What are Common Evaluation Approaches and Types?

Scholars differ greatly in their views of what evaluation is or is not, how it should be conducted, reported, and utilized. Some opt for a systematic approach to collecting information to assist decision makers. Others view evaluation as synonymous with professional judgment based on expert opinion.

Evaluators hold different philosophies about knowing or establishing truth, popularly known as epistemology. Epistemology is the theory of knowledge or the study of the nature of knowledge, philosophy of knowing, or establishing truth. Based on epistemology, evaluation approaches could be grouped under two categories – objectivism and subjectivism.

Worthen, Sanders and Fitzpatrick (1997) present some key elements of each epistemology. According to them, objectivism requires that evaluation information be "scientifically objective." In other words, use data-collection and analysis techniques that yield results reproducible and verifiable by others. This notion is derived from empiricism. Information or data collected are value-free, which means that the inquirer removes all subjective elements (values) from the situation, and only objective facts are presented. Evaluators are aware of experimental or control groups, subject characteristics, and structured observational protocols.

Subjectivism, on the other hand, bases its validity on "an appeal to experience rather than to scientific method. Knowledge is conceived as being largely tacit rather than explicit (Worthen, Sanders, & Fitzpatrick, 1997, p.65)." Evaluation procedures are "internalized," existing largely within the evaluator in ways that are not explicitly understood or reproducible by others. Information could be value-bound, which means that the inquirer is aware of the roles that values play in a given study. Therefore, it is important for the readers or users of the information to take into consideration both the inquirers' and participants' values and beliefs.

Over the years, evaluation scholars and practitioners have espoused different approaches based on purpose, time of evaluation, methodological backgrounds, and who conducts the evaluation. Some commonly described approaches to evaluation include the following:

Objective-oriented Evaluation Approach: Developed by Ralph W. Tyler in the 1930s to evaluate educational programs and projects, the basic assumption of this approach is that evaluation is a process of determining the extent to which the objectives of a program or project have been attained (Worthen, Sanders, & Fitzpatrick, 1997). Program evaluation involves the establishment of objectives in behavioral terms, finding situations in which achievement of objectives can be shown, collection of performance data, and comparison of performance data with stated objectives.

Management-oriented Evaluation Approach: Developed by Daniel Stufflebeam in the 1960s, this context, input, process and product (CIPP) approach views evaluation as "the process of delineating, obtaining, and providing useful information for judging decision alternatives" Worthen, Sanders and Fitzpatrick (1997, p. 98). The CIPP approach may help program managers to make four different kinds of decisions:

- Context evaluation, to serve planning decisions: What needs to be addressed by a program to help in defining the objectives?
- Input evaluation, to serve structuring decisions: What resources are available and what alternatives should be considered?
- Process evaluation, to serve implementing decisions: How well is the plan being implemented? What barriers threaten its success?

- Product evaluation, to serve recycling decision: What results were obtained? How well were the needs reduced? What should be done with the program after it has run its course? (p. 98)

Consumer-oriented Evaluation Approach: Espoused by Michael Scriven in the 1970s, the focus of this approach is on developing evaluative information such as rating scales for [educational] products for use by consumers in choosing among competing products or services.

Expertise-oriented Evaluation Approach: This is a widely used approach that involves direct application of professional expertise to judge the quality of whatever endeavor is evaluated. Accreditation and formal/informal review panels are examples of this approach.

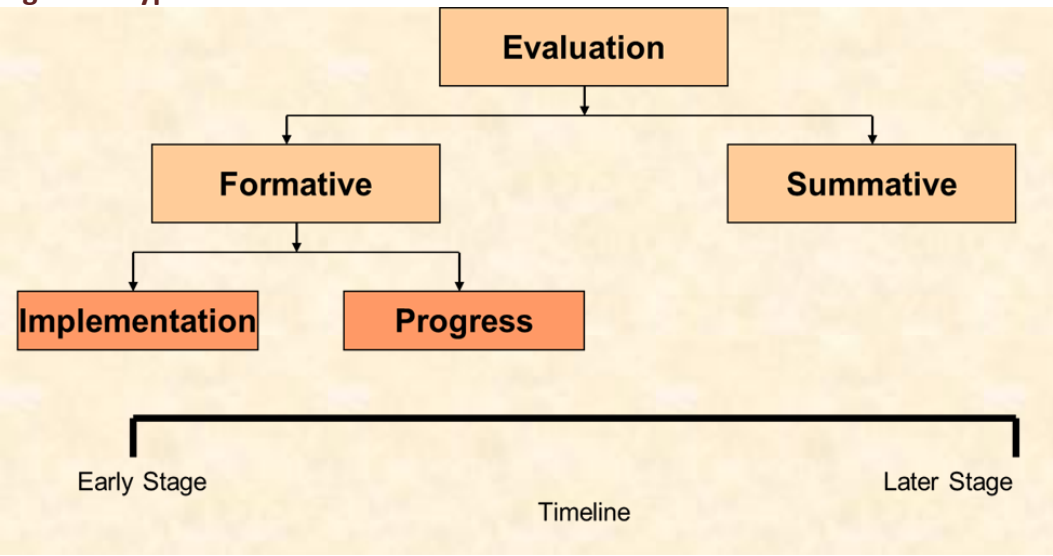
Participant-oriented Evaluation Approach: Developed in the 1970s, this approach considers involvement of participants as central in determining the values, criteria, needs, and data for the evaluation. The practice of stakeholder engagement in planning and evaluation, for example, resulted from this thinking.

The above approaches represent individuals' conceptions about the field of evaluation. They are neither models nor theories, and there is no one best approach to follow. Evaluation contexts and needs are different and it is difficult, if not impossible, to conceive of any one approach to be relevant to all situations. The choice of approach should be based on context, purpose, and resource availability, including expertise.

Based on general purpose, evaluations can be classified in two broad categories – formative and summative.

- A *formative* evaluation is conducted during the life of a program to identify its strengths or weaknesses and to enhance its quality and effectiveness.
- A *summative* evaluation is conducted at the end of a program to help decision makers decide a program's impact and its future. The focus is on determining program results and effectiveness (merit and worth). It serves the purpose of making major decisions about a program – continuation, expansion or reduction, and funding.

Figure 2: Types of Evaluation



Source: Frechtling, Frierson, Hood, & Hughes (2002)

Formative evaluation may take many forms. Usually these are evaluative studies conducted during the early stage of a program or project such as needs assessment, baseline studies, evaluation of on-going activities, or monitoring. Summative studies are usually planned during the later stages of a project (Figure 2).

Based on use of findings, Patton (1997) classifies evaluation into three categories: judgment-oriented, improvement-oriented, and knowledge-oriented. Impact assessments have judgment-orientation. Management generally values improvement-oriented studies. Donors and policy-makers seek knowledge-oriented evaluations and studies that answer accountability questions.

Based on timing and the specific purpose of the evaluation, evaluative studies could be grouped under three broad categories: evaluative studies conducted at the planning stage, during program implementation, and after the program ends.

EVALUATIVE STUDIES CONDUCTED AT THE PLANNING STAGE

Planners and policy makers conduct evaluative studies to gather input from various stakeholders of the extension program under consideration. These evaluative studies contribute to the development of program/project goals, objectives, strategies, and time lines. Results are used to develop feasible and realistic implementation plans. These studies also ensure that all stakeholders, including advisers, share a common vision of the program plan and of the evaluation plan. They help establish the groundwork for formative and later summative evaluations by developing measurable indicators and benchmarks.

The following types of evaluative studies often are conducted at the planning stage:

Context Evaluation: Context evaluation defines the bio-physical, socioeconomic, and cultural environment of the community in which the program/project will be presented. Its goal is to diagnose the needs or issues the program seeks to solve.

Needs Assessment: A needs assessment is an environmental monitoring process through which information is obtained that can be used to design timely, relevant, and reasonable programs (Fear 1988). It helps to establish which messages or media would work best to meet the identified need(s).

Input Evaluation: Input evaluations examine the human resources, budget, technology and equipment, facilities, and other resources that are necessary and available to deliver a program.

Feasibility Study/Market Analysis: The purpose of a feasibility study or market analysis is to determine if a program is feasible and/or desirable, and whether the available inputs and ideas can be crafted into a real-world program to benefit the intended audience. It also assesses how likely the program is to be successful in light of any other service providers who may be offering similar programs.

Benefit/cost analysis and rate of return on investment analysis are common methods used to determine the economic feasibility of new technology or development alternatives.

Baseline Study: Baseline studies measure the status quo, establishing a benchmark against which to judge future changes or program outcomes and impacts.

EVALUATIVE STUDIES CONDUCTED DURING PROGRAM IMPLEMENTATION

In-progress or Formative Evaluation: Formative evaluation looks at how a program is implemented and how the implementation process might be improved. This type of evaluation is conducted to make immediate changes or adjustments in the program. Formative evaluation usually takes place early in implementation and helps program managers find the strengths and weaknesses in a program while it is still going on. This type of evaluation is helpful for program improvement.

Mid-term or Midstream Evaluation: Mid-term evaluations are used when program managers want to adjust a program that is already underway. For a five-year project, mid-term evaluation is usually planned at the end of the second year so that adjustments in program design or delivery can be made to achieve project goals.

Monitoring: Monitoring involves gathering periodic information on project inputs and activities to ensure that the project is implemented as planned and enable management to take corrective actions when necessary. Program managers track resources (e.g., funds, personnel, and supplies) and processes (e.g., occurrence of meetings, demonstrations, and publications). For evaluation purposes, monitoring tracks key indicators of progress over the course of a program as a basis on which to evaluate outcomes of the intervention (Khandker, Koolwal, & Samad, 2010).

Misra (1998) offers 10 principles for monitoring:

1. ***Monitoring must be simple****.* A complex or complicated monitoring system is self-defeating.
2. ***Monitoring must be timely*** so that appropriate action may be taken.
3. ***Monitoring must be relevant*** to program objectives and generate useful information.
4. ***Information provided through monitoring should be dependable.*** Management will rely on monitoring findings only if the information is believed to be reasonably accurate.
5. ***Monitoring efforts should be participatory****.* It should include all concerned with extension, be they field-level personnel, subject-matter specialists, or extension's clients (the farmers).
6. ***Monitoring must be flexible****.* It is iterative in nature. It also becomes routine over time.
7. ***Monitoring should be action-oriented****.* It should follow pragmatic approaches, keeping the requirements of extension's clients uppermost in view.
8. ***Monitoring must be cost-effective****.*
9. ***Monitoring efforts should be top management-oriented****.* Monitoring units should keep in mind the requirements of top management when designing and operating a monitoring system.
10. ***Monitoring units represent specialized undertakings****.* Monitoring is not merely concerned with the collection and analysis of data, but with diagnosing problems and suggesting alternative practical solutions.

Operation Evaluation: Similar to monitoring, operation evaluation seeks to understand whether implementation of a program unfolded as planned. The aim is to compare what was planned with what was actually delivered to determine whether there are gaps between planned and realized outputs (Khandker, Koolwal, & Samad, 2010).

EVALUATIVE STUDIES CONDUCTED AT THE END OF THE PROGRAM OR AFTER

The purpose of program wrap-up, or summative evaluations, conducted at the end of or after a program, is to determine whether or not project objectives were met. These studies look for evidence of the value or success of a program. They measure the effects or impacts of a program. Summative evaluations also are called product, completion, or final evaluations. They supply unbiased information on the impacts, benefits, and cost-effectiveness of a program.

In some extension settings, this is the only evaluation conducted during the life cycle of an educational program. It summarizes what has occurred in the program, asks for end-of-program reactions, and attempts to assess success in meeting program objectives. It is used for program accountability purposes.

There are several different types of summative evaluation. Some of the commonly used include the following:

Output Evaluation: Output evaluation looks at basic program outputs, such as number of training programs conducted, extension bulletins published and distributed, number of male and female farmers reached, and program costs.

Outcome Evaluation: Outcome evaluation often measures progress in learning, such as changes in awareness, knowledge, attitudes, skills or behaviors. Usually, these studies focus on short-term impacts of a program, such as learning and medium-term impacts on people or policy. One example is development of a new policy to establish farmers' markets through cooperatives.

Impact Evaluation: Impact evaluation seeks to measure lasting impacts of programs or projects on important indicators such as crop yields, farm profitability, family income, or livelihood improvement. Such assessments also may focus on broad and long-term program effects, such as changes in ecological, social, economic, or community conditions.

Follow-up Study: A follow-up study is conducted long after a program is completed. This stage of evaluation looks at the long-term benefits of a program or policy. For example, participants in a leadership development program are contacted 5 or 10 years after completion of the training to determine whether and to what extent the training program was a factor in their career accomplishments.

QUALITATIVE VERSUS QUANTITATIVE STUDIES

Based on evaluation philosophy and methodology, evaluative studies can be qualitative or quantitative and follow ex ante or ex post facto methods.

Quantitative methods measure a finite number of predetermined outcomes and are appropriate for judging effects, attributing cause, comparing or ranking, classifying, and generalizing results. Quantitative methods are:

- suitable for large-scale projects;
- useful for judging cause and effect;
- accepted as more credible than qualitative methods by those who are oriented towards numbers; and
- applicable for generalizing to a larger population.

Quantitative methods commonly used in evaluation of extension programs include, but are not limited to:	
Existing information (e.g., census data)	Testing information & knowledge
Surveys (mail, telephone, online)	Economic models
Group-administered questionnaires	Personal survey/interviews

Qualitative methods take many forms, including rich descriptions of people, places, conversations, and behavior. The open-ended nature of qualitative methods allows the person being interviewed to answer questions from his or her own perspective. Qualitative methods are appropriate for:

- understanding the context in which a program takes place;
- understanding complex problems and process issues;
- clarifying relationships between program objectives and implementation;
- identifying unintended consequences of a program;
- gathering descriptive information;
- understanding operations and effects of programs; and
- conducting in-depth analyses of program impacts.

Qualitative methods commonly used in evaluation of extension programs include, but are not limited to:	
Existing information such as newspaper stories, family history, Focus group interviews	Key informant and semi-structured interviews Rapid rural appraisal
Participant observation	Case study

Mixed methods combine qualitative and quantitative methods within one evaluation study. This combination can be used to offset biases and complement each other by contributing the strengths of the varied methods. When using multiple methods, care should be taken to ensure that the selected methods are appropriate to the evaluation questions and that resources are not stretched too thinly. Multiple methods are appropriate for:

- understanding complex social phenomenon;
- allowing greater plurality of viewpoints and interests;
- enhancing understanding of both the typical and unusual case; and
- generating deeper and broader insights.

INTERNAL AND EXTERNAL EVALUATIONS

Based on who conducts the evaluation, either formative or summative, the evaluation could be conducted internally or externally to the organization offering a program.

- *Internal evaluations* are conducted by program/project employees. Generally, internal evaluations are used to monitor progress toward goals and for organizational learning.
- *External evaluations* are conducted by outsiders who are considered experts in the subject-matter. Generally, external evaluations offer more credibility than internal evaluations.

How do managers decide whether to use "in-house" staff to conduct evaluations or to hire external evaluation consultants? Table 1 below shows the advantages and disadvantages of using internal and external evaluators.

Table 1: Using Internal Versus External Evaluators

Options	Advantages	Disadvantages
Internal staff as evaluators	• Familiar with organization • Credible within organization • Develops institutional memory • Can follow up on evaluation recommendations	• Potential for lack of independence • Perceived organizational bias • Burden of additional tasks • Potential lack of power • May lack evaluation skills
External consultants as evaluators	• Specialized skills • New perspectives • Independence and objectivity • Readily available skills • Facilitation of program accountability	• Lack of knowledge of organization • Limited access to information and people • Lack of ability to follow up on recommendations • Can be expensive

Source: Boyle and LeMaire (1999)

When evaluators are evaluating their own programs, there are fewer problems involved in implementing findings. However, when evaluators are not the ones conducting the program, the likelihood of evaluation findings being ignored is greater.

One option is a "hybrid" – using internal staff and, as appropriate, contracting out portions of the evaluation to a professional evaluator. This hybrid approach helps organizations develop their internal evaluation capacity if external evaluators serve as coaches or facilitators.

EVALUATIVE STUDIES TO MEASURE LONG-TERM CHANGES

Agricultural extension and rural advisory services are long-term programs to serve farmers and agribusiness operators. As indicated earlier, many of these programs focus on education and technology transfer. Because learning and behavior change occur over time, evaluators generally are faced with the challenge of measuring long-term changes in participants or communities due to the program in place. The evaluators make use of longitudinal surveys to measure changes over time.

At a minimum, evaluators like to compare current conditions with past conditions using the same or similar measures and tools. Usually, they make the comparison with baseline data, census data, results of content analysis, or they track monitoring data from office records to assess change via select indicators. For example, they compare pre- and post-project data about awareness and knowledge of new farming practices, level of adoption or practice change, cost of production and yield, work-load for women members of household, marketing practices and costs, net profit, and improvement in livelihood.

Longitudinal Surveys: Longitudinal studies help track long-term changes. Three types of survey designs are popular: panel studies, trend analyses, and cohort studies.

> **Panel Studies:** In panel studies, the same subjects are surveyed at different times over an extended period. Although keeping track of people over time is difficult, data gathered from the same sample at different points in time offer meaningful conclusions.
>
> **Trend Analyses:** Different people from the same general population are surveyed at different times to detect if a trend is developing (e.g., adoption of IPM practices in vegetable production, adoption of hybrid corn, or membership in local food marketing cooperatives). Secondary data can show a trend, but cannot explain the cause. Evaluative studies that use a randomized experimental design with control group have the power to explain whether the trend is due to a new agricultural extension program or a new agricultural policy.
>
> **Cohort Studies:** These are longitudinal studies in which a specific population is followed over a period of time. Whereas trend studies sample a general population that changes in membership over time, a cohort study samples a specific population whose members do not change over the course of the study (Ary, Jacobs, & Razavieh, 2002). Cohort studies show if and how things have changed within a segment of the population over time.

Cross-sectional Studies: Cross-sectional studies are studies conducted at a single point in time from a specified population (Weiss, 1998). They measure a certain characteristic(s) or indicator(s), such as knowledge, attitude, or adoption behavior, using surveys, tests, or exams. For example, the evaluator may assess farmers' perceptions of hybrid maize in 2011, 2014 and 2017.

Pre- and Post-program Studies: Pre- and post-program assessment is the common approach to measure changes in specified variables as measured "before" and "after" a program takes place. A pre- and post-program study using a randomized experimental design with control group can explain whether the program has had any effect on the participants. Sample questions for pre-/post-program assessment are: Do farmers who attend extension meetings adopt hybrid corn varieties earlier than those who do not? Do they generate higher net income per hectare?

Does Evaluation Involve Research?

Evaluation professionals use a wide array of research methods, as diverse as casual observation and randomized experimental design. The word "data" refers not only to numbers, but also to content from

interviews, document reviews, observation, and case studies. Although knowledge about in-depth statistics is not often necessary, program managers should be able to identify cause-and-effect relationships between an activity designed to induce change (such as a workshop) and a particular desired outcome (such as increased knowledge of participants).

There are two schools of thought about evaluation of social programs such as agricultural extension. One school believes that evaluation involves value judgments and, thus, absolute accuracy is neither necessary nor attainable. Therefore, evaluation should be structured to serve as a learning process. Evaluators should be careful in the use of evaluation principles to improve judgments and decisions.

The other school considers that evaluation is useful only insofar as it provides credible evidence to inform real-world decision making. This requires (a) sound evaluation design, (b) valid and reliable information gathering about a program's outcomes and impacts, (c) credible methods to analyze and interpret data, and (d) conclusions derived from analysis of empirical data. To guide such evaluative studies, extension managers and policy makers must understand basic research methods and designs.

USAID puts much emphasis on measuring and documenting project achievements and shortcomings through evaluation. The new evaluation policy of USAID (2011) defines impact evaluation as that which:

> "... measures the change in development outcome that is attributable to a defined intervention; impact evaluations are based on models of cause and effect and require a credible and rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change. Impact evaluations in which comparisons are made between beneficiaries that are randomly assigned to either a "treatment" or a "control" group provide the strongest evidence of a relationship between the intervention under study and the outcome measured" (p. 4).

Most evaluators consider evaluation as applied research. Research and evaluation both are grounded in empirical techniques. Both apply a systematic inquiry process. However, the intent of research and evaluation is different. The aim of research is the discovery of knowledge; evaluation in extension focuses on the impact of the application of knowledge on livelihoods and society as a management tool for decision making.