

1. Details of Module and its structure

Module Detail	
Subject Name	Education
Paper Name	Educational Administration, Management & Leadership in School Education
Module Name/ Title	Program Evaluation
Module Id	e-PG EDN 13.22
Pre-requisites	Issues related to the management of public institutions
Objectives	<p>After completion of this module, learners will be able to:</p> <ol style="list-style-type: none"> 1. Explain the concept of Program Evaluation. 2. Explain different types of Program Evaluation. 3. Identify the role of stakeholders in Program Evaluation. 4. Discuss the benchmarks of Credible Evaluation.
Keywords	Evaluation, Evaluator, Program, Stakeholders

2. Development Team

Role	Name	Affiliation
Principal Investigator	Prof. P.K. Sahoo	Department of Education, University of Allahabad, U.P.
Co-Principal Investigator	Prof. Rajaram S. Sharma	Central Institute of Educational Technology, NCERT, New Delhi
Co-Principal Investigator	Prof. Dhananjay Yadav	Department of Education, University of Allahabad, U.P.
Paper Coordinators	Prof. Vidya Agarwal,	Department of Education, University of Allahabad,

		U.P.
	Dr. J.S. Dhillon	Principal, Khalsa College of Education, Amritsar
Content Writer/Author (CW)	Ms. Sukhmandeep Kaur	Khalsa College of Education, Amritsar
Content Reviewer (CR)	Ms. Navdeep Saini	Khalsa College of Education, Ranjit Avenue, Amritsar
Language Editor (LE)	Dr. Aerum Khan, Dr. Astha Saxena	Central Institute of Educational Technology, NCERT

TABLE OF CONTENTS

1. Introduction
2. Learning Outcomes
3. Program Evaluation
4. Purposes of Program Evaluation
5. Reliability, Validity and Sensitivity in program evaluation
 - 5.1 .Reliability
 - 5.2. Validity
 - 5.3. Sensitivity
6. Planning a program evaluation
7. Internal versus external program evaluators
 - 7.1.Internal Evaluators
 - 7.2.External Evaluators
8. Three paradigms
 - 8.1.Positivist

8.2. Interpretive

8.3. Critical-emancipatory

9. Types of Program Evaluations

9.1. Process Evaluations

9.2. Outcome Evaluations

9.3. Impact Evaluations

10. A framework for program evaluation

10.1. Steps in Evaluation Practice

10.1.1. Engage stakeholders

10.1.2. Describe the program

10.1.3. Focus on the evaluation design

10.1.4. Gather credible evidence

10.1.5. Justify conclusions

10.1.6. Ensure use and share lessons learned

10.2. Standards for "good" evaluation

10.2.1. Utility Standard

10.2.2. Feasibility Standard

10.2.3. Propriety Standard

10.2.4. Accuracy Standard

11. Summary

1. INTRODUCTION

As educational programs have increased greatly in size and expense, taxpayers and public officials increasingly urge that these programs be made more accountable to public. Indeed, accountability for expenditures of public funds has become the hue and cry of an ever-increasing number of social reformers. In several countries, policy makers at both national and local levels now routinely authorize funds to be used for the explicit purpose of evaluating educational programs to determine their effectiveness. Evaluation is the systematic application of scientific methods to assess the design, implementation, improvement or outcomes of a program. The term "program" may include any organized action such as media campaigns, service

provision, educational services, public policies, research projects, etc. Thus, program evaluation has come into being as both a formal educational activity and as a frequently mandated instrument of public policy.

2. LEARNING OUTCOMES

After completion of this module, learners will be able to:

1. Explain the concept of Program Evaluation.
2. Explain different types of Program Evaluation.
3. Signify the role of stakeholders in Program Evaluation.
4. Discuss the benchmarks of Credible Evaluation.

3. PROGRAM EVALUATION

Program evaluation is a systematic method of collecting, analysing, and using information to answer questions about programs particularly about their effectiveness and efficiency. In both the public and private sectors, stakeholders often want to know whether the programs they are funding, implementing, voting for, receiving or objecting to are producing the intended effect. Program evaluations can involve both quantitative and qualitative methods of social research. People who do program evaluation come from many different backgrounds, such as sociology, psychology, economics, social work, and public policy. Some graduate schools also have specific training programs for program evaluation.

Program evaluation consists of those activities undertaken to judge the worth or utility of a program in improving some specified aspect of an educational system. Evaluations may be conducted for programs of any size or scope, ranging from an arithmetic program in a particular school to an international consortium on metric education. Examples of program evaluations might include evaluation of a national bilingual education program, a university's pre-service program for training urban administrators, a ministry of education's staff development program, or a local parent education resource centre.

Key Considerations:

Consider the following key questions when designing a program evaluation:

1. For what purposes is the evaluation being done, i.e., what do you want to be able to decide as a result of the evaluation?
2. What are the kinds of information needed to make the decision or enlighten your intended audiences?
3. From what sources should the information be collected?

4. How can that information be collected in a systematic and reasonable fashion, e.g., questionnaires, interviews, examining documentation, observing customers or employees, conducting focus group discussions among customers or employees, etc.
5. When is the information needed?
6. What resources are available to collect the information?

4. PURPOSES OF PROGRAM EVALUATION

Most program evaluators agree that program evaluation can play either a formative purpose (helping to improve the program) or a summative purpose (deciding whether a program should be continued). The main purposes of program evaluation are:

- Demonstrate program effectiveness to funders
- To contribute to decisions about program installation
- To contribute to decisions about program continuation, expansion or certification
- To contribute to decisions about program modifications
- To contribute to the understanding of basic psychological, social and other processes
- Improve the implementation and effectiveness of programs
- Better manage limited resources
- Document program accomplishments
- Justify current program funding
- Support the need for increased levels of funding
- Maintain ethical responsibility towards clients and demonstrate positive and negative effects of program participation
- Document program development and activities to help ensure successful replication

5. RELIABILITY, VALIDITY AND SENSITIVITY IN PROGRAM EVALUATION

It is important to ensure that the instruments used in program evaluation are as reliable, valid and sensitive as possible. According to Rossi et al. (2004, p. 222), 'a measure that is poorly chosen or poorly conceived can completely undermine the worth of an impact assessment by producing misleading results. Only if the outcome measures are valid, reliable and appropriately sensitive can the impact assessments be regarded as credible'.

5.1. Reliability

The reliability of a measurement instrument is the 'extent to which the measure produces the same results when used repeatedly to measure the same thing' (Rossi et al., 2004, p. 218). The more reliable a measure is, the greater its statistical power and the more credible its findings. If a measuring instrument is unreliable, it may dilute and obscure the real effects of a program, and the program will 'appear to be less effective than it actually is' (Rossi et al., 2004, p. 219). Hence, it is important to ensure the evaluation is as reliable as possible.

5.2. Validity

The validity of a measurement instrument is 'the extent to which it measures what it is intended to measure' (Rossi et al., 2004, p. 219). This concept can be difficult to accurately measure: in general use in evaluations, an instrument may be deemed valid if accepted as valid by the stakeholders (stakeholders may include, for example, funders, program administrators, etc.).

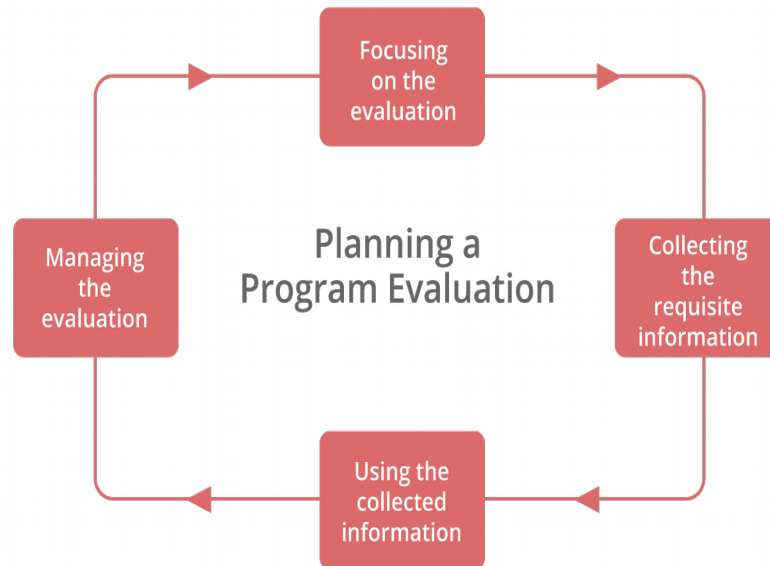
5.3. Sensitivity

The principal purpose of the evaluation process is to measure whether the program has an effect on the social problem it seeks to redress; hence, the measurement instrument must be sensitive enough to discern these potential changes (Rossi et al., 2004). A measurement instrument may be insensitive if it contains items measuring outcomes which the program couldn't possibly effect, or if the instrument was originally developed for applications onto individuals (for example, standardised psychological measures) rather than to a group setting (Rossi et al., 2004). These factors may result in 'noise' which may obscure any effect the program might have had.

Only measures which adequately achieve the benchmarks of reliability, validity and sensitivity can be said to be credible evaluations. It is the duty of evaluators to produce credible evaluations, as their findings may have far reaching effects. A discreditable evaluation which is unable to show that a program is achieving its purpose when it is in fact creating positive change may cause the program to lose its funding undeservedly.

6. PLANNING A PROGRAM EVALUATION

Planning a program evaluation can be broken up into four parts: focusing on the evaluation, collecting the information, using the information, and managing the evaluation.



Program evaluation involves reflecting on questions about evaluation purpose, what questions are necessary to ask, and what will be done with the information gathered. Critical questions for consideration include:

- What am I going to evaluate?
- What is the purpose of this evaluation?
- Who will use this evaluation? How will they use it?
- What questions is this evaluation seeking to answer?
- What information do I need to answer the questions?
- When is the evaluation needed? What resources do I need?
- How will I collect the data I need?
- How will the data be analysed?

7. INTERNAL VERSUS EXTERNAL PROGRAM EVALUATORS

The choice of the evaluator chosen to evaluate the program may be regarded as equally important as the process of the evaluation. Evaluators may be internal (persons associated with the program to be executed) or external (persons not associated with any part of the execution/implementation of the program).

7.1. *Internal evaluators*

Advantages:

- May have better overall knowledge of the program and possess informal knowledge about the program
- Less threatening as already familiar with the staff
- Less costly

Disadvantages:

- May be less objective
- May be more pre-occupied with other activities of the program and not give the evaluation complete attention
- May not be adequately trained as an evaluator.

7.2. *External evaluators*

Advantages:

- More objective towards the process, offer new perspectives, different angles to observe and critique the process
- May be able to dedicate greater amount of time and attention to the evaluation process
- May have greater expertise and knowledge about the evaluation methods

Disadvantages:

- May be more costly and require more time for the contract, monitoring, negotiations etc.
- May be unfamiliar with program staff and may create anxiety among them about being evaluated
- May be unfamiliar with organisation policies, and certain constraints affecting the program

8. THREE PARADIGMS

8.1. *Positivist*

Potter (2006) identifies and describes three broad paradigms within program evaluation. The first, and probably most common, is the [positivist](#) approach, in which evaluation can only occur where there are objective, observable and measurable aspects of a program, requiring predominantly quantitative evidence. The positivist approach includes evaluation dimensions such as needs assessment, assessment of program theory, assessment of program process, impact assessment and efficiency assessment (Rossi, Lipsey and Freeman, 2004).

8.2. *Interpretive*

The second paradigm identified by Potter (2006) is that of interpretive approaches, wherein it is essential that the evaluator develops an understanding of the perspective, experiences and expectations of all stakeholders. This would lead to a better understanding of the various meanings and needs held by stakeholders, which

is crucial before one is able to make judgments about the merit or value of a program. The evaluator’s contact with the program is often over an extended period of time and, although there is no standardised method, observation, interviews and focus groups are commonly used.

8.3. Critical-emancipatory

Potter (2006) also identifies critical-emancipatory approaches to program evaluation, which are largely based on [action research](#) for the purposes of social transformation. This type of approach is much more ideological and often includes a greater degree of social activism on the part of the evaluator. This approach would be appropriate for qualitative and participative evaluations. Because of its critical focus on societal power structures and its emphasis on participation and empowerment, Potter argues this type of evaluation can be particularly useful in developing countries.

9. TYPES OF PROGRAM EVALUATION

All program evaluations share common traits of rigorous planning, careful execution, thoughtful analysis, and thorough reporting.

Table 1: Common research questions asked at different program stages.

Program stage	Common research questions	Evaluation type
Early stage of program or new initiative within a program	<ul style="list-style-type: none"> • Is the program being delivered as intended to the targeted recipients? • Is the program implemented as intended? • Have there been any feasibility or management problems? • What progress has been made in implementing changes or new provisions? 	Process
Mature, stable program with well-defined program model	<ul style="list-style-type: none"> • Are desired program outcomes obtained? • What, if the program produced any 	Outcome

	unintended side effects? <ul style="list-style-type: none"> Do outcomes differ across program approaches, components, providers, or client subgroups? 	
	<ul style="list-style-type: none"> Did the program cause the desired impact? Is one approach more effective than another in obtaining the desired outcomes? 	Impact

9.1. *Process Evaluations*

Process evaluations, also called implementation evaluations, are the most frequently used type of evaluation. They review how a program is implemented and focus on how a program actually operates. Process evaluations can be beneficial throughout the life of a program; however they are often used when a program is implemented to ensure compliance with statutory and regulatory requirements, program design requirements, professional standards, and customer expectations.

9.2. *Outcome Evaluations*

Outcome evaluations, as the name implies, assess program outcomes. Outcomes can be immediate effects of a program or more distal. In general, the closer an outcome is to program outputs, the clearer the linkage between the two. That is, outcomes measured immediately after outputs are generated are less likely to be affected by outside factors that can cloud the relationship between outputs and outcomes. A simple [scenario](#) is provided to illustrate the added complexity of measuring outcomes as they become more distal from the program.

9.3. *Impact Evaluations*

Impact evaluations are designed to measure the net effect of a program by comparing actual program results with counterfactual data. Excluding all potential causes of an outcome can be a difficult and expensive proposition and is sometimes impossible. Because of their cost and required expertise, and often the need to plan the

evaluation during initial program design rather than after program implementation, impact evaluations are not common. Although impact evaluations should be planned during program start-up, they should not be undertaken until program operations are mature so that the true effect of the fully implemented program can be assessed.

10. A FRAMEWORK FOR PROGRAMME EVALUATION

The framework described below is a practical non-prescriptive tool that summarises in a logical order the important elements of program evaluation.

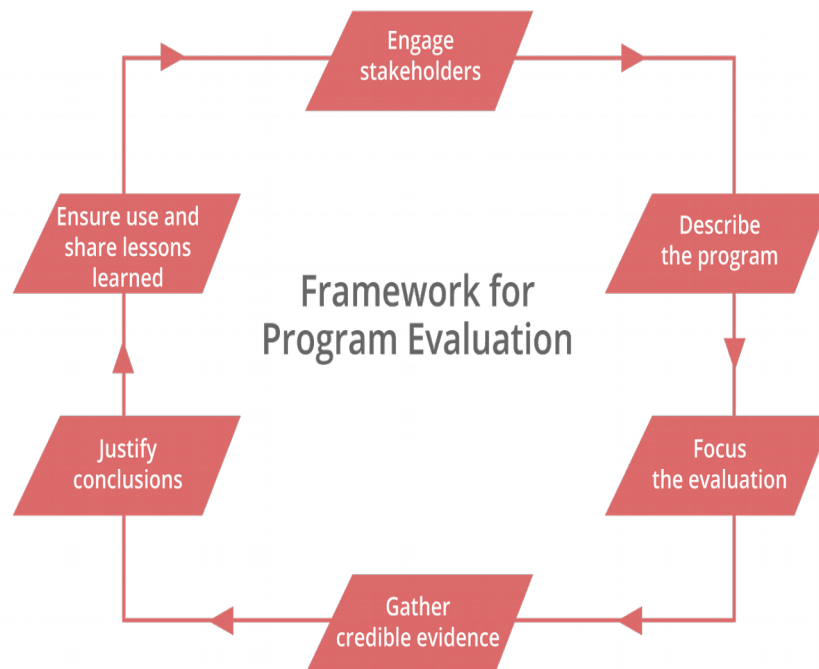
The framework contains two related dimensions:

- Steps in evaluation practice, and
- Standards for "good" evaluation

10.1. Steps in evaluation practice

The six connected steps of the framework are actions that should be a part of any evaluation. They are intended to serve as starting points around which community organisations can tailor an evaluation to best meet their needs.

- Engage stakeholders
- Describe the program
- Focus the evaluation design
- Gather credible evidence
- Justify conclusions
- Ensure use and share lessons learned



10.1.1. Engage Stakeholders

Stakeholders are people or organisations that have something to gain or lose from what will be learned from an evaluation, and also in what will be done with that knowledge. Evaluation cannot be done in isolation. Stakeholders must be part of the evaluation process in order to ensure that their unique perspectives are understood and included. When stakeholders are not appropriately involved, evaluation findings are likely to be ignored, criticised, or resisted. However, if they are a part of the process, people are likely to feel a good deal of ownership for the evaluation process and its results. They will probably want to develop it, defend it, and make sure that the evaluation really works.

10.1.2. Describe the Program

A program description is a summary of the intervention being evaluated. It should explain what the program is trying to accomplish and how it tries to bring about those changes. The description will also illustrate the program's core components and elements, its ability to make changes, its stage of development, and how the program fits into the larger organisational and community environment.

10.1.3. Focus the Evaluation Design

By focusing on the evaluation design, we mean, doing advance planning about where the evaluation is headed, and what steps it will take to get there. It isn't possible or useful for an evaluation to try to answer all questions for all stakeholders; there must be a focus. A well-focused plan is a safeguard against wastage of time and resources. Depending upon your objective or the area of evaluation, some types of evaluation will be better suited than others. However, once data collection begins, it may be difficult or impossible to change what you are doing, even if it becomes obvious that other methods would work better. A thorough plan anticipates intended uses and creates an evaluation strategy with the greatest chance to be useful, feasible, proper, and accurate.

10.1.4. Gather Credible Evidence

Credible evidence is the raw material of a good evaluation. The information learned should be seen by the stakeholders as believable, trustworthy, and relevant to answer their questions. This requires thinking broadly about what counts as "evidence." Such decisions are always situational; they depend on the question being posed and the motives behind asking it. For some questions, a stakeholder's standard for credibility could demand having the results of a randomised experiment. For another question, a set of well-done, systematic observations such as interactions between an outreach worker and community residents will have high credibility. The difference depends on what kind of information the stakeholders want and the situation in which it is gathered.

Having credible evidence strengthens the evaluation results as well as the recommendations that follow from them. Although all types of data have limitations, it is possible to improve an evaluation's overall credibility. One way to do this is by using multiple procedures for gathering, analysing, and interpreting data. Encouraging participation by stakeholders can also enhance perceived credibility. When stakeholders help define questions and gather data, they will be more likely to accept the evaluation's conclusions and to act on its recommendations.

10.1.5. Justify Conclusions

The process of justifying conclusions recognises that evidence in an evaluation does not necessarily speak for itself. Evidence must be carefully considered and examined from a number of different stakeholders' perspectives to reach conclusions that are well-substantiated and justified. Conclusions become justified when they are linked to the evidence gathered and evaluated against agreed-upon values set by the

stakeholders. Stakeholders must agree that conclusions are justified in order to use the evaluation results with confidence.

10.1.6. *Ensure Use and Share Lessons Learned*

It is naive to assume that lessons learned in an evaluation will necessarily be used in decision making and subsequent action. Deliberate effort on the part of evaluators is needed to ensure that the evaluation findings will be used appropriately. Preparing for their use involves strategic thinking and continued vigilance in looking for opportunities to communicate and influence. Both of these should begin in the earliest stages of the process and continue throughout the evaluation process.

10.2. *Standards for "good" evaluation.*

The second part of the framework is a basic set of standards to assess the quality of evaluation activities. There are 30 specific standards, organised into the following four groups:

- Utility
- Feasibility
- Propriety
- Accuracy

10.2.1. *The utility standards are:*

- Stakeholder Identification: People who are involved in (or will be affected by) the evaluation should be identified, so that their needs can be addressed.
- Evaluator Credibility: The people conducting the evaluation should be both trustworthy and competent, so that the evaluation will be generally accepted as credible or believable.
- Information Scope and Selection: Information collected should address pertinent questions about the program, and it should be responsive to the needs and interests of clients and other specified stakeholders.
- Values Identification: The perspectives, procedures, and rationale used to interpret the findings should be carefully described, so that the bases for judgments about merit and value are clear.
- Report Clarity: Evaluation reports should clearly describe the program being evaluated, including its context, and the purposes, procedures, and findings of the evaluation. This will help ensure that essential information is provided and easily understood.
- Report Timeliness and Dissemination: Significant midcourse findings and evaluation reports should be shared with intended users so that they can be used in a timely fashion.

- **Evaluation Impact:** Evaluations should be planned, conducted, and reported in ways that encourage follow-up by stakeholders, so that the evaluation will be used.

10.2.2. Feasibility Standards

The feasibility standards are to ensure that the evaluation makes sense - that the steps that are planned are both viable and pragmatic.

The feasibility standards are:

- **Practical Procedures:** The evaluation procedures should be practical; to keep disruption of everyday activities to a minimum while needed information is obtained.
- **Political Viability:** The evaluation should be planned and conducted with anticipation of the different positions or interests of various groups. This should help in obtaining their cooperation so that possible attempts by these groups to curtail evaluation operations or to misuse the results can be avoided or counteracted.
- **Cost Effectiveness:** The evaluation should be efficient and produce enough valuable information so that the resources used can be justified.

10.2.3. Proprietary Standards

The propriety standards ensure that the evaluation is an ethical one, conducted with regard for the rights and interests of those involved. The eight propriety standards follow.

- **Service Orientation:** Evaluations should be designed to help organisations effectively serve the needs of all of the targeted participants.
- **Formal Agreements:** The responsibilities in an evaluation (what is to be done, how, by whom, when) should be agreed to in writing, so that those involved are obligated to follow all conditions of the agreement, or to formally renegotiate it.
- **Rights of Human Subjects:** Evaluation should be designed and conducted to respect and protect the rights and welfare of human subjects, that is, all participants in the study.
- **Human Interactions:** Evaluators should respect basic human dignity and worth when working with other people in an evaluation, so that participants don't feel threatened or harmed.
- **Complete and Fair Assessment:** The evaluation should be complete and fair in its examination, recording both strengths and weaknesses of the program being evaluated. This allows strengths to be built upon and problem areas addressed.

- **Disclosure of Findings:** The people working on the evaluation should ensure that all of the evaluation findings, along with the limitations of the evaluation, are accessible to everyone affected by the evaluation, and any others with expressed legal rights to receive the results.
- **Conflict of Interest:** Conflict of interest should be dealt with openly and honestly, so that it does not compromise the evaluation processes and results.
- **Fiscal Responsibility:** The evaluator's use of resources should reflect sound accountability procedures and should otherwise be prudent and ethically responsible, so that expenditures are accounted for and appropriated.

10.2.4. Accuracy Standards

The accuracy standards ensure that the evaluation findings are considered correct. There are 12 accuracy standards:

- **Program Documentation:** The program should be described and documented clearly and accurately, so that what is being evaluated is clearly identified.
- **Context Analysis:** The context in which the program exists should be thoroughly examined so that likely influences on the program can be identified.
- **Described Purposes and Procedures:** The purposes and procedures of the evaluation should be monitored and described in thorough detail so that they can be identified and assessed.
- **Defensible Information Sources:** The sources of information used in a program evaluation should be described in enough detail so that the adequacy of the information can be assessed.
- **Valid Information:** The information gathering procedures should be chosen or developed and then implemented in such a way that they will assure that the interpretation arrived at is valid.
- **Reliable Information:** The information gathering procedures should be chosen or developed and then implemented so that they will assure that the information obtained is sufficiently reliable.
- **Systematic Information:** The information from an evaluation should be systematically reviewed and any errors found should be corrected.
- **Analysis of Quantitative Information:** Quantitative information - data from observations or surveys - in an evaluation should be appropriately and systematically analysed so that evaluation questions are effectively answered.
- **Analysis of Qualitative Information:** Qualitative information - descriptive information from interviews and other sources in an evaluation should be appropriately and systematically analysed so that evaluation questions are effectively answered.
- **Justified Conclusions:** The conclusions reached in an evaluation should be explicitly justified, so that stakeholders can understand their worth.

- **Impartial Reporting:** Reporting procedures should guard against the distortion caused by personal feelings and biases of people involved in the evaluation, so that evaluation reports fairly reflect the evaluation findings.
- **Meta-evaluation:** The evaluation itself should be evaluated against these and other pertinent standards, so that it is appropriately guided and, on completion, stakeholders can closely examine its strengths and weaknesses.

11. SUMMARY

There is a growing need for accountability of government funds budgeted for development programs. Taxpayers and government officials are interested in knowing exactly how money is being spent and what impact is being made. One strategy to improve accountability for government funds is enforcing program evaluation. Evaluations detail program inputs, outputs, and the outcomes and impacts that track the use of such funds. However, the consistency of rigorous evaluations at the level of outcomes and impacts is limited, as conducting evaluations often relies upon availability of data, funds, and the interest of donors and program management. Evaluation is a powerful strategy for distinguishing programs and interventions that make a difference from those that do not. It is a driving force for developing and adapting sound strategies, improving existing programs, and demonstrating the results of investments in time and other resources. It also helps to determine if what is being done is worth the cost it incurs.

This recommended framework for program evaluation is both a synthesis of existing best practices and a set of standards for further improvement. It supports a practical approach to evaluation based on steps and standards that can be applied in almost any setting. Because the framework is purposefully general, it provides a stable guide to design and conduct a wide range of evaluation efforts in a variety of specific program areas. The framework can be used as a template to create useful evaluation plans to contribute to understanding and improvement.

INTERESTING FACTS

1	Evaluation is the systematic application of scientific methods to assess the design, implementation, improvement or outcomes of a program
2	Program evaluations can involve both quantitative and qualitative methods of social

	research.
3	The framework of program evaluation contains two related dimensions: Steps in evaluation practice, and Standards for "good" evaluation
4	The benchmarks of credible evaluations are reliability, validity and sensitivity
5	The feasibility standards are to ensure that the evaluation makes sense - that the steps that are planned are both viable and pragmatic
6	There are 30 specific standards of program evaluation organised into four groups- Utility, Feasibility, Propriety and Accuracy