

## *A Critique of Evaluation*

Many Extension people who want their programs  
evaluated assume unconsciously that  
findings will be favorable

FRANK D. ALEXANDER

A MOST serious difficulty which evaluative research faces is acceptance by those who are deeply involved in conducting the program (teaching). The Extension worker's (teacher's) self-image is involved, and no amount of moralizing about objectivity can necessarily bring about acceptance. The likelihood that results of evaluation will not be accepted is increased if administrative decisions relating to job security are made as a result of findings. The author, while serving as an AID consultant to the Ministry of Agriculture in Jamaica, observed the backlash of such a decision on the research unit in the Ministry, a backlash which had serious effects on the subsequent role of the unit.

There are several defense mechanisms which the teacher may display toward unpalatable findings when his work is the subject of evaluation. He may claim that the teaching program is better than anything previously developed and that revealing weaknesses through the process of evaluation will only handicap it. He may maintain that the wrong questions were asked, or, if tests were used on the learners, that they were inadequate and failed to test what the learner really got from the educational experience. The data-collecting instrument (i.e., test items or questions) to which program participants responded may be criticized as ambiguous. Finally, if the findings are too distasteful, he may label the interpretation as being false. It is altogether possible that such defense reactions may be wholly or partially justified because of the poor quality of

---

FRANK D. ALEXANDER is Professor of Extension and Rural Sociology and Administrative Specialist in Extension Studies, Cornell University, Ithaca, New York.

the evaluation research. Even so, this is the kind of emotionally charged situation with which evaluative research in Extension is often confronted. It may be well to remember that many, perhaps most Extension people, who want their program evaluated make an unconscious assumption that the findings will be favorable.

However, evaluation as an exercise in Extension teaching need not be done by someone other than the teacher. It can cover a wide range of activities extending from simple, structured introspections by the teacher concerning a talk given at a meeting, to a controlled, experimental piece of research focused on two or more methods for accomplishing a given teaching objective. Because the process of evaluation is viewed unrealistically in many Extension circles, a critique seems timely. The purpose of this article is to attempt such a critique.

Most Extension workers have limited resources and time for conducting evaluation in depth. Such research is usually left to the Extension studies specialist who has the necessary time and resources. It is evaluation at this level that often encounters the kind of limited acceptance referred to in the opening statements. But evaluation does not have to be restricted to efforts at the "research" level. Extension agents, supervisors, or specialists who are immediately concerned with the teaching function may undertake limited evaluation activities which can provide important insights into the effects of teaching efforts. Moreover, from time to time, all types of Extension workers may very well be involved in more complex evaluative research, conducted by Extension studies specialists or other similarly trained researchers. The results of such research can be especially relevant to the teaching functions of Extension workers. Consequently, one purpose of this critique is to analyze the scope of evaluation at various levels in a manner that all types of Extension workers may see the function as a whole and identify the types of evaluation that are especially applicable to their respective roles.

#### LEVELS OF EVALUATION

Frutchey has pointed out that there are degrees of evaluation from *casual everyday evaluation* to *scientific research*.<sup>1</sup> This article will present a more detailed continuum. Twelve levels of evaluation are plotted on a continuum based on a subjective/objective scale, with the subjective end referring to evaluations that are personal introspections and the objective end referring to well designed and

<sup>1</sup> Darcie Byrn (ed.), *Evaluation in Extension* (Topeka, Kansas: H. M. Ives & Sons, Inc., no date), p. 2.

controlled research involving highly sophisticated measuring devices and refined data analysis techniques. The scale is a progressive identification of levels based on the author's judgment. However, since the dividing lines between levels may not be completely clear, the levels are classed under five broader headings.

An important aspect of the progression from one level of the scale to another is the degree of exactness and the objectivity of the methodology employed. Consequently, an important consideration relating to each level is the data collecting device (i.e., questionnaire or test) associated with it. Under the five broad categories, the 12 levels are identified in Table 1, along with a brief explanation of typical kinds of instruments employed and the types of Extension workers most likely to be involved.

Levels 1 and 2 are in the category of "habitual but unorganized" and are likely employed by anyone who teaches. Since they represent unsystematic approaches to evaluation, they will be disregarded in this discussion. Levels 3, 4, and 5, classified as "simple guides," are easily within the resources of those who teach or who are closely associated with teaching (i.e., agents, supervisors, and specialists). Evaluation at these levels requires relatively elementary skills. Agents, supervisors,<sup>2</sup> and specialists who want to conduct more refined evaluation at these levels may find it desirable to seek the assistance of Extension studies specialists or similarly qualified researchers.<sup>3</sup>

Levels 6, 7, 8, and 9 require greater skill and hence may very well be jointly planned with Extension studies specialists or similarly qualified researchers. However, the conduct of studies at these levels can usually be handled by agents or specialists. There are, nevertheless, instances in which evaluation on these levels may require the leadership of an Extension studies specialist or a researcher with similar skills.

Levels 10, 11, and 12 are more complex research exercises and require a high degree of attention and skill. These levels fall within

<sup>2</sup>For participation in evaluation at levels 4 and 5, supervisors appear to have a natural role. Supervisors were not included in levels 6, 7, 8, and 9 since these levels do not seem to require their participation in any important way. They may, however, stimulate agents and specialists to engage in evaluation at these levels and may, at times, participate in planning the exercise.

<sup>3</sup>The author feels strongly that there is a distinct need for an Extension studies specialist on every state Extension staff. In this article the existence of such a position is assumed. While in a few states there is a position of this kind, in other states the study function is performed by a person who is also responsible for training. In still other states Extension studies are conducted by staff members of rural sociology or other departments who are part-time Extension workers or who receive compensation from Extension funds.

**Table 1. Levels of evaluation by broad categories, instruments appropriate, and by Extension workers typically involved.**

Levels	Instruments appropriate	Extension workers involved
<i>Habitual but unorganized</i>		
1. Habitual but unorganized introspections by teacher relative to a teaching situation		A common experience of agents and specialists
2. Panel or staff group who jointly and informally review a teaching situation without a list of questions or items to guide the review		A not infrequent experience of agents, supervisors, and specialists
<i>Simple guides</i>		
3. Teacher's rating by means of a list of questions or items of his teaching relative to what he did and the effects on those being taught	Typical questions or items relate to the degree of interest shown, extent of discussion, teacher's degree of satisfaction, and concepts or skills teacher thought were learned as judged by attention and discussion	Conducted by either agents or specialists with assistance, where needed, from Extension studies specialist in preparing list of questions or items
4. Observations of a teaching situation by an individual or by those taught, guided by a list of questions or items	Typical questions or items relate to attention of students, kinds and extent of each student's participation, judgments by observer of quality of teaching	Conducted by either agents, supervisors, or specialists with assistance, where needed, from Extension studies specialist in preparing list of questions or items
5. Observations of a teaching situation by a panel guided by a list of questions or items	Typical questions or items relate to observed attention of students, kinds and extent of each student's participation, indications of teacher's objectives as revealed in teaching, teacher's management of discussion	Same as for (4) above
<i>Reporting by students or teacher</i>		
6. A system used by the teacher for reporting the learning of individual students	A typical reporting system: cards for each student or a form listing each student on which topics discussed and the teacher's judgment as to learning relevant to the topics are recorded; system would be used primarily in class	Jointly planned with Extension studies specialist, but conducted by agents or specialists
7. Analysis of an educational experience by students using a well designed post-questionnaire	Typical questions relate to the student's awareness of the teacher's objectives; ratings of methods used by the teacher; indications of useful concepts and skills learned; stimulation for further study; relation of learning to job requirements	Jointly planned with Extension studies specialist, but conducted by agents or specialists; Extension studies specialist may in some instances assume leading role in conducting evaluation
<i>Post- or pre- and post-testing</i>		
8. A post-questionnaire or test which ascertains the knowledge, attitudes, and skills of students resulting from one teaching situation	Typical instrument includes true-false, multiple-choice, and application (imagined or real) questions or items which sample expected knowledge, attitudes, and skills	Same as for (7) above
9. A questionnaire or test which ascertains knowledge, attitudes, and skills of students resulting from one teaching situation, administered before and after the teaching situation	Content of instrument similar to (8) above; questionnaire or test same for before and after	Same as for (7) above
10. A questionnaire or test before and after more than one teaching situation (often several as in a program extending over time) which ascertains the knowledge, attitudes, and skills of students resulting from the exposure; desirable to be accompanied by an in-put study	Content of instrument similar to (8) above; same for before and after; instruments for in-put study seek to find what was done in the teaching program; contact records containing reference to topics discussed; teacher's diary, and possibly participant observer's notes	Planned and conducted by Extension studies specialist with assistance from agents, supervisors, or specialists
<i>Experimental design using control or comparative group</i>		
11. A questionnaire or test before and after an educational experience (either one or more exposure) measures knowledge, attitudes, and skills; a similar pre- and post-testing of a matched control group; also a study of the	Content of instrument similar to (8) above and the same for before and after; in-put instruments similar to (10) above	Same as for (10) above

*Content of instrument similar to (8) above and the same for before and after; in-put instruments similar to (10) above*

the scope of responsibility of the Extension studies specialist or similarly qualified researcher, but the successful conduct of evaluation studies at these levels depends in no small measure upon those associated with the teaching (agents, supervisors, specialists).

These 12 levels provide a picture of the total operational scope of Extension evaluation as it relates to teaching. It emphasizes the differential role of Extension workers at various levels; and while responsibilities at these levels should never be arbitrarily determined, it is clear that different types of evaluation are appropriate for different classifications of Extension workers. It is therefore important that each worker understand more clearly his particular role in the total scope of evaluation.

#### PURPOSE AND BY-PRODUCTS

The major purpose of an evaluation exercise in education is to ascertain the effects of teaching under given conditions on the knowledge, attitudes, and behavior of those being taught in order to provide a basis for improving, justifying, or discontinuing the teaching activity. Always, the major focus of evaluation is to attempt to determine what kind of individual emerges from the learning experience to which he is exposed. However, the techniques which have been developed for this central purpose lend themselves to other important uses. Such techniques serve the following purposes:<sup>4</sup>

1. *Clarification of objectives:* To evaluate any educational experiences, it is essential that the aims be defined in terms of the learner—his knowledge, attitudes, and behavior. Otherwise it is impossible to know precisely what measurements to make. If evaluation is to be done effectively, it must be part of the educational activity from the very beginning. If this principle is followed, the evaluation design will inevitably provide both a challenge and an opportunity to clarify educational objectives. The program planner, or teacher, or both, may resist this type of influence by evaluation design because he will say, "Our program is not being planned or conducted to satisfy an evaluation exercise." But if a teaching program or experience is so diffused or poorly planned that it cannot be measured, then the proposed program really needs to be thoroughly examined.
2. *Planning instruction or program on basis of before-testing:* If

<sup>4</sup>For the source of these ideas and their elaboration, along with other related ideas, see Edward J. Furst, *Constructing Evaluation Instruments* (New York: Longmans, Green and Co., 1958), pp. 4-14.

evaluation is designed so that subjects are pre-tested, the teacher can know at what point his students are starting. He is thus in a better position to plan his instruction so that he can adapt it to the needs of the learners (i.e., the extent of their knowledge, inaccurate concepts which they hold, and their attainment of prerequisite skills).

3. *Motivating learning*: If the teacher, through evaluation designing, has clarified goals so that he can make clear to learners what is expected of them, meaning and direction to learning are more fully assured.
4. *Providing guidance to learners*: If the results of evaluation are fed back to the learners, they can correct misconceptions and gain knowledge as to their progress toward goals set by their teachers.
5. *Development of teachers*: Evaluation possesses the potentiality for fostering objectivity and an experimental and creative approach to teaching.

#### COMPLEX LEVELS

Evaluation at the less complex levels should be encouraged. It needs to be conducted more frequently and systematically than is currently being done. At the same time greater attention needs to be given to the more complex levels of evaluation so Extension teaching will increasingly be improved through reliable and valid findings relative to its effectiveness. This does not mean that those who teach will be isolated from the research—the conduct of more complex evaluation is dependent on the full involvement of the teacher. The findings are his concern. Because of resources required for the higher levels of evaluation, problems to be investigated should be thoroughly screened and the study design carefully formulated in order that findings will have significance to Extension.

While such studies may be planned for on-going teaching programs, some attention should also be given to conducting strictly experimental studies using a control or comparative group in which the educational experience (one or more exposure) is planned for the purpose of studying what happens to the students (level 12). At this level, before and after questionnaires or tests are used and a study of the educational in-put is made.<sup>5</sup> An experimental study of

<sup>5</sup> For a study representing this level, see James W. Longest and William W. Gengenbach, *Otsego County Experimental Program for Testing Methods of Forming Farm Management Study Groups: A Progress Report*, Extension Study No. 8 (Ithaca, N.Y.: Office of Extension Studies, State Colleges of Agriculture and Home Economics, Cornell University, February, 1965).

this type is completely within the tradition of the Extension Service and its emphasis on demonstrations. Conducting such evaluation would mean that in order to discover its own effectiveness, Extension apply an approach not unlike what it has long used with its clientele.<sup>6</sup>

The more rigorous evaluative studies involving before and after testing of subjects and control groups pose difficulties for an educational agency whose students are often adults, and always volunteers. Because participation in Extension's educational activities is voluntary, it is difficult to design studies for which a participating group and a nonparticipating group (control) are obtained from the same population by randomization. If the researcher resorts to pair-matching to obtain a control group, the difficulty may be even greater because of inability to match on the factor of volunteering to participate. Moreover, since participation in Extension's educational activities is voluntary, attrition from both the participants and control group can be serious. Such attrition occurs because participation is a voluntary matter—those who agree to participate may drop out of their own accord while, on the other hand, members of the control group may decide to become participants. Despite these handicaps, if Extension workers really place as much value on sound evaluation as their lip-service indicates, imagination and creativity on the part of agents, supervisors, researchers, and administrators can partially overcome these handicaps and successfully execute evaluative research that is soundly designed.

Herzog comments that evaluative research at times "has been plagued by unrealistic expectations on the part of the research consumer and also of research producers. As the magnitude and complexity of the problems become evident, these expectations often give way to a sense of let-down on one side, and a considerable defensiveness on the other."<sup>7</sup> This statement is a concise indication of a fundamental problem of evaluative research growing out of ineffective communication with the consumers of such research in the initiating and planning stages. In considering an evaluation study of a Farm and Home Management Program, an Extension administrator, a state leader of agents, and an agricultural econom-

<sup>6</sup>A comprehensive discussion which goes beyond the possibilities for experimental studies in Extension but which can provide valuable guidance in designing evaluative studies applicable to Extension is to be found in N. L. Gage (ed.), *Handbook of Research on Teaching* (Chicago: Rand McNally & Co., 1964), pp. 171-246.

<sup>7</sup>Elizabeth Herzog, *Some Guide Lines for Evaluative Research* (Washington, D.C.: Children's Bureau, Social Security Administration, U. S. Department of Health, Education and Welfare, 1959), p. 94.

ics specialist who were involved in policy formulation for the program were asked what questions they expected to be answered by the evaluation. Of 20 different questions listed by these three staff members, only three could be dealt with adequately and directly by the findings of the study. Although a number of them could be given judgmental answers by the researchers, such answers were hardly at a scientific level.<sup>8</sup>

#### CONCLUSION

There are approaches to evaluative research which can help Extension personnel properly appraise what such research can accomplish and on what basis they can and should accept findings. Involvement of those planning and conducting teaching activities is basic not only to accepting findings, but also to effective evaluation. Communication between the researchers and program personnel should be continuous. This includes the researcher's reviewing the data-collecting instruments with program people and reporting preliminary findings to them before final reports are completed.

When a serious long-range evaluation program is being planned, it may be necessary to define the research as self-evaluation in the beginning, thus placing considerable responsibility for studies on program personnel. The role of evaluative researchers at this stage is to serve as advisors on design, instrument construction, and interpretation. In his Jamaican experience, the author found this the most effective approach in the attempt to develop an evaluation unit in the Jamaican Ministry of Agriculture following the internal conflict over an evaluation study referred to previously.

In order for the findings of Extension evaluation research to be utilized in making decisions as to the continuation, revision, or discontinuation of educational programs, the research operation must be planned so that the deadline for its completion is established with respect to administrative deadlines for making decisions. While program people may be criticized for impatience expressed by making decisions before the findings of a study are available, researchers, no less, must be held responsible for unrealistic delays in conducting and bringing research to completion.

<sup>8</sup> See James W. Longest and Frank D. Alexander, "Evaluation Study of the Farm and Home Management Program in New York State: Design and Methodology," in *National Extension Research Seminar, April 18-21, 1961*, Darcie Papp (ed.) (Washington: Division of Extension Research and Training, Federal Extension Service, U.S.D.A., ER&T-55, March, 1962), pp. 58-59.