review possible technologies you need to get as realistic an estimate as possible of the resources (time and money) that will be demanded of you and the span of time required.

STEP #4: EVALUATE ALTERNATIVES
This step, as we have suggested, involves evaluating alternatives in the light of your understanding of the problems you are trying to alleviate and the objectives/criteria that seem to describe what a useful solution looks like. You should be using weightings that express your view of relative importance. Also important in this step is to establish an estimate of the resources (time and money) that you have available to work with. This itself may help screen out alternatives that are clearly too costly.

NEXT STEPS
Given your evaluation of alternatives and your estimate of resources available the next step would be to select an alternative and implement it. However, you will likely find yourself recycling through all of the steps above before you make a final decision. For example, if you are stuck trying to evaluate a particular approach you might want to get back to understanding the problems you face better. Also, the technology can often be tailored to your specific needs. We will leave you here to grapple with these steps if you find it warranted.

A FINAL QUESTION
Has this brief monograph helped? It is absolutely clear that you are the final judge on this particular evaluation question. We hope we have helped you feel more comfortable with the family of conceptual tools discussed: the model of a decision, some different meanings of the “How-did-it-go” question, performance measurement and input, output and outcome goals. If so, we have met one of our output goals. To the extent that you find ways to use these tools, we will have met our other output goal. We would welcome--and appreciate--your assessment.
[NOTES:
1. Stephen M. Drezner and William B. McCurdy, *A Planning Guide for Voluntary Human Service Delivery Agencies* (New York, N.Y.: Family Service Association of America, 1979). [Alas, it is out of print! If you know a family and children’s agency that was in FSAA membership at that time they may still have their copy. The successor organization to FSAA is the Alliance for Children & Families, 11700 W. Lake Park Dr., Milwaukee, WI 53224. They may have information on where you can see a copy.]