

Practical Assessment, Research & Evaluation

A peer-reviewed electronic journal.

Copyright is retained by the first or sole author, who grants right of first publication to *Practical Assessment, Research & Evaluation*. Permission is granted to distribute this article for nonprofit, educational purposes if it is copied in its entirety and the journal is credited. PARE has the right to authorize third party reproduction of this article in print, electronic and database forms.

Volume 8, Number 20, October, 2003

ISSN=1531-7714

Evaluability Assessment: A Primer

Michael S. Trevisan and *Yi Min Huang*

Washington State University
Pullman, Washington

Conducting evaluations of programs that are useful to decision makers is the hallmark of successful evaluation. Appropriate program implementation and operation are critical to this work. A strategy that can be used to determine the extent to which a program is ready for full evaluation, is known as evaluability assessment. Initially developed by Wholey (1979), evaluability assessment (EA) seeks to gain information from important documents and input from stakeholders concerning the content and objectives of the program. Outcomes from EA include clear objectives, performance indicators, and options for program improvement. Wholey (1979) recommended EA as an initial step to evaluating programs, increasing the likelihood that evaluations will provide timely, relevant, and responsive evaluation findings for decision makers.

This paper has two purposes. The first is to increase awareness among policymakers and practitioners for the power and utility of EA, particularly at the state and local level. To this end, the background and rationale for this evaluation strategy is documented. The second is to support and promote its use. The article provides an outline of the procedures for conducting effective EA. While there are detailed resources available (e.g., Nay & Kay, 1982; Smith, 1989; Wholey, 1983), we provide a simplified, accessible presentation of the EA procedure and framework. An example is also provided. Issues and benefits are discussed.

Background and Rationale

The 1960s marked the development of many federal programs to address national problems associated with such issues as poverty, education, transportation, housing, and health care. To provide feedback to program managers concerning the effectiveness of programs in their charge, and to policymakers regarding their policy choices, program evaluations were conducted across a wide variety of agencies and programs. What quickly became apparent, however, was that all too often evaluations provided little in the way of useable information for both policymakers and program personnel. Impact statements from these evaluations were shallow, as there was little effect shown from instituting programs designed to address important national issues. As a result, some began to question the wisdom of evaluation, and whether these procurements were worth the expenditure of additional tax dollars.

Equally apparent to evaluators were the challenges and problems inherent in conducting evaluations. The complex policy and management environment in which programs are developed and administered created uncertainty with regard to program objectives, allocation of resources, and type of information needs (Wholey, 1979). Programs often stated goals that could not be measured or were irrelevant, had no apparent logic that connected program resources and activities to stated outcomes, and were inflexible to program modification for unstated political or ideological reasons (Jung and Schubert, 1983). Even identifying the government official(s) ultimately responsible for a particular program was sometimes problematic.

In short, evaluators were caught between the information needs of policy makers on the one hand and the political environment in which programs operate on the other. Thus, evaluations often reported findings with little utility to stakeholders or decision makers.

In response, Wholey (1979) developed EA as a procedure and method to determine whether or not programs were ready for evaluation. According to Wholey (1979):

Evaluability Assessment explores the objectives, expectations, and information needs of program managers and policy makers; explores program reality; assesses the likelihood that program activities will achieve measurable progress toward program objectives; and assesses the extent to which evaluation information is likely to be used by program management. The products of evaluability assessment are: (1) a set of agreed-on program objectives, side effects, and performance indicators on which the program can realistically be held accountable; and (2) a set of evaluation / management options which represent ways in which management can change program activities, objectives, or uses of information in ways likely to improve program

Wholey (1979) argues that when used wisely, EA can save scarce evaluation resources. This is done through the EA process by recommending evaluation only when programs are ready, and targeting evaluation resources for essential evaluation needs.

EA was originally developed to support summative evaluation. EA has also proved useful in developing the program itself by clarifying goals and objectives and establishing a program theory; that is, identifying a reasonable model of the program so that one can ascertain whether or not the attainment of specified outcomes is plausible (Wholey, 1987).

EA has been effective in enhancing program improvement. This is accomplished by developing a shared understanding about the purpose of the program among key stakeholders and program actors, and preparing key personnel to use evaluation results (Smith, 1989). Inclusion of stakeholders in particular, tends to foster buy-in, increase the likelihood that the program will be developed and refined to meet real needs, and helps to ensure that the evaluation will be shaped to obtain findings that inform important concerns.

Several agencies within the federal government used EA extensively during the late 1970s and early 1980s. Since that time, EA has been used sparingly. Rog (1985) and Smith (1989) argue that this phenomenon may in part be due to the fact that major proponents of EA in the federal government (e.g., Wholey) are no longer working in federal agencies and as a consequence, its use has decreased. Nevertheless, many prominent evaluators continue to maintain that EA is an indispensable (albeit underutilized) method and process for the development and evaluation of programs (e.g., Scriven, 1997; Smith, 1989; Wholey, 1987).

Evaluation today has taken center stage in deliberation, legislation, funding, and management of almost all federal and state programs. The recent reauthorization of the elementary and secondary education act (No Child Left Behind (NCLB) Act of 2001), for example, calls for strong evaluation and accountability requirements of programs and services. These requirements include use of scientifically based research, assessment of objective data, the establishment of a set of performance measures, and a focus on results. These requirements underscore not only the need to establish and implement quality programs, but also to ensure that programs are ready for the evaluation expectation. Thus, the time may be right to re-consider the benefits of EA and establish EA as a priority practice when evaluating programs.

NCLB legislation also maintains a new role for states in the management and evaluation of programs. For instance, the 21st Century Community Learning Centers Program, designed to provide coordinated after school programs, is an example of a program for which states are required to take a stronger role in management and evaluation (Harris & Little, 2003). States must provide systematic evaluation and technical assistance activities to local projects. Also under the rubric of state activity and responsibility, local projects must conduct periodic evaluations and use the findings to strengthen and improve project performance measures established by the state.

The new requirements place a premium on planning useful evaluations. Thus, as states work to meet this new evaluation challenge and struggle to establish frameworks for evaluation under tight budget constraints, EA may become an essential strategy in the repertoire of state evaluation.

Method and Process

Conducting EA can be complex, require understanding of the political and organizational context in which the program resides, and attention to subtleties in the data. Some have referred to EA as a process only and recommend against specifying steps. Smith (1989) suggests that EA has evolved into: (a) an evaluation tool that can be used to understand stakeholder awareness of the program, its components, goals and objectives, and what is needed to obtain stated outcomes; and (b) a program development tool. As such, Smith (1989) argues that providing methodological steps for users is needed to simplify the process and make EA more accessible to a variety of potential users.

The most recent set of steps is offered by Smith (1989). The steps are not meant to suggest that EA is a lock-step linear process. Depending on the context and purpose, some steps could be omitted or re-ordered. Nevertheless, the steps are offered to potential users to improve their understanding of the often complex undertaking inherent in EA and a means to start the process. These steps, with short explanations, are as follows:

Step 1: *Determine Purpose, Secure Commitment, and Identify Work Group Members.* A clearly articulated purpose will help secure buy-in and foster commitment. Seven to nine members are recommended with representation from important stakeholder groups, and program administration.

Step 2: *Define boundaries of Program to be Studied.* This step sets limits on EA work and further clarifies the purpose of the EA and role of the team. Boundaries may vary based on such factors as geographic location or program objectives. Boundaries might also be constructed to focus the EA on a program component(s) rather than the entire program.

Step 3: *Identify and Analyze Program Documents.* Documents could include legislation authorizing a program, grant applications, evaluations, audits, and internal memoranda. Documents provide a sense of the

intent of the program as well as what is actually occurring. They can provide a sense of the underlying politics surrounding the program. Secure permission to examine documents as soon as possible.

Step 4: *Develop/Clarify Program Theory*. Developing a program theory is fairly straightforward (see Wholey, 1987, for an extended treatment). Identifying assumptions and values, available resources, program activities, objectives, and how these components relate to one another to produce outcomes, are the major features of developing a program theory.

Note, that the theory depicts a “logic” of how components interact to produce outcomes, and shows performance indicators for the objectives.

The simplicity of developing a program theory can be a seductive feature of EA, particularly for those seeking a straightforward representation of the program. EA users are cautioned against oversimplification of program reality or using depictions of program theory as all encompassing illustrations of the program. As Smith (1989) suggests, programs are often recursive rather than linear; contain vague and illusive assumptions, values, and expectations rather than clear connections between activities and outcomes; and may reasonably contain multiple representations of program theory rather than one.

The important point in developing a program theory is to construct a reasonable depiction of how a program works so that the plausibility of the model can be assessed.

Step 5: *Identify and Interview Stakeholders*. Identification of key stakeholders is critical for program survival as they can provide insights and support for program continuation. Interviews should focus on what stakeholders know and perceive to be true about the program. It is a good idea to develop interview guides. Careful selection and training of interviewers is also recommended.

Step 6: *Describe Stakeholder Perceptions of Program*. Descriptions and comparisons of stakeholder perceptions is important for further understanding of the program.

Step 7: *Identify Stakeholder Needs, Concerns, and Differences in Perceptions*. Differences in perception, needs, and concerns can indicate misperceptions of the program and intent, or a program that is not sufficiently meeting the needs of one or more stakeholder groups.

Step 8: *Determine Plausibility of Program Model*. Data from program staff, documentation, and stakeholder interviews are used to determine plausibility of the program. That is, data are analyzed to determine the extent to which the program is properly implemented, sufficiently developed, and activities appropriate, to reasonably predict that desired outcomes will be met.

Step 9: *Draw Conclusions and Make Recommendations*. The EA team makes conclusions and recommendations. Conclusions and recommendations are drawn from the data. EA teams are encouraged to guard against validity threats, such as personal bias.

Step 10: *Plan Specific Steps for Utilization of EA Data*. The next step might be to continue with an evaluation of the program, revise the program, or that no action be taken.

An Example

A recent example illustrates the utility of EA for state level program development and evaluation priorities. Youtie, Bozeman, and Shapira (1999) document the use of EA for the evaluation of the Georgia Research Alliance (GRA). As stated by Youtie, Bozeman, and Shapira (1999), “The GRA is a collaborative initiative among six research universities in Georgia to use research infrastructure invested in targeted industry areas to generate economic development results” (p. 58). These investments are in various technologies such as advanced telecommunications, environmental technologies, and human genetics. Prominent faculty in these fields are recruited to Georgia universities, given supplemental funding from GRA, with the idea that these faculty will establish research collaborations with industry and develop commercially viable products. While the program has been in operation for several years, no formal evaluation plan was in place, wide agreement on objectives was not apparent, nor was there a common framework to understand how various components of the program function.

As part of Georgia’s performance-based budgeting process, the authors were asked to develop a possible evaluation plan with methods and strategies for GRA that in turn would be adopted and implemented by an external evaluator. Thus, EA was initiated.

As an initial task, the EA team conducted a literature review of statewide evaluation practices for technology development programs. From this literature, 11 different evaluation methodologies were identified. The strengths and limitations of each were articulated, particularly for research and development programs, such as GRA.

The EA team conducted interviews of important stakeholders, including key faculty members, university

administrators, and business executives dealing with the industry-university partnerships in relevant technology areas. The findings showed that these stakeholders differed in their perceptions of the program, sometimes dramatically. Universities, for example, saw GRA as a means to increase research productivity while businesses viewed GRA as a means to develop marketable products.

The EA team also discovered challenges in documenting outcomes due to the nature of the GRA enterprise. For instance, a time lag of 7-15 years exists between the start of a partnership and when significant results occurred. A complex set of factors influence GRA such that, links between investments and outcomes were found to be indirect. And other factors outside the control of the partnership, such as economic cycles, impact the attainment of outcomes.

With this knowledge, the EA team developed an evaluation plan that maximized the strengths of various evaluation methodologies obtained from the initial literature review and that accounted for program characteristics and dynamics discovered through the EA process. This plan included a periodic assessment to meet the needs of stakeholders and program management, and a comprehensive evaluation for policy makers.

In short, the EA process was used to develop an optimum evaluation plan for a program already in operation. This increased the viability of the evaluation and in turn, ensured that policy makers, stakeholders, and program participants, received timely and useful evaluation findings.

Common Issues

There are two common issues when conducting EA that can be problematic if not recognized and properly managed. First, EA is typically conducted by a team. As mentioned, the ideal team is composed of members from stakeholder groups, program implementers, and administration. This ensures representation by broad program constituency and builds into the process, a comprehensive view of the program. If the group does not function well however, the integrity of the EA is undermined and the exercise can be costly. Thus, building group cohesion at the outset will likely payoff with an efficient and productive EA process.

Second, EA can be time consuming. Time is consumed because of scheduling conflicts, difficulty in gaining commitments from key stakeholders, or because program documentation is unorganized. Detailed planning, fair distribution of the workload, and supervision of all tasks and activities, are strategies that can be employed to help control the amount of time spent on EA.

Benefits

A successful EA can lead to more appropriate and realistic outcomes, stable program implementation, and a viable evaluation. Smith (1989) and Wholey (1979) maintain additional benefits to EA that are worth noting for those interested in adopting the process and method. These include: (a) the ability to distinguish between program failure and evaluation failure, (b) accurate estimation of long term outcomes, (c) increased investment in the program by stakeholders, (d) improved program performance, (e) improved program development and evaluation skills of staff, (f) increased visibility and accountability for the program, (g) clearer administrative understanding of the program, (h) better policy choices, and (i) continued support.

In summary, EA is a method and process designed to increase the probability that evaluations will be timely, relevant, and responsive. Investment in EA before an evaluation is conducted is a cost effective strategy to increase the quality of program implementation, conserve evaluation resources, and target those resources to essential evaluation needs. States in particular, are urged to consider EA as an initial strategy in meeting their new role in management and evaluation of programs.

References

- Harris, E., & Little, P. (2003). Evaluating the 21st century community learning centers program – A view from the states. *The Evaluation Exchange*, 9(1), 6-7.
- Jung, S. M., & Schubert, J. G. (1983). Evaluability assessment: A two-year retrospective. *Educational Evaluation and Policy Analysis*, 5(4), 435-444.
- Nay, J. N., & Kay, P. (1982). *Government oversight and evaluability assessment*. Lexington, MA: Lexington Books.
- No Child Left Behind Act of 2001, Public Law No. 107-110, 115 Stat. 1425 (2002).
- Rog, D. (1985). *A methodological analysis of evaluability assessment*. Doctoral Dissertation, Vanderbilt University, Nashville, Tennessee.
- Scriven, M. (1991). *Evaluation thesaurus* (4th Ed.). Newbury Park: Sage Publications.
- Smith, M. F. (1989). *Evaluability assessment: A practical approach*. Clemson: Kluwer Academic.

Wholey, J. S. (1979). *Evaluation: Promise and performance*. Washington, DC: The Urban Institute: Author.

Wholey, J. S. (1983). *Evaluation and effective public management*. Boston: Little, Brown, & Co.

Wholey, J. S. (1987). Evaluability assessment: Developing program theory. In L. Bickman (Ed.), *Using program theory in evaluation*. New Directions for Program Evaluation, No. 33. San Francisco: Jossey-Bass.

Youtie, J., Bozeman, B., & Shapira, P. (1999). Using an evaluability assessment to select methods for evaluating state technology development programs: The case of the Georgia Research Alliance. *Evaluation and Program Planning*, 22(1), 55-64.

Author Contact

Michael S. Trevisan, ELCP
P.O. Box 642136
Washington State University
Pullman, WA 99164-2136

email: trevisan@mail.wsu.edu

Descriptors: Evaluation Methods; Evaluation Research; Program Evaluation; Evaluability Assessment

Citation: Trevisan, Michael S. & Yi Min Huang (2003). Evaluability assessment: a primer. *Practical Assessment, Research & Evaluation*, 8(20). Available online: <http://PAREonline.net/getvn.asp?v=8&n=20>.