# A Phased Approach for Conducting Program Evaluations*

*Alan Winberg*
*Indian and Northern Affairs Canada*

**RÉSUMÉ**

Pour être utiles, les évaluations de programme doivent être conduites de façon suffisamment rigoureuses afin d'en assurer la credibilité. Aussi, elles doivent être complétées au bon moment, c'est-à-dire, être disponibles au moment où on en a besoin pour prendre des décisions. Une façon d'attaquer le problème d'être prêt au bon moment est d'adopter une approche echelonnée pour la conduite et la remise des rapports d'évaluations de programme.

Dans cette optique, les études d'évaluations se découlent par étapes. Chaque étape répond à des questions séparées. À la fin de chaque étape, on rend compte du progrès au client et on planifie la prochaine étape. Ainsi, les gestionnaires n'ont pas besoin d'attendre la complétion de l'étude avant de recevoir des renseignements et des analyses utiles sur les questions à l'étude. À la complétion de chaque étape, il est possible de décider si l'on doit procéder à l'étape suivante, et si oui, de quelle manière.

**ABSTRACT**

To be useful, evaluations have to be of sufficient rigor that they are credible. Also, they must be timely – that is, available when needed for decision making. One way of addressing the "timeliness" issue is through the adoption of a phased approach for the conduct and reporting of evaluations.

Under this approach, studies are broken into discrete phases, each addressing and reporting formally to the client on specific evaluation issues and each laying the groundwork for the next phase. Thus, management does not have to wait until completion of the entire study for timely information and analysis on these issues. Completion of each phase provides a decision point on whether to proceed with the next phase, and if so, in what specific manner.

## Introduction

Program evaluators are usually faced with a fundamental problem in planning and carrying out studies. Evaluation studies have to be credible: the data and analysis must provide a firm basis for the findings and recommendations of the studies. In addition, if they are to be useful to management, in the first place, the issues considered in the studies must be those which contribute most to addressing the key problems facing management (Patton, 1978); secondly, reports must be timely. They must be available when needed for decision-making (Rich, 1979; Datta and Perloff, 1979; Lindbloom and Cohen, 1979; Weiss, 1972). The problem, thus, is one of being simultaneously credible, useful and timely. Producing effective stud-

* An earlier version of this paper was presented at a seminar organized by the Office of the Comptroller General of Canada in Ottawa, November 7, 1985.

ies, ones that are actually used for decision-making, requires that the evaluator find the right balance. He/she must be practical. Studies must be of sufficient reliability and rigor that they are credible to the clients of the evaluation (Alkin et al., 1979).

One of the ways in which this issue can be addressed is through the adoption of a phased approach for the conduct and reporting of evaluations. This paper describes the typical phases of an evaluation and discusses the nature of the phased approach.

**Phased Approach**

Under a phased approach, evaluation studies are broken into discrete phases, each addressing and reporting formally to the client of the study on specific evaluation issues, and each laying the groundwork for the next phase. In addition, the completion of each phase may provide a decision point on whether it is in fact useful and practical to proceed with the next phase, and if so, in what manner.

As will be considered in this paper, use of a phased approach has been found practical as a way of:

i) making available to clients the best information and analysis possible, given limited resources and time (permitting evaluation clients to better assess the pros and cons of making partial decisions and/or deferring other aspects of decisions until the full study is completed);

ii) focusing the evaluation work in a practical manner which allows scarce evaluation resources to be applied to the highest priority issues in a given limited timeframe;

iii) improving the utilization of evaluation results through:

   a) permitting early introduction and formal explanation in a written report of innovative ideas for change,

   b) providing a window for obtaining client reaction and views on the evaluation work as it progresses (Patton, 1978),

   c) providing an opportunity for mid-stream correction thus increasing the likelihood that the evaluation work is on track and relevant to the informational and analytic needs of the client (Patton, 1981); and

iv) providing an opportunity to terminate formal evaluation studies once the most important evaluation issues have been answered with sufficient rigor and credibility for management decision-making on a given government program. This has allowed limited in-house evaluation resources to be used on the most pressing issues and enhanced the use of evaluation as a management tool.

To make these ideas more concrete, it is useful to consider the typical phases in conducting an evaluation of a government program.

## Typical Phases to an Evaluation

Most evaluation research will generally follow a number of steps, each of which could be considered a separate "phase" within the context of this discussion (Datta and Perloff, 1979; Rossi et al., 1979). For discussion purposes, six phases can be identified:

1. A *planning phase* is undertaken to understand the fundamentals of the entity to be evaluated and the context within which it functions. Another purpose of this stage is to set out the major issues to be addressed and assess the optional methods that could be used for data collection and analysis as well as the various ways the study could be designed. This phase usually involves a documents review, a literature review and a limited number of in depth interviews. The output of this phase is a plan or a feasibility (evaluability assessment) study (Rutman, 1980; Wholey, 1977).

2. A *decision-making phase* follows to select the most appropriate scope for the study and how the work is to be undertaken. This phase involves the evaluator reviewing the evaluation options with the client and the client making a decision. The output of this phase is terms of reference for the study (Canada, Office of the Comptroller General, 1981).

3. A *preliminary data collection phase* is then undertaken to get an initial fix on the issues. This phase usually involves analysis of secondary data obtained from a directed file search, a directed documents review, or a directed literature review. Generally, some primary data would be collected in interviews with key respondents. This phase may produce preliminary findings on program objectives, program logic, results indicators and unintended impacts. During this phase, the key evaluation issues become clearer. The output of this phase would be a formal report to the client, and if appropriate, recommended modifications to, or confirmation of, the study plan.

4. A *preliminary analysis phase* involving more in-depth data collections, directed at specific issues, would follow. This phase would usually include in-depth consultations with a broader number of respondents, larger scale surveys, and preliminary analysis of the data obtained. Detailed findings would result on some issues. A wide range of options open to management would be identified. Each of these options would address certain identified issue areas in a different manner. The anticipated range of possible outcomes may be estimated in a preliminary manner. Also, during this phase a preliminary range of valuations may be estimated for the various outcomes. The output of this phase would be a report, generally of very high interest to the client. This is his/her first chance to get a reading on the range of possible findings. This report should contain recommended modifications to, or confirmation of, the study plan.

5. An *in-depth comparison of options phase* would follow involving continued treatments and data collections, all in more depth and focusing on specific issues. Case studies of "normal" cases as well as "deviant" cases could be undertaken at this point in the work. Detailed findings on more issues would generally result. During this phase, the options set out in the previous phase would be analyzed against the findings. The expected consequences or outcomes of each of the various options would be predicted, and a range of valuations would be estimated for the various outcomes. At this time, some options could be eliminated from further analysis based on this initial valuation. A great many evaluations would be considered to be completed at this point in the work.
6. The last phase would be an *implementability analysis* which would include aggregating the diverse findings and analysis and making a comparison of a small number of detailed options. This comparison would usually include estimating expected outcomes and a range of valuations of each outcome. Questions related to the implementation of the recommended option(s) would be analyzed in depth. Pilot tests of the best options could also be undertaken as part of the final stage(s) of the study.

These phases are set out graphically in Exhibit I. Exhibit II details the evaluation research tasks normally associated with each phase. In essence, this exhibit shows that the number of data sources and the amount of data increases in each phase. Also, as one moves further down the chart, random selection is used increasingly, thus improving the reliability and validity and therefore the credibility of the findings. In addition, the "implementability" questions that may relate to any given recommendations receive increasing attention and analysis as we proceed down the chart (Reichardt and Cook, 1980; Rich, 1981).

Some studies, of course, would terminate at the end of any of these phases, and normally a report would be produced upon termination. Also, in some cases it may make good sense to combine phases, the result being that fewer formal reports (i.e., say three or four rather than five) would be issued during evaluation. The key point about the phased approach here presented, however, is that formal reports are presented at several points during the study.

### An Example in Which a Phased Approach Was Used

In an evaluation of Canada's Metric Commission, carried out in 1982 and 1983, the work was broken into five phases (Canada, Consumer and Corporate Affairs, 1983). The Commission had been established in 1970 to serve as a catalyst to facilitate the conversion of the system of measurement used in Canada to the international, metric system. The Planning Phase for the evaluation was carried out over a three-month period. In Phase Two, terms of reference for the study were approved.

The terms of reference set out the key issues that had been identified during the Planning Phase. The report issued at the end of Phase Three recommended clarifications for the role, strategies, and operational objectives for the Commission. The basis for these recommendations was a file

search, a documents review, a literature review and a set of interviews with key respondents. The Phase Three work, while not providing definitive answers, served to narrow the range of issues to be addressed in the rest of the study.

In Phase Four, the achievements of the Commission were assessed and program priorities were defined. Methodologies used included more in depth file reviews, documents reviews, a literature review, interviews, surveys, case studies and expert opinion. Options which set out where the Commission should target its work in order to achieve the greatest expected results over a specific timeframe were assessed. Although five options were presented, the Phase Four report found that one option appeared to stand out over the others. The report recommended that in the next phase the consequences of implementing this option be considered fully, and the other options be assessed in considerably less detail.

## Exhibit I

## Typical Phases of Evaluation Studies

| | |
|---|---|
| **Phase I** | Planning phase — outlining the program's description, key issues and evaluation plan. |
| **Phase II** | Client review and decision-making on plan to produce terms of reference for an evaluation. |
| **Phase III** | Clarification of objectives, and results indicators. |
| **Phase IV** | Preliminary (descriptive) assessment of general problem areas. Identification of alternatives and preliminary analysis. |
| **Phase V** | In depth assessment of most promising alternative methods of addressing key program problems:<br>• prediction of outcomes<br>• valuation of outcomes |
| **Phase VI** | Detailed comparison of key alternatives, analysis of "implementability" of limited number of viable options. |

## Exhibit II
## Typical Evaluation Tasks

| | |
|---|---|
| **I Planning** | Identification of secondary data bases<br>General documents review<br>Limited number of face-to-face interviews with key respondents<br>(Elaboration of study plan) |
| **II Terms of Reference** | Communication of plan to Client(s) |
| **III Preliminary Data Collection** | Directed preliminary examination of specific program files<br>Directed documents review<br>Directed literature review<br>Interviews with key respondents<br>Natural observation, journalistic analysis<br>Preliminary analysis of objectives, program logic, potential impacts and effects |
| **IV Preliminary Analysis** | In depth consultations on specific issues with larger number of responsents (may include random selection)<br>Surveys of various parties impacted by program (may include random selection)<br>Experiments, Quasi-experiments conducted<br>Natural observation may continue<br>Options identified for program change<br>Preliminary analysis. Detailed findings on some issues. |
| **V In depth Comparison of Options** | In depth analysis of data<br>In depth case studies undertaken, additional experimentation, natural observation if required<br>Aggregation of findings, analysis |
| **VI Implementability Analysis** | Pilot testing of key alternatives<br>"Implementability" analysis through expert opinion, interviews with key respondents, natural observation |

The Phase Four report, which was completed over a four-month period, was accepted by Senior Management soon after it was completed. This meant that the implementability analysis, to be done in Phase Five, could be very focused and completed over a two-month period. This work involved the preparation of action plans, extensive interviewing and focus group meetings to insure that the recommendations could be carried out.

As a result of the work undertaken through the phased approach, the evaluation found that there was significant momentum for conversion to metric in Canada and that the foundation had been laid for the successful completion of the conversion. The study identified the program tasks to be carried out by the Metric Commission up to March 31, 1985 and concluded that the Commission could be phased-out by March 31, 1985 without affecting the pace of the conversion process in the following years. The study recommended three priority areas where the Commission should concentrate its efforts over the phase-out period. All of these recommendations were accepted and implemented over the recommended timeframe.

From this example, one can see that the specifics of the various phases must be tailored to fit the circumstances of the program being evaluated and the issues of concern to management. In the Metric study, five phases were used; in a recent evaluation of a labelling program (Canada, Consumer and Corporate Affairs, 1985), four phases were used. So that all concerned are aware of this phased approach, there are additional advantages to setting out this phased-reporting approach in the terms of reference for the study.

**The 80/20 Rule**

Someone wise once said that the first eighty percent of the study findings can be had for about twenty percent of the total costs and that that last twenty percent, including the "finishing touch", requires eighty percent of the effort. Maybe the percentages are not all that exact, but anyone doing this kind of work, will recognize the basic truth behind the concept. For example, going from a first draft to a perfectly-finished product may require a large amount of effort and expense, but only rarely results in major and significant changes to the "core" or the basic ideas of the report.

A phased approach to evaluation reporting often may provide an opportunity to complete work and provide good quality advice without the very expensive effort to produce the "perfectly-polished" final product. Certainly, part of that last twenty percent is often worth doing. But confirming only those vital points or controversial issues identified in the previous phase is certain to be less expensive and time consuming than a wall-to-wall validation of every evaluation finding.

It is noted that the intensity of effort, and the resources required for each phase of the evaluation process increase in a geometric fashion. Each phase, in effect, serves to increase the reliability of the information being gathered. It is to be expected that for any given increase in reliability, or stated differently, any given reduction in uncertainty, increasing costs are involved. For most governmental programs, no evaluation can provide the "right" answer. There are no single, pure truths. There are no certainties that a

given event will in fact occur; rather there are probabilities that can be attached. There are no absolute true measures; rather there are ranges of uncertainty surrounding any social measure. It is widely accepted that these ranges of uncertainty, or stated differently, confidence intervals, should be pointed out when reporting evaluation findings (Alkin et al., 1979; Chen and Rossi, 1983; Cooper, 1982; Glass et al., 1981; Rich, 1981; Scherill, 1984).

While this is all straightforward to social scientists, it may be a new way of thinking to many potential clients of evaluation. A phased approach to evaluation reporting drives home these basic concepts to the client in a very dramatic and physical manner. A formal report is produced on completion of each phase. It is well understood that if greater reliability (reduced uncertainty) is required, the next phase of the work should be undertaken. The cost of this information, and the cost of delaying a given management decision, can therefore, be weighed against the expected value of the additional information in terms of reduced uncertainty. The client is in a much better position at that stage to make this trade off.

Any analytic effort, no matter how meticulous, can be criticized; in a phased approach, critics of the study are given an opportunity to start to make their cases early on, rather than on completion of a final report. Valid criticism of methods or design can be accommodated in subsequent phases of an evaluation. Simply put, a phased approach allows for mid-stream correction and reduces the possibility of new objections being raised on completion of the work or of the work getting off track in the first place (Rutman, 1980; Weiss, 1972; Weiss and Bucuvalas, 1980).

**Breadth, Depth, Time (BDT)**

In a given amount of time, with a given amount of resources for an evaluation, a series of trade offs must be made. The trade offs are generally along the dimensions of study breadth, study depth and timeliness. Breadth would include the number of issues examined, or the number of alternative options considered in the study. Depth refers to the amount of data that would be collected on a given issue or option, and the quality of that data, or stated differently, the level of confidence required that the data that has been collected is valid. Timeliness refers to the presentation of evaluation findings and results at that particular point when they can most usefully assist management decision making. Exhibit III illustrates the type of trade-offs that can be considered (Cronbach, 1980; Glass et al., 1981; Pillemer and Light, 1980).

At point "A" shown on the chart, studies that would be "quick and dirty" would be ones that look at only a few key issues in a rather superficial manner. In general, such studies are highly qualitative and are produced relatively quickly in response to an urgent need for information prior to a decision which must be made in a short timeframe. As we add issues to be considered in a study, or as we increase the required level of depth on any number of issues, we increase the time required and other resources required to complete the study.

"Quick and dirty" studies can also be seen from the vantage point as being the initial phase(s) of a more rigorous study. In this sense, if useful to

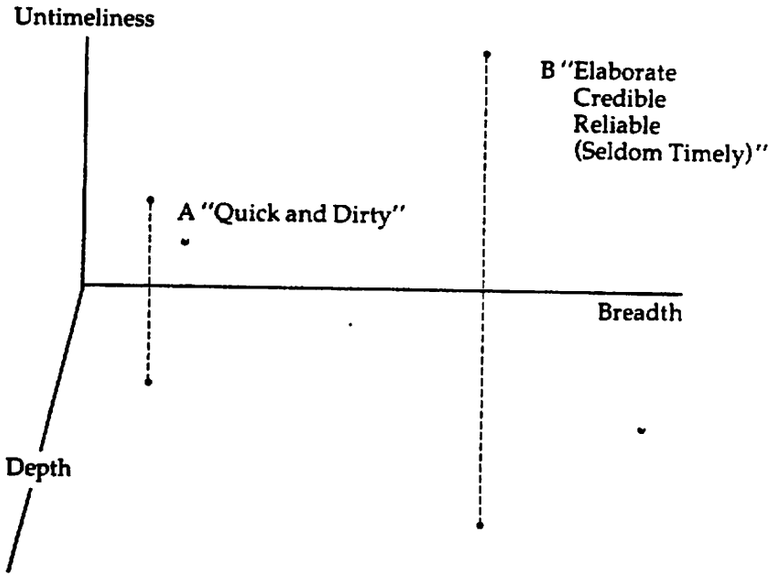
management, even the most rigorous of studies, if "phased", could provide for a preliminary analysis in a report, once the initial phases of the evaluation were completed. Based on this report, the next phases of the evaluation could be accelerated, or on the other hand, reduced, deferred or cancelled if no longer required. In short, according to the value of the information to management at various points in time, the study findings and analysis could be reported when most valuable in their best possible form at that time. If it helps decision making there is nothing wrong with providing interim results as long as they are reported as such and as long as the range of uncertainties or confidence interval is clearly set out.

At point "B" shown on the chart, we are in the area of very expensive, "wall-to-wall", long term studies. Here of course time and resources are relative terms and should be taken in the context of what is being measured and assessed and in what context. However, in general, there is really no limit to how much research could be undertaken on a specific complex issue. Although from a practical viewpoint, there most likely will come a time when further research would not be expected to produce information whose incremental value (in terms of reducing uncertainties surrounding a given issue) would be likely to exceed its cost.

**Exhibit III**

**Trade Offs For Planning Phase**

**The BDT (Breadth-Depth-Timeliness) Problem**



A non-phased approach, or even a more prudent two-phased approach (evaluability assessment followed by a study), requires the client for a study to decide rather early on where the evaluation should be on the BDT graph shown in Exhibit III (Patton, 1978; Reichardt and Cook, 1980; Rutman, 1980; Wholey, 1977). A phased approach keeps the options open longer and provides more flexibility. Also, it can reduce costs because each phase provides an opportunity to focus the work in the next phase, and importantly, to receive feedback from the client on whether, in his view, this focusing is reasonable and appropriate. Significantly, each phased report allows an opportunity to terminate the evaluation project gracefully once enough information has been collected on the issues with an adequate degree of reliability and credibility for decision making.

## Context

This approach was developed to suit a particular organization in the Canadian federal government. Evaluation in this context is seen as corporate work, intended as a management tool. A phased approach may be less appropriate in different contexts where evaluation is seeking to provide information to clients who may not be involved in managing a given program.

In evaluation work undertaken to date at the Canadian Department of Consumer and Corporate Affairs, use of a phased approach has permitted the introduction of new ideas or ideas for change at an early stage in the evaluation. These new ideas can then be discussed with client groups, and the various issues associated with them can be assessed further in later phases of the study. This has served to reduce the threatening aspects that may be attached to some innovative ideas in some bureaucratic institutions where "change" can often raise unnecessary concerns. This has been a major feature of the phased evaluation approach often used in this organization where the internal program evaluation unit was established as an agent or catalyst for change where changes were appropriate and necessary. The function has received the full and complete support of senior management in undertaking this role.

## Conclusion

In the experience to date using this approach for corporate program evaluations as a management tool, each phase in the evaluation studies has typically served to narrow the focus of the evaluation efforts of the next phases of the study. In turn, generally this has allowed either more in depth treatment of the priority issues or a reduction in the resources required to complete the work as planned. This has been particularly true in cases where the evaluation was aiming, through systematic data collection and analysis, to produce timely and credible information on key issues for decision making rather than to provide scientific-type proof on some aspect of a program or policy issue.

## References


Alkin, M.C., Daillak, R., and White, P. (1979). *Using evaluations: Does evaluation make a difference?* Beverly Hills, CA: Sage.

Canada. Department of Consumer and Corporate Affairs. (1985). *Energuide Evaluation.* Ottawa: Department of Consumer and Corporate Affairs.

Canada. Department of Consumer and Corporate Affairs. (1983). *Metric Evaluation.* Ottawa: Department of Consumer and Corporate Affairs.

Canada. Office of the Comptroller General. (1981). *Guide on the Program Evaluation Function.* Ottawa: Supply and Services Canada.

Chen, H.T., and Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review, 7,* 283-302.

Cooper, H.M. (1982). Scientific guidelines for conducting integrative research reviews. *Review of Educational Research, 52,* 291-302.

Cronbach, L.J. (1980). *Toward reform of program evaluation: Aims, methods, and institutional arrangements.* San Francisco: Jossey-Bass.

Datta, L. and Perloff, R. (1979). *Improving Evaluations.* Beverly Hills, CA: Sage.

Glass, G.V., McGaw, B., and Smith, M.L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

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

Patton, M.Q. (1978). *Utilization-focused Evaluation.* Beverly Hills, CA: Sage.

Patton, M.Q. (1981). *Creative Evaluation.* Beverly Hills, CA: Sage.

Pillemer, D.B. and Light, R.J. (1980). Synthesizing outcomes: How to use research from many studies. *Harvard Educational Review, 50,* 175-195.

Reichardt, C.S. and Cook, T.D. (1980). Paradigms Lost: Some Thought on Choosing Methods in Evaluation Research. *Evaluation and Program Planning, 3,* 229-236.

Rich, R.F. (ed.) (1979). *Translating Evaluation into Policy.* Beverly Hills, CA: Sage.

Rich, R.F. (1981). *The Knowledge Cycle.* Beverly Hills, CA: Sage.

Rossi, P.R., Freeman, H.E., and Wright, S.S. (1979). *Evaluation: A systematic approach.* Beverly Hills, CA: Sage.

Rutman, L. (1980). *Planning Useful Evaluations.* Beverly Hills, CA: Sage.

Scherill, S. (1984). Toward A Coherent View of Evaluation. *Evaluation Review, 8,* 443-466.

Weiss, C.H. (1972). *Evaluation research: Methods of assessing program effectiveness.* Englewood Cliffs, NJ: Prentice-Hall.

Weiss, C.H. and Bucuvalas, M.J. (1980). *Social Science Research and Decision Making.* New York: Columbia University Press.

Wholey, J.S. (1977). Evaluability Assessment, in L. Rutman (ed.), *Evaluation Research Methods: A Basic Guide.* Beverly Hills, CA: Sage, pp. 42-46.