

Program Development and Evaluation



Developing a Concept of Extension Program Evaluation

Mohammad Douglah

What is evaluation?

Those who have studied the concept of evaluation define it in different ways. Below are some of their conclusions.

The term **evaluation** refers to the activity of systematically collecting, analyzing and reporting information that can then be used to change attitudes or to improve the operation of a project or program. The word **systematic** stipulates that the evaluation must be planned (Alkin 1990).

Evaluation is the assignment of worth or value according to a set of criteria and standards, which can be either explicit or implicit (House in Alkin 1990).

Evaluation is the process of providing reliable, valid, relevant and useful information to decision makers about the operation and effects of social programs or other institutional effects (McLaughlin in Alkin 1990).

Evaluation includes the following elements:

1. Systematic collection of information
2. For use by specific, identifiable people or groups for the purposes of
3. Making decisions about and/or improving program effectiveness (Patton in Alkin 1990).

Evaluation does not aim to replace decision makers' experience and judgment, but rather offers systematic evidence that *informs* experience and judgment (Weiss in Alkin 1990).

Evaluation refers to the process of determining the merit, worth or value of something, or the product of that process (Scriven 1991).

The systematic examination of events occurring in and consequent on a contemporary program—an examination conducted to assist in improving this program and other programs having the same general purpose (Cronbach 1980).

What does evaluation mean to you?

If ten extension stakeholders were asked what program evaluation meant to them, there would likely be ten different responses. This is understandable. Program evaluation means different things to different people.

To some, it means determining the extent to which the extension program is achieving the goals and objectives it set out to achieve. To others, it means judging the merit or worth of a program. Still others may think of program evaluation as providing elected officials, influential members of the community, extension administrators and customers information that can help them make important decisions about extension's present and future status. Other people's views of program evaluation may be shaped by their belief that it is essentially just "going through the motions" since important program decisions are usually not based on evaluation, but on other considerations such as political expedience.

These and many other responses to the question "What is program evaluation?" are partially correct. However, taken singly, they do not portray a complete and accurate picture of program evaluation. In fact, there is no all-inclusive definition. There are many ways of conceptualizing evaluation: its nature, philosophy, purpose, function or methodology. Most definitions place more emphasis on one or more of the above dimensions than on others. For example, the following two definitions highlight the nature of the evaluation process:

- Evaluation is the act of comparing what should be (criteria, standards, goals or objectives) with what is (evidence, data, information) for the purpose of ascertaining the worth or value of what is being evaluated (Steele, 1991).
- Evaluation involves the identification, clarification, and application of defensible criteria to determine an evaluation object's value (worth or merit), quality, effectiveness, or significance in relation to those criteria (Worthen, Sanders and Fitzpatrick, 1997).

In contrast, these two definitions emphasize the purpose and function of evaluation:

- Evaluation is the process of delineating, obtaining, and providing useful information for judging decision alternatives (Stufflebeam, 1973a).
- Evaluation is a process used to determine whether the design and delivery of a program were effective and whether the proposed outcomes were met (Caffarella, 1994)

Just as there is no one definition for evaluation, there is no single correct way to evaluate that is best suited for all situations. It is likely that a particular evaluation approach using certain methods and techniques will provide better answers to specific questions than some other approaches.

What is important for extension professionals to realize is that program evaluation is more than simply filling out post-activity reaction questionnaires (though these have a definite and useful place in the larger and more comprehensive effort of program evaluation). It is more useful to think of evaluation as a continuous process of inquiry—a process of asking questions about social, economic and environmental conditions and circumstances within which the extension program is being developed. Such questions help in assessing the needs, goals and objectives the program is attempting to achieve (known as **context** or **planning** evaluation).

It also includes asking questions about how the program is being designed and delivered to its intended audiences (**input** or **process** evaluation). And last, but not least, it means asking questions about the extent to which the program is producing desired results (**outcome, product** or **impact** evaluation).

It is equally important for extension professionals to realize that developing a genuine understanding of program evaluation starts with examining one's attitude toward the idea. Do we accept evaluation as a useful activity which is an integral part of our total programming mission, or do we consider it something we do in response to organizational demands? Do we believe evaluation can be approached as a learning process that can help us grow and develop personally and professionally, in addition to satisfying the public's demands for accountability? Or do we view it as a waste of time?

We suggest that program evaluation is first and foremost a state of mind, a readiness to subject programming decisions and practices to careful scrutiny so that successes and achievements can be celebrated and built upon, and failures and shortcomings can be corrected and overcome.

Informal vs. formal evaluation

On the job, extension professionals are continuously engaged in making quick gut-level decisions about the worth or value of their programming efforts. Such decisions are likely to be the outcome of more informal evaluations and, oftentimes, they are sufficient for gaining useful personal insights. However, other extension stakeholders, such as public officials, donors and administrators expect more formal program evaluations because they are generally assumed to be more accurate and objective. It is important to remember that an evaluation's formality is a matter of degree. Program evaluation is more, or less, formal to the extent that evaluators:

- Rely on criteria (standards) which clearly, explicitly and comprehensively indicate the expectations (goals and objectives) of the program.
- Apply systematic procedures in gathering trustworthy evidence (data, information) which is relevant to the goals and objectives of the program.
- Weigh the evidence against the criteria carefully and objectively in the process of determining the programming effort's value.

For your program evaluations to be credible, they need to be judged as more formal than informal; that is, when you decide what aspect of your program you wish to evaluate, you take time to articulate specifically what it is you want the evaluation to reveal.

The informal/formal evaluation continuum

Informal ←	→ Formal
Casual	Deliberate
Spontaneous	Takes time
Criteria are implicit	Criteria are explicit & clear
Evidence is quickly & insufficiently ascertained	Evidence is systematically collected
Judgment is hastily made	Judgment is carefully made

You also need to identify the necessary information, the source(s) that will provide it and the methods and tools that will be used to gather it. Finally, you need to plan for analyzing and reporting the information in light of what you want the evaluation to accomplish. The more resources (time, energy, money) you invest, and the more critical thinking processes you employ, the more formal, and consequently more credible, your evaluation will be.

Why evaluate?

There are perhaps as many reasons for evaluating educational programs as there are definitions of evaluation. However, a survey of rationales for conducting evaluations indicates that the reasons can be grouped into two categories: 1) those that emphasize the value of evaluation as an instrument to *improve* programs; and 2) those that point to its use as a means to *prove* the effectiveness of programs. Michael Scriven (1967), coined the terms **“formative”** and **“summative”** evaluation to refer to the former and latter, respectively.

Some evaluations are carried out with the primary aim of enhancing the quality of educational programs or activities—not to demonstrate their usefulness, cost-effectiveness or the need for their continuation to skeptical stakeholders.

When a family living agent, for example, evaluates the effectiveness of a newsletter for the purpose of deciding whether certain changes are needed to enhance its usefulness as an educational tool, she undertakes a formative evaluation. However, when the same newsletter is evaluated because the county staff and some members of the county extension committee are asking questions about its usefulness and whether it should be continued, the evaluation is being conducted for summative purposes.

Useful as these two types of evaluation may be in distinguishing between the different forces that drive evaluation, they are, nevertheless, not mutually exclusive. Results of formative evaluation can be used for such purposes as accountability and gaining public support and visibility, which are usually associated with summative evaluation.

Likewise, summative evaluation can yield information that can be utilized to change the focus of the program, establish more collaborative relationships with other agencies, or employ more up-to-date technologies in program delivery—purposes usually associated with formative evaluation. However, when the original intent is primarily to improve a program, the evaluation is formative, and when the intent is to prove the worth and value of a program, the evaluation is summative. That is why formative evaluation usually examines the context, inputs, strategies or methodologies of programs, while

summative evaluations focus on outcomes, benefits or impacts of programs.

Is evaluation applied research?

To some people, especially academics, making distinctions between basic research, applied research and evaluation is important. To them, basic research is different from applied research, and both are different from evaluation. Academics think of researchers as scientists—basic researchers responsible for generating “new” knowledge that helps us better understand phenomena, while applied researchers help make things work better and more efficiently.

Evaluators, on the other hand, are viewed as expert methodologists who apply the tools of research to answer specific questions about the quality of what they evaluate. Worthen and Sanders (1987) differentiate between research and evaluation on the basis of these eight criteria:

1. Motivation of the inquirer (researchers are curious; evaluators are concerned)
2. Objectives of the inquiry (research seeks conclusions; evaluation leads to decisions)
3. Autonomy of the inquirer (researchers are independent; evaluators work on behalf of clients)
4. Role of explanation (research is done to answer “why” questions; evaluation probes “what” questions)
5. Rigor of inquiry (research employs more controls to enhance prediction)
6. Properties of phenomena being studied (evaluation assesses value of a thing; research generates scientific knowledge)
7. Generalizability of phenomena being studied (aim of research is to generalize across time and space; evaluation assesses the value of one thing at a time)
8. The contribution to law making (the ultimate goal of research is to derive laws that explain relationships between variables; evaluation describes the workings of what is being evaluated to assess its quality).

Other people don’t see the distinctions between research and evaluation as being so clear cut, especially those between applied research and evaluation. They contend that just because evaluators are concerned about phenomena doesn’t necessarily mean they are less curious; that researchers are not as autonomous as some people think they are; that sound evaluations don’t stop at describing; that accurate evaluations employ as rigorous

methodologies as research. Finally, when the phenomenon being examined relates to some aspect of human behavior, it is extremely difficult, if not impossible, to conduct the inquiry in a manner that accounts for all contextual variations so that one can generalize and derive laws.

The most useful distinction between research and evaluation relates to the type of phenomena being studied and the purpose of the inquiry. For example, if the phenomena we want to study is the relationship between participation in extension and learning styles of participants, and the purpose is to make a contribution to a theoretical body of knowledge about the educational behavior of adult learners, the resulting inquiry is more likely to be labeled research. On the other hand, if the inquiry focuses on comparing participation in an extension program through distance education with participation through traditional modes of contact (for example face-to-face) for the purpose of determining learning outcomes and cost-effectiveness, the comparative study will likely be viewed as evaluation. What really matters is to recognize that just as the quality of research can be placed on a continuum, the same can be said about evaluation.

Research and formal evaluation are both modes of inquiry. While employed to examine different phenomena for different purposes, they are both subject to the same standards of excellence if their findings, conclusions and recommendations are to be taken seriously.

Some ethical considerations

Because the practice of evaluation has become more widespread in education and other fields, the American Evaluation Association deemed it necessary to develop standards to guide the profession (Joint Committee on Standards for Educational Evaluation 1994). One of these standards¹ addresses the issue of propriety. Propriety refers to evaluation being conducted legally, ethically and with regard for the welfare of those involved in the evaluation as well as those affected by its results. Royse (1992) summarizes the practices that ensure the preservation of the dignity and welfare of all people involved in or affected by the evaluation:

- Evaluation subjects must be volunteers. Subjects must understand that they are freely participating in the evaluation and that they can withdraw from it at any time without penalties or loss of program services or benefits, now or in the future.
- Potential subjects should be given sufficient information about the evaluation, including its purpose, duration, procedures, benefits (if any), and possible risks or discomforts, to enable them to freely choose whether or not to participate.
- No physical, emotional or psychological harm should result as a consequence of participation in the evaluation.
- All sensitive information about subjects should be protected. The privacy of human subjects is protected by:
 - a. allowing subjects to respond **anonymously**, if at all possible;
 - b. where complete anonymity is not possible (for example, because pretest and post-test scores have to be matched for each subject), protection is provided by separating any personally identifying information from the evaluation data through the use of numeric or other special codes.
- The privacy of human subjects is further protected by not collecting or reporting any personal information (names, addresses, telephone numbers, or personal descriptors such as Mayor, Eminence, Police Chief) that would result in evaluation subjects being identified by persons other than the evaluator.

¹ The other three standards are: 1) *utility*—an evaluation serves the information needs of the intended users; 2) *feasibility*—an evaluation will be realistic, prudent, diplomatic and frugal; and 3) *accuracy*—an evaluation will reveal and convey technically adequate information about features related to the worth or merit of the program being evaluated.

- Subjects should be assured that their views and opinions, likes and dislikes, suggestions and recommendations will be treated **confidentially** and that only people closely associated with the evaluation will have access to the data.

Extension professionals have an obligation to observe the above guidelines in all their evaluation endeavors. Some evaluators believe it is a good practice to accompany each request for information with a consent form that clearly explains the above guidelines and asks subjects to sign and return requests to the evaluator.

The philosophical perspective

Michael Patton (1995), addressing the topic “Political, Ethical, and Interpersonal Aspects of Evaluation” to the national audience of a satellite course on evaluation, pointed out that one of evaluation’s main purposes is to serve as a check on reality. He went on to explain how reality is not absolute and one-dimensional, but rather relative and subject to multiple interpretations. For example, when we conduct a program evaluation as extension professionals, we, in essence are attempting to compare our perception of the reality of the program, (our informal evaluation) with the perception of our stakeholders (through a formal evaluation).

The point Patton was making is that program evaluation is greatly influenced by the philosophical beliefs of those who do the evaluation. The decisions they make in regard to what to evaluate, how to evaluate it, who should be involved in the evaluation, and other decisions, are a reflection of their conscious and unconscious beliefs about a number of evaluation issues. Examples of such issues include:

1. Whether program participants and other stakeholders should have a role in the design, conduct and use of program evaluation (**participatory** evaluation) or whether evaluations should be undertaken by trained specialists (**expert-based** evaluation).
2. Whether information collected for evaluation should consist of what is tangible, observable and measurable, (numbers of people, businesses or communities adopting certain practices; increases in family income; decrease in crime rate—**quantitative** data), or whether it should also address program characteristics or outcomes that cannot be easily quantified and tabulated, such as collaboration with other agencies and programs, or feelings of empowerment and self-esteem (**qualitative** information).
3. Whether evaluation is an activity that can be conducted by neutral and objective evaluators (**value-free** evaluation) or whether it is always affected by personal, professional and political beliefs and values (**advocacy** evaluation).

The way you respond to these and other evaluation issues forms your philosophy of evaluation and its role in the extension programming process. Whatever your philosophy, it is best to subject it to critical reflection and potential modification every time your program undergoes an evaluation.

References

- Alkin, Marvin C. (1990). *Debates on Evaluation*, Newbury Park, Sage Publications.
- Cronback, Lee J. & Associates (1980). *Toward Reform of Program Evaluation*. San Francisco, Jossey-Bass.
- House, Ernest R. (1993). *Professional Evaluation*. San Francisco: Sage Publications.
- Patton, Michael Q. (1995). Political, ethical, and interpersonal aspects. Lecture delivered as part of agricultural satellite course on Program Evaluation in Adult Education and Training. S. Kay Rockwell, course instructor, University of Nebraska–Lincoln.
- Royse, David. (1992). *Program Evaluation*. Chicago: Nelson-Hall Publishers.
- Scriven, M. (1991). *Evaluation Thesaurus*, 9th Ed. Newbury Park, Sage Publications.
- Scriven, M. (1973). The methodology of evaluation. In B. R. Worthen and J. R. Sanders *Educational Evaluation: Theory and Practice*. Belmont, CA:Wadsworth.
- Steele, Sara M. (1991). Introduction to Evaluation. Class notes, Department of Continuing and Vocational Education: UW-Madison.
- The Joint Committee on Standards for Educational Evaluation. (1994). *Educational Evaluation: Standards*. San Francisco: Sage Publications.
- Worthen, Blaine R. And James R. Sanders. (1987). *Educational Evaluation: Alternative Approaches and Practical Guidelines*. New York: Longman.

Author: Mohammad Douglah is a professor of continuing and vocational education and a faculty associate with the Department of Agricultural Journalism at the University of Wisconsin–Madison and the University of Wisconsin–Extension, Cooperative Extension.

Issued in furtherance of Cooperative Extension work, Acts of May 8 and June 30, 1914, in cooperation with the U.S. Department of Agriculture, University of Wisconsin–Extension, Cooperative Extension. University of Wisconsin–Extension provides equal opportunities in employment and programming, including Title IX and ADA requirements. If you need this information in an alternative format, contact the Office of Equal Opportunity and Diversity Programs or call Extension Publishing at (608)262-2655.



© 1998 by the Board of Regents of the University of Wisconsin System doing business as the division of Cooperative Extension of the University of Wisconsin–Extension. Send inquiries about copyright permission to: Director, Cooperative Extension Publishing, 201 Hiram Smith Hall, 1545 Observatory Dr., Madison, WI 53706.

You can obtain copies of this publication from your Wisconsin county Extension office or from Cooperative Extension Publications, Room 170, 630 W. Mifflin Street, Madison, WI 53703, (608)262-3346. Before publicizing, please check on this publication’s availability.

**G3658-7 Program Development and Evaluation,
Developing a Concept of Extension Program Evaluation**

I-4-98-.5M-200