an emerging concept of program evaluation

Sara Steele

For the past few years, Extension, like some other areas of education, has used a limited definition of evaluation—that of examining the results of a project or even more specifically, determining whether a project met its objectives. In other fields, demands for program evaluation in the late 1960s found that such a concept wasn't sufficient and new ideas and new evaluation frameworks have emerged.

One of the most useful of the merging concepts is that of program evaluation as a specific and unique type of evaluation. It includes much more than the prevailing 1960s concept.

This article addresses itself to the concept of program evaluation and its differences and similarities to the prevailing concept of evaluation. It discusses how program evaluation is (1) different from project evaluation, (2) different from program research, (3) a process rather than a procedure, (4) more than examining the attainment of objectives, (5) more than evaluating the results of a program, (6) a management tool, and (7) people-centered.

This article won't provide you with a procedure for doing program evaluation. It will expand your understanding of your concept of what program evaluation is so that you can devise the procedures most useful to you in your programming situation.

Different from Project Evaluation

Project evaluation is concerned primarily with a specific project or program activity. It controls and improves the specific programmer-participant relationship. *Program* evaluation is concerned with the additive effects of a series of instructional components. What are the combined results of a farmer attending several meetings held by Extension? *Project* evaluation is usually

Sara Steele: Professor and Coordinator, Division of Program and Staff Development, University of Wisconsin-Extension.

most concerned with knowledge, skills, and attitude change. *Program* evaluation is concerned with the impact that those changes and program participation have on the person and those he's in contact with. *Project* evaluation is more apt to deal with how the program satisfies the specific needs of individual learners; *program* evaluation is more apt to deal with how the program meets the needs of the community, or a subsection of society.

Program evaluation includes project evaluation, but deals with additional things. It's concerned, for example, with establishing priorities among projects. It's concerned with whether adequate resources are being applied to the right programs. It deals with such issues as whether to be all things to all people, or whether to concentrate resources in certain priority thrusts. It's concerned with the extent to which the agency is carrying out its mission.

Program evaluation doesn't substitute for project evaluation; and project evaluation doesn't substitute for program evaluation. It's possible to do a good job of project evaluation—without ever really doing program evaluation.

Different from Program Research

Program evaluation and program research use similar methods. Both are important in programming. The difference between them is important in understanding the best procedures to use in program evaluation.

Program research looks for new and generalizable knowledge. It's trying to formulate theories and principles of programming. If any benefits occur for the program that is the focus of that research, those benefits are peripheral. The research isn't being carried out to help the program and its participants.

Program evaluation deals with old questions about programs. It's trying to get sound and reliable information to use within the context of a specific program. Its primary audience is individuals closely involved with that program or those responsible for funding or supporting the specific program.

Instrumentation and statistical processes must be tested for validity and reliability as tools in evaluation rather than automatically transferred. Considerable adaptation may be needed. For example, with audiences with varied needs, average gain of the group may be less meaningful than numbers of people achieving specific things important to them. When conclusive proof isn't needed, a .25 or .10 rather than a .05 level of statistical significance may be enough. If 25 out of 100 times the finding is apt to occur because of something other than chance occurrence, this may be strong enough to make some decisions about programs.

From one perspective, data-gathering activities are separate from evaluation. They are essential to it, but can be considered as prior and prerequisite to the actual act of evaluation. Data in and of themselves give few answers. Data

must be accurately and immediately interpreted within a specific context. Interpretation rests with people. Wisdom and experience in forming and using criteria, in assessing the limitations of data and the potential consequence those limitations pose, and finally in making, communicating, and defending judgments, are more important skills in evaluation than are the skills of data gathering and statistical analysis. Scientifically produced data are a valuable input in evaluation, but shouldn't stand alone as the output from it.

The most important judgment to make in evaluation, and the most difficult one, is determining the value of the program.

Criteria are the basic organizing framework of evaluation as hypotheses are in research. Criteria indicate what information is to be presented, organized, and interpreted. Program criteria must be probed for clarity, reality, and the extent to which they portray the most valuable things in programs and programming. Programs can appropriately be judged on a wide range of criteria, some of which are in conflict with others. Criteria must be ordered in terms of their importance. For example, which program is more successful—the one that reaches a large number of people with a moderate amount of help to them or the one that gives great help to a very small number?

Criteria are debatable. Confrontation on criteria clears up some of the ambiguities and frustrations that plague programs.

Reports of evaluation must be designed to get action rather than to uphold research tradition. For example, putting conclusions and recommendations at the beginning rather than the end of the report may attract the attention of busy decision makers who don't routinely read through complete research-type reports.

Process Not Procedure

Evaluation is most easily handled and has greatest utility if it's considered a generic term and used as such.

The two most prevalent ways of looking at evaluation as a generalizable process are:

1. As a process of forming judgments about programs using criteria or standards of comparison and descriptions of what occurred and resulted in the program.
2. As a process of using information in comparing alternatives in reaching program decisions.

The first concept places the emphasis on judging and forming conclusions about program activities and results of one program. The second concept emphasizes identifying alternatives and then using evaluation to help choose among those alternatives. The first concept would deal with a question

like, "How important is this program?" Or would examine how effective a given approach is in carrying out the agency's mission. The second would compare two or more approaches to see which is most effective.

More Than Examining Objectives

Although examining the attainment of objectives is still one important aspect of evaluating a program, three major expansions of the concept of evaluating results are occurring.

Sometimes the value or harm of the unanticipated results are more important than the intended objectives attained in the program. Negative as well as positive results must be identified. The person who masters the content of the meeting, but comes to dislike the Extension staff in the process has both a negative and a positive result of the program.

Although results may be judged against what was expected as stated in objectives, many of the broader approaches compare the amount of results produced to: (1) the needs that initiated the program, (2) the kind and amount of results that must be produced if the agency is to attain its mission, (3) broad competency standards, or (4) statements of the kind of results that a program conducted with the particular clientele using a certain amount of input should be expected to produce.

Objectives are symbols or transmittal links to one or more of these broader definitions of results such as meeting a need or carrying out a mission. If objectives are well chosen and directly on target, results judged against the need or mission will also show that the objectives have been attained. Evaluation that's concerned with the overall effectiveness of a program is concerned not only with results in terms of behavioral changes in people, but also with the proportion of the potential clientele that's reached, the balance in types of people reached, the extent to which the results deal with urgent and continual need, and the care with which participant, agency, and societal resources are used.

... it's essential to remember that evaluation is a tool and must return its cost from the way it helps better manage and improve programs and the way it helps your clientele get more results of greater value ....

There are two major processes involved in examining the results of the program—description and evaluation. *Description* provides evidence of what occurred. *Evaluation* produces judgments as to the adequacy of what occurred.

Evaluation involves such sticky questions as: Are the results important? Do they contribute more to the participants and society than if the resources had been invested in other things? Were they produced at a reasonable cost? Are the results sufficient in terms of the overall need? Are they sufficient

in the expectations of the participants and the amount of time and energy they invested? Is there any evidence that it's realistic to expect a program to produce more results than this one has, given the same budget, personnel, and working conditions? If the results are insufficient, does it mean that the program isn't effective or that changes need to be made in the way the program is carried out?

Many so-called "evaluation reports" only provide descriptions in that they stop at presenting statistical evidence of what has occurred without either making the judgments required for evaluation or providing evidence of what judgments others have made after studying the program and its results. Such reports may be useful in creating program image with outsiders. However, evaluation, as well as description, must occur for the information to improve decision making.

The most important judgment to make in evaluation, and the most difficult one, is determining the value of the program. What did participating in a program and using practices learned in the program actually mean to the participants, their community, and society as a whole? Was the benefit produced from their having and/or using the content of the program of greater value that what could have been produced if the program resources had been applied to something else? Evaluation of worth and value isn't satisfied with attainment of objectives or with the fact that sizable results have been produced. It's concerned with the value of these results.

More Than Results Evaluation

Evaluation in Extension has been primarily emphasized and practiced as a *summative* activity in which the results of a program were described. *Formative* evaluation—evaluation conducted *before* and during the program and used to influence the program while it's in progress—is equally important. Formative evaluation emphasizes evaluating the first decisions involved in programming and isn't merely another name for evaluating the means used in programming. Formative evaluation includes evaluating the extent of need.

Evaluating the expected program outcomes to see if they're realistic and focus on the most crucial needs, evaluating the plans for the program to see if adequate inputs are being marshalled and if the activities planned can produce the type of results expected, and monitoring the processes involved to see that they stay on target and that repetition or reinforcement are added to the original plan if such activity is necessary—all are essential if a program is to achieve maximum results. Evaluation that influences ongoing developments in the program has great value—it improves and gives immediate benefits.

Program evaluation as a generic term can include the evaluative activities that focus on either process or product. But the

emerging emphasis is on frameworks that consider both process and product and, more importantly, the interrelatedness of the two. Program evaluation deals with the program as a functioning, producing system. Program evaluation serves both during the actual program operation and afterwards in retrospectively analyzing how the particular level of results was produced. This broader track of program evaluation emphasizes developing and improving programs. Approaches that focus on results are primarily concerned with questions about ending or continuing the program. Both are essential. One doesn't substitute for the other.

Management Tool

Evaluation can be a powerful working tool in programming. It's not an end in itself, something to engage in because it's intrinsically good, but a way to get things done. With the clearer understanding of evaluation's role in decision making comes a better understanding of the value of evaluation in guiding and managing program activities by providing sounder answers to everyday crises and decisions. Evaluation provides a basis for better program choices and for more rapid response to needs for improvement. It can be a tool in improving total operating efficiency as well as providing clientele and society with more effective programs.

Evaluation is natural human process—a continual programming activity rather than an episodic or extensive but infrequent effort. Extension needs to pay more attention to improving the accuracy of this process by introducing more system into the way it's done rather than by replacing it completely with periodic, systematic, formal evaluations.

Data in and of themselves give few answers. Data must be accurately and immediately interpreted within a specific context. Interpretation rests with people.

Evaluation in terms of intuitive judgments does and should go on continually. It should be an automatic part of an adult educator's professional skills. He should form the basis for making accurate judgments as part of gaining professional competence, in the same way a doctor makes accurate diagnoses. Just as the doctor uses lab tests when he feels they're warranted, so the adult educator should know when more extensive, systematic processes are worth their cost. Both natural and systematic evaluation play important roles; neither replaces the other. Both are essential in managing programs.

Evaluation has to be well managed if it's to serve as a management tool. The extent of time invested in evaluation, degree of formality, and objectivity of procedures should be matched with the degree of value the particular exploration

can produce. Evaluation should be designed for the context in which it will be used. What questions must be answered or decisions made? By whom? By what date? A clear understanding of why you're evaluating and what you want to accomplish by that evaluation is essential in effectively using evaluation as a management tool.

People-Centered

If evaluation is to be used, *people* have to use it. Evaluation as input into decision making emphasizes the need of interface and interaction. Evaluation for program improvement recognizes that those who must make the improvements must be actively involved in the evaluation. Criteria come from people and judgment is made by people. Recognizing close relationships between evaluation and politics and policy emphasizes the human element. Even in terms of data, the need for involving a variety of people in interpreting data so that a more complete picture can be secured is becoming more apparent.

For a few years, experts tried to take evaluation out of the hands of program people, hoping to increase objectivity. Lately, however, there's been a return to active involvement of many minds as conclusions are formulated and decisions made. We need ways of increasing objectivity while relating closely to the people involved during the evaluation process.

Many Ways of Doing Evaluation

My recent summary, *Contemporary Approaches to Program Evaluation*, includes more than 50 frameworks. Some focus only on evaluation of attainment of objectives; others concentrate on evaluation of results with emphasis on total outcomes and effects. Still others focus on programs as functioning, producing systems.

In part, the proliferation of approaches represents a groping and a search for relevant models that was triggered when the established approaches to evaluation were found wanting. The number of approaches may decrease as acceptance and new stabilization occurs. On the other hand, the proliferation also represents the awareness that many kinds of evaluation are needed.

The situation can be likened to examining a mountain area. The geographer is interested in topography. The geologist looks for rock samples. The mountain climber, the engineer designing a railroad, the pilot flying over the mountains, and the native going from one valley to another are all getting data about conditions in the mountain range to try to describe, evaluate, and deal with it. But, the data gotten and the approach to getting that information aren't the same.

In many ways, programs are like mountains. They're complex with many planes and facets. These planes and facets can be examined in a variety of ways. Conditions in mountains change just as program situations change. People have varying

---

needs and purposes in relation to describing and evaluating them. Many types of information exist that are useful to those purposes. Therefore, just as there are many ways to describe, analyze, and evaluate a mountain, so are there many ways to examine and evaluate programs. The approach you take depends on your particular needs.

How many approaches should an Extension agent, specialist, or administrator be familiar with? It's much easier to have someone hand you one approach like the Tylerian model as though it were the only one. However, perusal of different frameworks helps to develop a kit of evaluative tools that we can draw on as the need arises in a wide range of situations.

Sometimes the value or harm of the unanticipated results are more important than the intended objectives attained in the program.

As we develop that kit of evaluative tools, it's essential to remember that evaluation is a tool and must return its cost from the way it helps better manage and improve programs and the way it helps your clientele get more results of greater value to them, their families, communities, and society.

Summary   This article has presented a concept of program evaluation that can be extremely useful to you if you find your own ways of operationalizing it. The extensive bibliography that follows gives you clues on how others have operationalized evaluation. However, in addition, you'll want to apply your own common sense to figuring out the soundest procedures for coming to the most important judgments of your program. Evaluation isn't something someone can give you as an easy procedure to follow. You've got to think it through yourself.

Bibliography   Bennett, Claude F. *Theory and Strategy for Evaluation of Knowledge Dissemination*. Washington, D.C.: Extension Service—USDA, Program and Staff Development, 1973.

Borus, Michael E., and William R. Tash. *Measuring the Impact of Manpower Programs: A Primer*. Ann Arbor: The University of Michigan—Wayne State University, Institute of Labor and Industrial Relations, 1970.

Bruce, Robert L. "Monitoring Evaluation." Ithaca, N.Y.: Cornell University, 1972. (Mimeographed)

Byrn, Darcie, ed. *Evaluation in Extension*. Topeka, Kan.: H. M. Ives & Sons, Inc., 1959.

Caro, Francis G., ed. *Readings in Evaluation Research*. New York: Russell Sage Foundation, 1971.

Duft, Ken D. "Systems Planning for Extension." *Journal of Cooperative Extension*, VII (Fall, 1969), 168-78.

Essert, Paul L. *Creative Leadership of Adult Education.* Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1951.

Freeman, Howard E., and Clarence C. Sherwood. *Social Research and Social Policy.* Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1970.

Glass, Gene V. "Educational Product Evaluation: A Prototype Format Applied." *Educational Researcher*, I (January, 1972), 7-10

_____. *The Growth of Evaluation Methodology.* AERA Monograph Series on Curriculum Evaluation. Chicago: Rand McNally, 1972.

_____ and B. Worthen. "Educational Evaluation and Research: Similarities and Differences." In *Curriculum Theory Network*, J. Weiss, ed. Toronto: The Ontario Institute for Studies in Education, 1972.

Guba, E. G. "The Failure of Educational Evaluation." *Educational Technology*, IX (1969), 29-38.

Harris, Chester W. "The Appraisal of a School—Problem for Study." *Journal of Educational Research*, XLI (November, 1947), 172-82.

Hayes, Samuel P., Jr. *Evaluating Development Projects.* Place de Fontenoy, Paris: United National Education, Scientific & Cultural Organization, 1959.

Knox, Alan B. "Continuous Program Evaluation." In *Administration of Continuing Education*, Nathan C. Shaw, ed. Washington, D.C.: National Association for Public School Education, 1969.

Lange, Robert R. "Evaluation As Management Methodology." In *A Strategy for Evaluation Design*, Casper F. Paulson, Jr., ed. Oregon State System of Higher Education, 1970, Section VII.

Lindvall, C. M., and Richard Cox. *The IPI Evaluation Program.* AERA Monograph Series on Curriculum Evaluation. Chicago: Rand McNally, 1970.

Merwin, Jack C., and Frank B. Womer. "Evaluation in Assessing the Progress of Education To Provide Bases of Public Understanding and Public Policy." In *Educational Evaluation: New Roles, New Means*, Ralph Tyler, ed. Chicago: The University of Chicago, 1970, pp. 306-12.

Phi Delta Kappa National Study Committee on Evaluation. *Education and Decision Making.* Itasca, Ill.: F. E. Peacock Publishers, Inc., 1971.

Popham, W. James. *An Evaluation Guidebook.* Los Angeles: The Instructional Objectives Exchange, 1972.

Provus, Malcolm. *Discrepancy Evaluation.* Berkeley, Calif: McCutchan Publishing Corporation, 1971.

Rippey, Robert M. "Can Evaluation Be Integral to Innovation?" In *Curriculum Evaluation: Potentiality and Reality*, Joel Weiss, ed. Curriculum Theory Network Monograph Supplement. Ontario, Canada: Ontario Institute for Studies in Education, 1972, pp. 45-58.

Steele, Sara M. *Contemporary Approaches to Program Evaluation.* Washington, D.C.: Capital Publications, Inc., Educational Resources Division, 1973.

_____. "Program Evaluation—A Broader Definition." *Journal of Extension*, VIII (Summer, 1970), 5-17.

_____. *Six Dimensions of Program Effectiveness*. Madison, Wis.: University of Wisconsin-Extension, Division of Program and Staff Development, 1972.

_____and Robert E. Brack. *Evaluating the Attainment of Objectives: Process, Properties, Problems and Prospects*. Syracuse, N.Y.: Syracuse University Publications in Continuing Education, 1973.

Stufflebeam, Daniel L. "The Relevance of the CIPP Evaluation Model for Educational Accountability." *Journal of Research and Development in Education*, V (February, 1971), 19-25.

Suchman, Edward A. *Evaluative Research*. New York: Russell Sage Foundation, 1967.

Tripodi, Tony; Phillip Fellin; and Irwin Epstein. *Social Program Evaluation*. Itasca, Ill.: F. E. Peacock Publishers, Inc., 1971.

Wilson, Meredith C., and Gladys Gallup. *Extension Teaching Methods*. Extension Service Circular 495. Washington, D.C.: Federal Extension Service—USDA, August, 1955.

Young, Stanley, "Organization As a Total System." In *Emerging Patterns of Administrative Accountability*, Leslie H. Browder, Jr., ed. Berkeley, Calif: McCutchan Publishing Corporation, 1971, pp. 311-30.