# Effectiveness of Evaluations in South Africa Organisations

Gcobani Mkwela[1]; Mandla Alfred Twala[2]; Marty Zikhali[3]

[1] Lion Pride Estates, Malibongwe Drive, 2088, South Africa

[2] Lebombo Avenue, Oakdene, 2190, South Africa

[3] University of Limpopo, South Africa

E-mail: Gcobani.Mkwela@gauteng.gov.za, Mandla.Twala@gauteng.gov.za, marty.zikhali@ul.ac.za

***Abstract***

This article seeks to investigate the effectiveness of evaluations in organizations. Evaluations contribute to the improvement of public policy interventions and expenditure programs by providing evidence-based assessments of their relevance and performance. Evaluations also serve to strengthen accountability by providing reliable information on progress in the achievement of public objectives to stakeholders, often identifying the key factors driving success or failure. Moreover, evaluation is a potent tool that assist organizations striving for a highly competitive world. It also provides quality control to support one's efforts and ensures that time, money, and resources are well spent. The more evaluation approaches become ingrained in our culture, the more chances there are to use evaluation to improve the effectiveness and efficiency of services, systems, and programs. The research study adopted "the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA)", as an item checklist, that filters relevant information using the four criteria, namely, identification, screening, eligibility, and inclusion. PRISMA collects and analyzes data from the studies that were included in the review and employs statistical and methodical methods to identify and evaluate relevant research. The findings revealed that evaluation is an important source for communication, and marketing efforts and produces results for decision-making. Evaluation becomes a tool and resource to provide organizations with direction and guides strategic decisions. Furthermore, it helps to build a culture of learning within organizations and promote continuous improvement. It is recommended that this will happen through the provision of evidence-based feedback and recommendations, hence evaluation encourages program managers and implementers to reflect on their practices and processes, and also identify opportunities for growth and development.

***Keywords:*** *Evaluation; Program; Effectiveness; Policy; Organization*

## Introduction

Organizations are increasingly being asked to explain their achievements and how they are making a difference in people's lives, as well as their return on investment. But the citizens are more interested in economic outcomes such as the Growth Domestic Product (GDP) and employment creation. On that note, evaluation is a process that helps to establish and implement policies and programmes that lead to a better future. The coexisting difficulties in the organization can be solved by investigating various techniques and implementing evaluation (Brousselle & McDavid, 2021; Brousselle et al., 2022; Patton, 2020; Rowe, 2018, 2019a; Uitto & Batra, 2020; Van Den Berg et al., 2021). This article seeks to provide a snapshot of the effectiveness of evaluations in organizations.

Several scholars evaluate the worth of a particular programme, project, or policy. This determines a specific intervention's effectiveness, as well as examines and improves its quality. Evaluation is used as a technique to assess the level of success in achieving targeted outcomes and objectives. Essentially, practitioners carefully write explicit measurable objective statements that allow evaluators to determine whether the intervention effects the targeted indicators and/or whether corrective measures were employed to assess effectiveness. Evaluation allows one to discover the critical parts of a specific intervention (such as activities, content, resources, and structure), adjust content and implement methods, and decide whether or not to invest more resources to improve business performance. Evaluation occurs in every aspect of life in some form. If the evaluation process is removed from human life, perhaps the purpose of life will be defeated. The evaluation value chain is the only way to determine how effectively a program, policy, or initiative is being executed.

Effective evaluation depends on how well the design and measurements chosen to fit the research questions and the population under study. However, measuring is essential to assessment because it gives the evaluator testable proof of participant growth and program performance, as well as the ability to determine whether any changes or improvements arise as a result of the intervention. Similarly, another finding of Gulis et al. (2022) is that, even with the current framework or guidelines, it is frequently difficult to interpret the evaluation results if the evaluator has the necessary content or industry expertise. According to Chikwe's (2017) theory, evaluation is the process of giving a phenomenon a symbol in order to describe its value or worth, typically in relation to scientific or societal norms. On the other hand, reviews are typically planned to be carried out for the donor's advantage in order to track spending, impact, and intervention. More precisely, carrying out a suitable and thorough assessment demonstrates the evaluator's accountability to the audiences and communities they serve, the organization they work for, the funding source for the project, and the public service field as a whole.

Evaluation provides a multitude of diverse functions in holding stakeholders accountable. Fundamentally, evaluation activities also assist in ascertaining whether or not predefined goals pertaining to performance metrics were met. Enhancing aspects related to program execution (such efficacy and efficiency) is another benefit of evaluation. Evaluation also advances our understanding of the factors that influence infrastructure, security, education, politics, the environment, finances, health, and other concerns, as well as the most effective and suitable solutions to address them. This information is very helpful in directing future studies and applications. Policy decisions at the organizational, municipal, regional, provincial, national, and international levels are also informed by evaluation.

While vital, monitoring just inquires as to whether we are carrying out our planned activities. We must do assessments in order to determine whether or not our strategies are having the desired effects and/or results, as well as the reasons why. Deep investigation of topics like effectiveness, efficiency, relevance, accountability, and value for money is required for evaluation. Organizations undoubtedly need to use the evaluation outcomes to guide changes to our strategy. According to Beeby (2007), evaluation is defined as "the systematic collection and interpretation of evidence leading as a part of process to a judgment of value with a view of action" and consists of four essential components.

**Historical Overview of Evaluation**

Even though many academic sectors see evaluation as a relatively recent profession, evaluation activities have been around for a while. According to Stufflebeam, Madaus, and Kellaghan (2000), the term "evaluation" first appeared in the United States of America in the 1800s when the government assigned inspectors to assess how well public facilities—such as jails, schools, hospitals, and orphanages—were operating. The assessment profession as we know it now began in the 1960s when scholars reported the findings of evaluations they did at the time. Beginning with the introduction of Lyndon Johnson's Great Society projects, which included the Elementary and Secondary Education Act and Head Start programs, which required program reviews. The heterogeneous academic roots of evaluation have further complicated its history. For example, psychologists are more closely associated with applied social research traditions, whereas educational evaluators come from a background in testing, assessment, and objectives-based evaluation (Mark, Greene & Shaw, 2006). Regarding the educational perspective Guba and Lincoln (1989) proposed four generations of assessment:

- First generation: Student testing and measurement
- Second generation: Tyler's study, mentioned in Stufflebeam et al. (2000), describes the objectives and tests.
- Third generation: Judgment: the models based on decisions, as Stufflebeam (1982), Scriven (1967a), and Stake (1983).
- Fourth generation: experimental, constructivist assessment.

Madaus et al. (2000), on the other hand, identified seven stages of evaluation development. "There are seven periods recognized by scholars: the first is the Age of Reform, which began before 1900; the second is the Age of Efficiency; the third is the Tylerian Age, which lasted from 1930 to 1945; the fourth is the Age of Innocence, which lasted from 1946 to approximately 1957; the fifth is the Age of Development, which lasted from 1958 to 1972; the sixth is the Professionalization Age, which lasted from 1973 to 1983; and the seventh is the Expansion and Integration Age, which lasted from 1983 to 2000.

A thorough and impartial assessment of a project, program, or policy, including its design, implementation, and outcomes, is what the Organization for Economic Co-operation and Development (OECD) (2002) defines as "evaluation." The objective is to ascertain the significance and accomplishment of goals, as well as the efficacy, impact, sustainability, and efficiency of development. Recipients and donors should be able to incorporate lessons learned into their decision-making process by using reliable and practical information from an evaluation (Patton, 1997). A few definitions of evaluation, program evaluation, project evaluation, and policy evaluation are given in the table 1 below.

Table 1: Definitions of Evaluation

| Source | Description / Definitions |
|---|---|
| Alkin (1970) | The concept evaluation means: *"Evaluation is the process of examining the performance of an organization, program, project, policy, or any other intervention to determine its relevance, adequacy, effectiveness, efficiency, and progress for the purpose of identifying areas for improvement. Essentially, it refers to the combination of evidence and values to determine whether an intervention has merit, worth, or significance.* |
| Mertens & Wilson, (2013); and Scriven, (2003) | In fact, there is no single definition for Evaluation: *"It is an important tool for understanding the effectiveness of a program or initiative. Evaluation helps identify what is working and what is not so that solutions be found to improve the project or initiative. It also reveals elements of a program or initiative that could be better utilized and determines if alternative methods should be adopted to support program goals.* |

| (Rossi & Freeman, (1993); and Short, Hennessy & Campbell, (1996) | **Programme evaluation** is the systematic application of scientific methods to assess the design, implementation, improvement, or outcomes of a program. |
|---|---|
| Steven et al. (1993) | **Project evaluation** is the process of measuring the success of a project, program or portfolio. This is done by gathering data about the project and using an evaluation method that allows evaluators to find performance improvement opportunities. |
| Gasper, D. (2018) | **Policy evaluation** is defined as a *"structured and objective assessment of an ongoing or completed policy or reform initiative, its design, implementation and results. It aims to determine the relevance and fulfilment of objectives, efficiency, effectiveness, impact and sustainability, as well as the worth or significance of a policy"*. |

### *Distinction between Evaluation and Research*

Concerns over the line that separates evaluation from study are widespread. Since evaluation and research are sometimes mistaken for one another, the researchers felt it was crucial to distinguish between the two. While research focuses on creating new knowledge or expanding on already acquired knowledge, evaluation entails assessing an existing program. Evaluation and research share commonalities in that they use comparable methodologies and approaches. For instance, they both use primary and secondary data, surveys, observations, and the review of books, papers, and journal articles, among other sources of information. Evaluation can be done quantitatively or qualitatively, much like research.

The graphic in Figure 1 below depicts the points in the process where research and assessment divide and where they converge. While there are many variations between the start and finish of research or assessment initiatives, techniques and analysis are where the disciplines come together. When it comes to methodology, researchers and evaluators are undoubtedly speaking the same language.
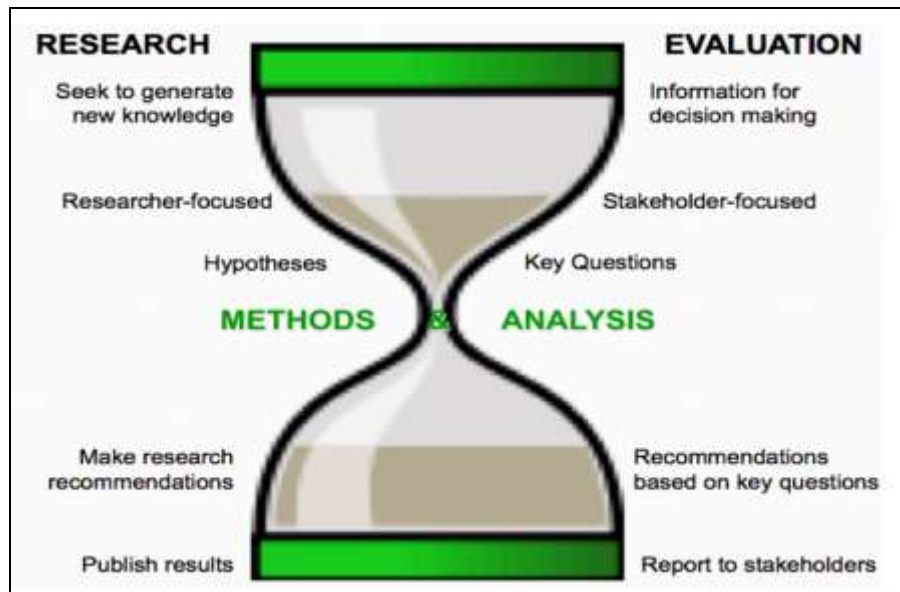


Figure: 1 Distinction between evaluation and research

"The systematic assessment of the operation and the outcomes of a program or policy, compared to a set of explicit or implicit standards, as a means of contributing to the improvement of the program or policy" is how Weiss, cited in Powell (2006), defines evaluation. Stated differently, assessment determines value by posing the questions of what functions, what doesn't, and why. Conversely, research is the methodical process of gathering, evaluating, and interpreting facts to deepen our comprehension of a topic (Leedy & Ormrod, 2010). The goal of research is to produce new knowledge. According to Lewis and Thornhill (2009), this indicates that people conduct research in order to learn more in an organized manner.

According to Powell (2006), evaluation research is a kind of study that applies conventional social research methods for evaluative goals. It may also be characterized as a particular research methodology or an assessment process that makes use of certain procedures that are only used in the evaluation of social programs. Powell further argues that evaluative research is more likely to be viewed as an action or applied research type than as fundamental or theoretical research when it is used as a research method. Though they contribute to various kinds of understanding, evaluation and research both aim to deepen our understanding in a number of areas.

According to Rozalis (2003), evaluation and research are two different and distinct disciplines even though they share some concepts, instruments, and procedures. While an evaluation's goal is to give program managers insightful input on a particular program, research aims to expand the corpus of scientific knowledge. While research is theory-dependent, evaluation is field-dependent, and theory is utilized to enhance and explain assessment findings. While research settings and data collection techniques are derived from theory, evaluation settings and methodologies are derived from the field (Rozalis, 2003). Rozalis goes on to say that researchers are typically active, whereas evaluators are more likely to be reactive.
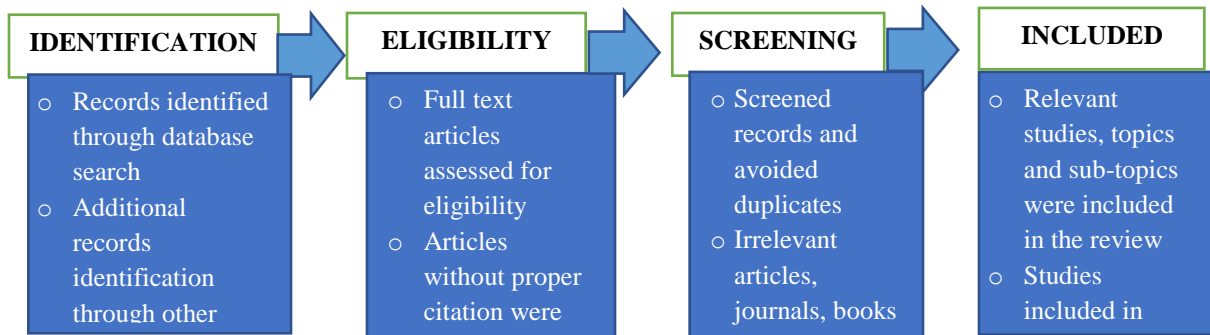
Research is done to further scientific understanding, while evaluation is mostly done for project measurements. Surveys might be strictly quantitative or completely open-ended in their use of evaluation. The primary frame of reference for the acknowledged logic of inquiry in science is theory (Rozalis, 2003). Nevertheless, Scriven (1991) asserts that theory is useless for evaluation.

Evaluation end by measuring value, and research concludes by measuring causes. Researchers and evaluators have similar values in that they both need to present a degree of impartial, objective value judgments. According to Kusters et al. (2017), evaluations are the end result of the process of evaluating something's merit, worth, or value. Evaluation and research are not similar except in the approaches they use. For instance, evaluation employs interactive and observational methods and contains aspects of action research. This indicates that the procedures for gathering, compiling, and analyzing data are comparable.

### *Methodology*

The screening procedure is summarized in the PRISMA flow diagram. To assess and report on the literature review, the researchers used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method. PRISMA flow diagram and checklist were used in the study. PRISMA collects and analyzes data from the studies that were included in the review, as well as employs statistical and methodical methods to identify and evaluate relevant research. The number of citations and journal impact factor were among the specified criteria used to filter the papers found for eligibility and relevancy. The reliability of the sources was evaluated by the authors using the PRISMA checklist, and only the most pertinent studies were considered in the study. The search was also narrowed down by using a number of restrictions, such as defining a time span to weed out publications that did not fall

inside that range. The PRISMA flow diagram, shown in the picture below, provides the writers with four phased criteria to help them with their research: identification, screening, eligibility, and inclusion.

| IDENTIFICATION | ELIGIBILITY | SCREENING | INCLUDED |
|---|---|---|---|
| o Records identified through database search<br>o Additional records identification through other | o Full text articles assessed for eligibility<br>o Articles without proper citation were | o Screened records and avoided duplicates<br>o Irrelevant articles, journals, books | o Relevant studies, topics and sub-topics were included in the review<br>o Studies included in |

Simplified flow chart showing the selection process for the articles

**Benefits of Evaluation**

Evaluation is crucial because the organization has a moral duty to give the public the finest programs and/or services possible. For instance, evaluation guarantees that your time, money, and resources are well-used and offers quality control to assist your efforts. Senior management and policymakers should turn to evaluation for reliable, unbiased analysis and suggestions to guide their decision-making (Solheim, 2013).

Through evaluation, an organization can determine what outcomes it is attaining and how, as well as whether it is operating effectively and efficiently toward its goals. For instance, organizations should be able to clearly identify which initiatives are most likely to produce the intended outcomes in a given situation before allocating funding accordingly. When there is a lack of compelling evidence, plans for gaining a deeper knowledge ought to be included from the start. Accordingly, evaluation results should be incorporated into strategic planning, including sector methods and country strategies (Solheim, 2013).

Three primary variables are needed for evaluation, according to Eresia-Eke & Ile (2019): cost, time, and quality. First and foremost, the cost-effective review must guarantee that no public resources are wasted. This means that management must make sure that costs are evaluated and that they are connected to concrete actions, outputs, and results. Second, time is a crucial component of evaluation since it allows for the planning and distribution of the project's required duration. The project may fail completely as a result of poor time management. Finally, quality guarantees that the service meets or beyond expectations and that the resources invested in the project were worthwhile. This implies that the development of quality indicators is necessary to guarantee the efficacy and efficiency of the in the implementation of projects.

Lessons from evaluation studies have shaped our current approach to development. For example, an evaluation study conducted in Ethiopia found that nutrition programmes were targeting the wrong children with supplements – reaching kids that were either too old to benefit or were not actually malnourished. These evaluation findings resulted in significant changes to the way the programmes were being implemented and ultimately improved health outcomes in the beneficiary communities. While monitoring, performance management and results reporting systems provide some information, in-depth evaluation is needed to explain how results are achieved (Solheim, 2013).

Ijeoma (2010) asserts that performance, quality, relevance, and impact are all covered by evaluation in general. Even if evaluation is helpful in determining relevance, efficacy, efficiency, impacts, and sustainability, stakeholder participation is still required in the methodical process of evaluating public

programs. In order to inform funders, stakeholders, and the general public on the outcomes of a program or project, evaluation is essential. This is done by:

o Offering evidence-based perspectives on the success and effects of a program, evaluations promote, increase awareness of, and push for change. Since it guarantees that evaluation results are comprehensible and available to a broad spectrum of stakeholders, effective communication is an essential component of evaluation. This comprises those involved in the project or program, such as financiers, legislators, program managers, and members of the community.

o Evaluation helps to guarantee that the results are understood and used to guide decision-making and resource allocation by conveying evaluation findings succinctly and clearly. A project or program's support can also be increased through effective communication, which highlights its accomplishments.

o Positive assessment results are shared with stakeholders and the larger community, which aids in increasing program awareness and fostering support for ongoing execution. By pointing out areas that require improvement, evaluation can also be used to promote change. Evaluation, which finds the program's or project's shortcomings and inadequacies, helps to organize support for changes by bringing attention to the need for change. Evaluation contributes to coalition building, awareness-raising, and change advocacy by offering evidence-based insights and succinctly conveying findings.

Over the course of a program, numerous evaluations are conducted for a variety of applications and objectives (Cloete, 2009). These include formative evaluation, which is carried out at the start of a program to determine its necessity, and impact evaluation, which is carried out after a program has been in place for some time to determine the degree to which the program has contributed to the changes it is meant to bring about (Potter, 2006). Program monitoring is carried out during the program's execution to offer information and comments on how effectively the program is proceeding as planned. Nonetheless, comparable procedures are used at each of these M&E levels (Cloete, 2009).

Ile, Eresia-Eke, & Allen-Ile (2012) state that there are often some connected tasks between monitoring and evaluation. Critically describing the relationship between the two processes as intertwined and indicating that monitoring should come before evaluation, monitoring gives information for assessment. They claimed in their research that, in general, it is nearly hard to conduct adequate evaluations in the absence of monitoring. As a result, monitoring can occur without assessment, whereas assessment almost always depends on monitoring. In addition, Nalubega & Uwizeyimana (2019) urge assessment scholars and practitioners to sharpen their intellectual faculties so they can take advantage of the opportunities posed by technological advancements to create workable, long-lasting solutions to M&E problems in the African context.

### *Discussion and Results*

### **Policy Framework**

The goal of the policy framework is to address the usage of evaluation by highlighting the effects of government initiatives while also enhancing accountability and transparency. The framework seeks to integrate assessment with planning and budgeting procedures and highlight the role evaluation plays in management and policymaking. It seeks to raise the standard of assessments conducted and guarantee that evaluation results are applied to enhance performance. Government employees, managers, and political figures are the document's primary intended readership. Under Section 195 of the Constitution, it is required that public administration principles include:

- Accountability, development orientation,
- Efficient, economical, and
- Effective resource usage is all necessary for public administration.
- Transparency is also encouraged by giving the public timely, accurate, and accessible information.

In addition, a legal foundation for the effective and efficient administration of public policies and programs is provided by the Public Finance Management Act (PFMA, 1999), the Municipal Finance Management Act (MFMA), and the Public Service Act (1994 as amended by Act 30 of 2007). Additionally, these Acts offer a legitimate foundation for carrying out the various evaluation kinds.

**Types of Evaluations**

Notably, this article elaborates on its focus on the sorts of evaluations, notwithstanding the importance of evaluation in the organization. Over time, the function of evaluation has changed. Evaluations can be classified into three primary categories based on their design and purpose: Formative Evaluation, Summative Evaluation, and Diagnostic Evaluation (Flay et al., 2005). Even though they differ, these categories are typically not independent. The distinctions have to do with why it is being done. The objective of the endeavor, the service delivery issue being addressed, the audience, the setting, and the timetable all influence the assessment design choice.

As a result, assessment techniques must be tailored to the subject matter and evaluation goals. Knowing the many evaluation techniques that can be utilized during a program's life cycle and when to employ them is crucial (McKenzie et al., 2009). The following are a few of them:

- Formative evaluation is usually conducted early in an initiative to "pilot test" in order to gather input from relevant parties, improve the intervention's content and components, and determine the initiative's future course. The most common issues in formative evaluation are practicality and suitability of tools and techniques. Formative assessment enables the initial testing and improvement of research hypotheses, data gathering tools, and statistical/analytical techniques. Before "going to scale" (i.e., allocating larger investments of time, effort, and resources), this form of evaluation is typically conducted on a small scale to ensure unexpected problems (e.g., glitches, breakdowns, lengthy delays, and departures from the design) are identified and the intervention quality is improved.
- Summative evaluation includes the intervention's overall effectiveness in terms of short-, intermediate-, and long-term results. Since anticipated results and objectives can only be attained if the intervention is implemented faithfully, as planned, this kind of evaluation also takes process evaluation into account in addition to intervention efficacy.
- Diagnostic evaluation is defined by The Guideline on Diagnostic Evaluation 2014 as preliminary research (also known as ex-ante evaluation) to determine the actual state of affairs before an intervention and to guide the design of the intervention. It examines the state of affairs as it is, the issues and chances that need to be resolved, the causes and effects—including those that the intervention is unlikely to produce—as well as the likelihood of success for various course of action. Umoinyang, Asim, Akwa & Bassey (2004) state that it offers an extensive search for potential causes of learning issues in addition to supplemental formative evaluation.

Table 2: Formative and Summative Evaluation

| **Formative evaluation** includes several evaluations | **Summative evaluation** also be sub-divided |
|---|---|
| ▪ needs assessment determines who needs the program, how great the need is, and what might work to meet the need.<br>▪ evaluability assessment determines whether an evaluation is feasible and how stakeholders help in shaping its usefulness.<br>▪ structured conceptualization helps stakeholders define the program or technology, the target population, and the possible outcomes.<br>▪ implementation evaluation monitors the fidelity of the program or technology delivery.<br>▪ process evaluation investigates the process of delivering the program or technology, including alternative delivery procedures. | ▪ outcome evaluations investigate whether the program or technology caused demonstrable effects on specifically defined target outcomes.<br>▪ impact evaluation is broader and assesses the overall or net effects - intended or unintended - of the program or technology as a whole.<br>▪ cost-effectiveness and cost-benefit analysis address questions of efficiency by standardizing outcomes in terms of their dollar costs and values.<br>▪ secondary analysis reexamines existing data to address new questions or use methods not previously employed.<br>▪ meta-analysis integrates the outcome estimates from multiple studies to arrive at an overall or summary judgment on an evaluation question. |

Most evaluations, according to Babbie & Mouton (2001), take place at particular points during the implementation process, like its start, middle, or finish. According to Mertens & Ginsberg (2009), evaluation is the methodical application of social research techniques to assess the merits and demerits of social interventions, encompassing initiatives, policies, people, organizations, and products. The evaluation kinds associated with formative assessment and the division of summative evaluation are depicted in the table below.

This article emphasizes that assessing complex system-based programs, such a governance system, using theory-of-change may only be viewed as a positive first step. It will assist us in moving toward the creation of critically important evaluation paradigms as well as cutting-edge approaches and strategies that may be able to assist in addressing the intricate issues that the governance system is currently facing.

Evaluation was largely expected to provide order to the disorganized realm of government, as noted by Weiss (1977). According to Weiss, a major factor contributing to the difficulty of evaluating complex programs is the inadequate or nonexistent identification of the underlying assumptions. This raises questions about how the initiative will play out. As a result, the formative review of early and mid-term indicators—which is necessary to achieve a longer-term goal—gets little attention. Weiss proposed a different kind of evaluation known as "theory-based evaluation," in which the evaluation's "theories of change" served as the foundation for the outcomes-based evaluation. Program theory (Weiss, 1998) and the program's theory of action (Patton, 2008) are other names for theory-of-change. The theory-of-change is also known as the program's theory of action (Patton, 2008) or program theory (Weiss, 1998). Theoretical presumptions about why a program will succeed (achieve its goals) or fail form the foundation of change theories.

**Moving Beyond the Theory-of-Change**

This article emphasizes that assessing complex system-based programs, such a governance system, using theory-of-change may only be viewed as a positive first step. It will assist us in moving toward the creation of critically important evaluation paradigms as well as cutting-edge approaches and strategies that may be able to assist in addressing the intricate issues that the governance system is currently facing.

Evaluation was largely expected to provide order to the disorganized realm of government, as noted by Weiss (1977). According to Weiss, a major factor contributing to the difficulty of evaluating complex programs is the inadequate or nonexistent identification of the underlying assumptions. This raises questions about how the initiative will play out. As a result, the formative review of early and mid-term indicators—which is necessary to achieve a longer-term goal—gets little attention. Weiss proposed a different kind of evaluation known as "theory-based evaluation," in which the evaluation's "theories of change" served as the foundation for the outcomes-based evaluation. Program theory (Weiss, 1998) and the program's theory of action (Patton, 2008) are other names for theory-of-change. Theories of change are predicated on theoretical presumptions about whether or not the program will be effective (achieve its goals).

While causal-loop diagrams are useful for showing feedback effects, they are not sufficient to alter a system as a whole. According to Patton (2008):

> "...evaluators may need to investigate, comprehend, and assess the ways in which being a part of larger bureaucracies affects program and project effectiveness when evaluating the effectiveness of government programs" (Patton, 2008).

On the other hand, evaluation occurs at designated times and allows for a longer-term review of a program's development. Evaluation monitors changes and concentrates more on the result and degree of impact. The following figure, which depicts the relationship between the planning cycle and the chain of inputs, outputs, results, and impacts, demonstrates this. Figure 2 below provides an example of how to apply Theory of Change (ToC).
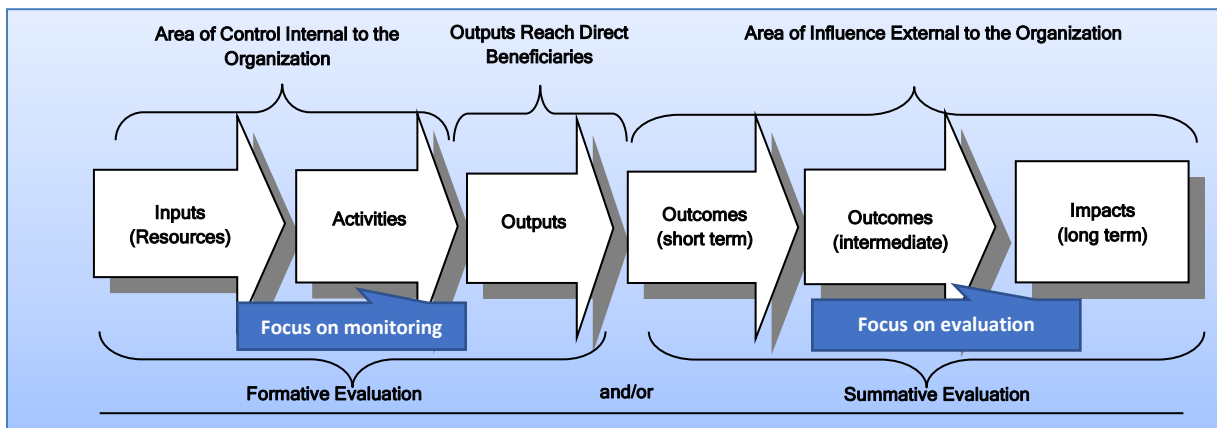


Figure 2: Using Theory of Change

The evaluation cycle provided a comprehensive structure for comprehending evaluation as an iterative process. The three primary steps of the evaluation approach guide, created by Daykin et al. (2013), are project planning, data collection and analysis, and reporting and dissemination.

Remarkably, the cycle's data collecting, and analysis phase takes up very little time, yet project planning, reporting, and distribution frequently take longer than anticipated. This accurately captures the procedures involved in evaluation. Moreover, the cycle has no ultimate conclusion. Phases are interconnected, with lessons learned from each evaluation—including ethical considerations—feeding into project design and subsequent review. The Evaluation Cycle or Framework is depicted in Figure 3 below.
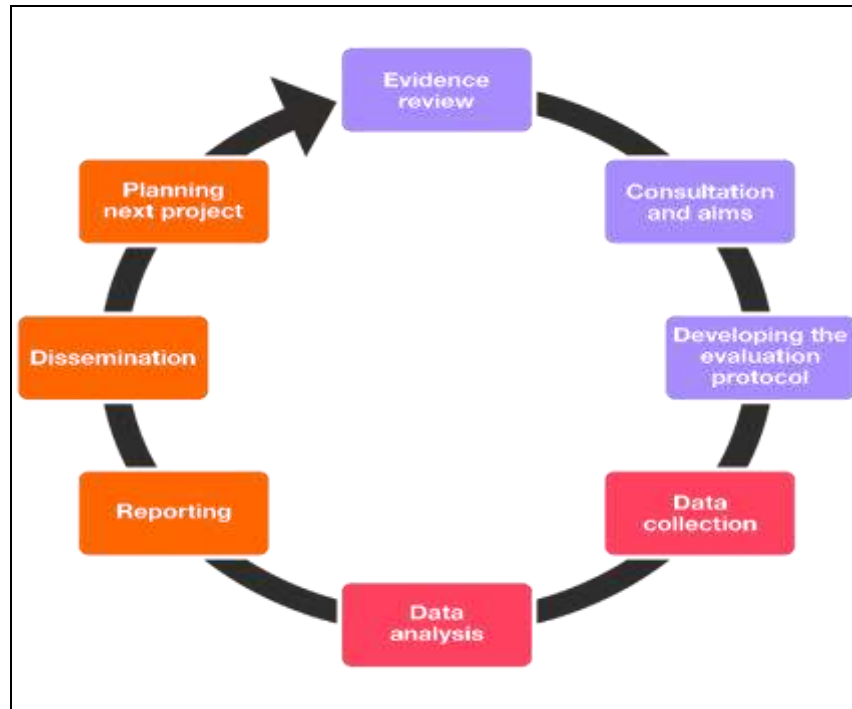


Figure 3: Evaluation Cycle or framework

**Terms of Reference for an Evaluation**

Creating a precise Terms of Reference (ToR) document that outlines every detail for the teams or evaluators doing the assessment is essential. It clearly describes the resources available to carry out the study, establishes the goals and parameters of the assessment, and delineates the roles and duties of the consultant or team. Creating a precise and detailed ToR is an essential part of overseeing an excellent assessment. The assessment of the ToR document establishes the standards by which the assignment's success will be judged and forms the foundation of a contractual agreement with one or more evaluators.

The Independent Evaluation Group (2011) states that the terms of reference have multiple functions and aid in defining the appropriate structure and content. The terms of reference were also explained in depth in this page. Step-by-step instructions for conducting an evaluation are provided in the table below.

Table 3: Terms of Reference for Evaluation

| Key Aspects | Descriptions |
|---|---|
| 1. Sharing background knowledge and providing the rationale for the evaluation | ▪ The project or program being evaluated is described along with key milestones in its history. This allows the ToR to show how the evaluation or review contextualised within the broader development strategy for a sector or country/region. A ToR typically includes a brief review of relevant available knowledge regarding the program and its effects to inform evaluation consultants or teams. |
| 2. Identifying the specific evaluation questions | ▪ ToR builds an understanding of the scope, process, and expectations for the desired task(s) by succinctly presenting information about why the evaluation is being conducted, its objectives, and its intended users. |
| 3. Defining the scope, approach, and methodology | ▪ Terms of reference typically specify the scope of the evaluation (time period, depth, etc.), but vary in terms of how much flexibility consultants or in-house evaluators have to propose their own methodologies or approaches to complete the designated tasks. However, at least the expected broad approach is outlined clearly in the ToR to set realistic expectations among all relevant parties engaged in the study. The degree to which the evaluator(s) propose additional or alternative methods for completing the task(s) should also be specified. |
| 4. Articulating the governance and accountability arrangements | ▪ This section highlights the governance and accountability arrangements. ToR document is crafted for an individual consultant or for a team. Various stakeholders will be engaged to facilitate or participate in the work, and parallel related tasks conducted by other consultants might be in progress. ToR outlines the roles and responsibilities envisioned to carry out the assignment and the management and coordination arrangements. The hierarchy for accountability and the structure and resources established for support are also explicitly stated in this section. |
| 5. Setting the guiding principles or values | ▪ Terms of reference specify research ethics or procedures that the evaluator(s) is expected to follow. |
| 6. Identifying the professional qualifications of the individual evaluator or team | ▪ ToRs present the expected profile of the evaluation team. This includes describing desired experience and credentials, as well as noting the minimum professional requirements or competencies. |
| 7. Defining the deliverables and schedule | ▪ ToR specifies the expected deliverables, timeline, and any work plan if available. ToR may ask the evaluator to provide a detailed timeline and milestones within the timeline specified. ToRs list any products that the evaluators should develop as part of their assignment. To the degree possible, this list includes details related to format, content, length, intended audience, and the expected review process. |
| 8. Defining the budget | ▪ ToR states the budget (and potentially other resources) available for the evaluation and what that budget covers. |

## Conclusion and Recommendations

In summary, the researchers aim to persuade decision-makers, program designers, politicians, and evaluators to consider the entire spectrum of value chains that resulted in effect assessments. An essential component of encouraging learning and growth in projects and programs is monitoring and evaluation. Along with evaluating a program or project's strengths and limitations, the evaluation offers insightful commentary on what is successful and what needs improvement. In order to help program administrators and implementers find opportunities for innovation and improvement, the input is essentially being utilized to guide decisions about program design, implementation, and resource allocation.

In certain cases, the evaluation also offers evidence-based criticism and suggestions, encourages continual improvement, and helps organizations create a culture of learning. Program managers and implementers are encouraged by evaluation to consider their procedures and practices and to look for areas where they may improve. Additionally, this results in better program design, more successful execution, and enhanced impacts and outcomes. By disseminating insights and lessons gained to other organizations and stakeholders, evaluation also aids in the promotion of organizational learning. This is mostly accomplished by sharing assessment results and recommendations; evaluation also adds to the body of knowledge in the field and fosters group growth and learning. Since evaluation offers insight into the effectiveness and efficiency of a program, it is crucial for promoting learning and development.

It is recommended that the results be shared with interested parties, organizations, and other relevant stakeholders. When it comes to the planning, execution, and future of initiatives and programs, evaluation will be essential to enable evidence-based decision-making. Evaluation will also support program managers and decision-makers in making well-informed decisions about how to allocate resources, enhance program design, and produce better results and impacts by offering trustworthy and valid evidence about the efficacy and efficiency of a program (addressing social, economic, and environmental challenges). Furthermore, it will also aid in identifying knowledge gaps and areas in which more studies  required. If the proposed recommendations are implemented, this will promote the more effective and efficient use of resources and strengthens the body of knowledge in the sector.

## References

Alkin, M. C. (1970): Evaluation Theory Development. Evaluation of Short-Term Training in Rehabilitation, 9-16.

Babbie, E. and Mouton, J. (2015): The Practice of Social Research. Republic of South Africa: Oxford University Press.

Babbie, Earl, and Johann, Mouton. (2001): The Practice of Social Research. Cape Town: Oxford University Press.

Beeby, S.T. (2007): Strategy for Evaluation Education Programmes (mineo). A paper presented at the Nigeria Educational Research Council (NERC) Inter – Departmental Seminar, Lagos.

Brousselle, A. and McDavid, J. (2020): Evaluators in the Anthropocene. Evaluation: The International Journal of Theory, Research and Practice 26(2): 190–204.

Brousselle, A. and McDavid, J. (2021): Evaluation for Planetary Health. Evaluation: The International Journal of Theory, Research and Practice 27(2): 168–183.

Brousselle, A., McDavid, J. and Curren, M., et al. (2022): A Theory-Based Approach to Designing Interventions for Planetary Health. Evaluation: The International Journal of Theory, Research and Practice 28(3): 330–355.

Chikwe, K.C. (2017): Fundamentals of Testing. Port Harcourt: Emmanuel ventures.

Cloete, Fanie. (2009): *"Evidence-Based Policy Analysis in South Africa: Critical Assessment of the Emerging Government-Wide Monitoring and Evaluation."* Journal of Public Administration 44 (2): 293-311.

Daykin, N., Attwood, M. and Willis, J. (2013): Supporting Arts and Health Evaluation: Report of a UK Knowledge Transfer Partnership. Journal of Applied Arts & Health. 4 (2): 179-190. DOI: http://dx.doi.org/10.1386/jaah.4.2.179_1.

Department of Planning, Monitoring and Evaluation. National Evaluation Policy Framework, November 2019. Pretoria.

DPME Guideline 2.2.10 Diagnostic Evaluation. 20 March 2014. The Presidency Republic of South Africa, Pretoria.

Flay, B.R., Anthony, B., Robert, F.B., Flelipe, G.C., Denise, G., Sheppard, K., Eve, K.M., Steven. S., Jeffrey, C.V. and Peter, Ji. (2005): Standards of Evidence: Criteria for Efficacy Effectiveness and Dissemination. Prevention Science Vol.6 No 3 September 2005. Dol: 10.1007/S11121-005-555-Y.

Gasper, D. (2018): "Policy Evaluation: From Managerialism and Econocracy to a Governance Perspective", in *International Development Governance*, Routledge, https://doi.org/10.4324/9781315092577-37.

Guba, E. G., and Lincoln, Y. S. (1981): Effective Evaluation: Improving the Usefulness of Evaluation Results through Responsive and Naturalistic approaches. San Francisco: Jossey-Bass.

Gulis, G., Krishnankutty, N., Boess, E.R, et al. (2022): Environmental Impact Assessment, Human Health and the Sustainable Development Goals. International Journal of Public Health 67: 1604420. Harris P, Viliani F and Spickett J, (2015) Assessing Health Impacts within Environmental Impact Assessments: An Opportunity for Public Health Globally which must not remain missed. International Journal of Environmental Research and Public Health 12:1044–9.

Joint Information Systems Committee (JISC) (2007): Six Steps to Effective Evaluation: A Handbook for Programme and Project Managers.

Kusters, C., Batjes, K., Wigboldus, S., Brouwer, J., and Baguma, S.D., (2017): Managing for Sustainable Development Impact: An Integrated Approach to Planning, Monitoring and Evaluation. Wageningen University and Research: Practical Action Publishing Ltd, Netherlands.

Leedy, P.D and Ormrod, J.E. (2010): Practical Research: Planning and Design. 9th edition. Boston, MA: Pearson Education International.

Lewis, S.M. and Thornhill, A. (2009): Research Methods for Business Students. Pearson Education International.

Ile,I., Eresia-Eke C & Allen-Ile, C. (2012): Monitoring and Evaluation of Policies, Programmes and Projects, Van Schaik Publishers, Pretoria.

Ile I.U, Eresia-Eke C. & Allen-Ile C. (2019): Monitoring and Evaluation of Policies, Programmes and Projects: Practical Guide to Public Policy Monitoring and Evaluation. 2nd edition. Van Schaik Publishers, Pretoria: South Africa.

Ijeoma, E.O.C. (2010): Mainstreaming Government-Wide Monitoring and Evaluation Policy in South Africa. Journal of Public Administration. 45(2):344–360.

*Independent Evaluation Group (IEG) annual report (2011): Results and Performance of the World Bank Group (Vol. 3) : Management action record (English).* Results and Performance (RAP) Washington, DC: World Bank. http://documents.worldbank.org/curated/en/758641468319776005/Management-action-record.

Madaus, G.F., Stufflebeam, D.L., & Kellaghan, T. (2000): Evaluation models: Viewpoints on educational and human services evaluation. (2nd ed.). Hingham, MA: Kluwer Academic Publishers.

Mark, M. M., Greene, J. C., & Shaw, I. E. (2006): The Evaluation of Policies, Programs, and Practices. In I. E. Shaw, J. C. Greene, & M. M. Mark (Eds.), The SAGE Handbook of Evaluation (pp. 3–32). Thousand Oaks, CA: SAGE.

McKenzie, J.F., Neiger, B.L., and Thackeray, R. (2009): Planning, Implementing, & Evaluating Health Promotion Programs: A Primer. 5th ed. San Francisco: Benjamin Cummings. 464 p.

Mertens, D.M. and Ginsberg, E.P. (2009): Performing Qualitative Cross-Cultural Research. Cambridge: Cambridge University Press.

Mertens, Donna. M. & Wilson, Amy. T. (2019): Program Evaluation Theory and Practice: A Comprehensive Guide. 2nd edition. New York London. The Guilford Press.

Mouton, J. (2017): Designing Evaluation Studies. CREST, Stellenbosch University.

Nalubega, T. & Uwizeyimana, D.E. (2019): 'Public Sector Monitoring and Evaluation in the Fourth Industrial Revolution: Implications for Africa'. Africa's Public Service Delivery and Performance Review Vol 7(1).

OECD, (Organisation for Economic Co-operation and Development). (2002): "Glossary of Key Terms in Evaluation and Results-Based Management." Paris: Development Assistance Committee.

Patton, Michael Q. (1997): Utilization-Focused Evaluation. London: Sage Publications.

Patton, M.Q. (2020): Blue Marble Evaluation. Premises and Principles. New York: The Guilford Press.

Patton, M.Q. (2008): Utilization-Focused Evaluation. 4th ed. Thousand Oaks, California: Sage.

Potter, Charles. (2006): *"Programme Evaluation."* In Research in Practice: Applied Methods for the Social Sciences, 409-429. Cape Town: UCT Press.

Powell, R.R. (2006): Evaluation Research: An Overview Library Trends, 55(1), pp.102-120.

Public Finance Management Act (PFMA, 1999): The Public Service Act (1994 as amended by Act 30 of 2007) and the Municipal Finance Management Act (MFMA).

Public Services Commission (2008): Basic Concepts in Monitoring and Evaluation. PSC Commission House: Pretoria. Retrieved from: http://www.psc.gov.za.

Rossi, P. H., & Freeman, H. E. (1993): Evaluation: A Systematic Approach. 5ᵗʰ ed. Newbury Park, CA: Sage Publications, Inc.

Rossi, Peter. H., Howard, E. Freeman, and Mark, W. Lipsey. (1999): Strategies for Impact Assessment. New Delhi, Sage Publications. Inc.

Royse, D., Thyer, B., and Padgett, D. (2009): Program Evaluation: An Introduction. Belmont, CA: Cengage Learning. 416 p.

Rothman, J. and Thomas, E.J. (eds) (1994): *Intervention Research: Design and Development for Human Service*. New York: Haworth.

Rowe, A. (2018): Ecological thinking as a route to sustainability in evaluation. In: Hopson R and Cram F (eds) Tackling Wicked Problems in Complex Ecologies: The Role of Evaluation. Stanford, CA: Stanford University Press, pp. 25–44.

Rowe, A. (2019a): Sustainability-Ready Evaluation: A Call to Action. New Directions for Evaluation 162: 29–48.

Rozalis, M.L. (2003): Evaluation and Research: Differences and Similarities. Canadian Journal of program Evaluation, Vol 18(2). Ben-Gurion University of the Negev Beer-Sheva: Israel.

Scriven, M. (1967b): The Methodology of Evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), Perspectives of Curriculum Evaluation (AERA Monograph Series— Curriculum Evaluation). Chicago: Rand McNally.

Scriven, M. (1991): *Evaluation thesaurus* (4th ed.). Sage Publications, Inc. [Access 27 March 2024].

Scriven, M. (2003): Evaluation in the New Millennium: The Transdisciplinary Vision. In S. I. Donaldson & M. Scriven (Eds.), Evaluating Social Programs and Problems (pp. 19–41). Mahwah, NJ: Erlbaum.

Short, L., Hennessy, M., & Campbell J. (1996): Tracking the Work. *In Family violence: Building a Coordinated Community Response: A Guide for Communities.*

Solheim, E. (2013): Evaluating Development Activities: Providing Evidence on Results for Learning and Decision Making, OECD.

Stake, R. E. (1983): The Case Study Method in Social Inquiry. In G. F. Madaus, M. Scriven, & D. L. Stufflebeam (Eds.), Evaluation Models (pp. 279–286). Boston: Kluwer- Nijhoff.

Stevens, F., Lawrenz, F., & Sharp, L. (1993): User Friendly Handbook for Project Management: Science, Mathematics, Engineering and Technology Education. Washington DC: National Science Foundation.

Stufflebeam, D. L. (1982): A Review of Progress in Educational Evaluation. American Journal of Evaluation, 3, 15–27.

Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (2000): Evaluation Models. Boston: Kluwer Academic.

Tukwayo, P.E. (2018): Evaluation of the Effectiveness of the Implementation of the Staff Performance Management System in the Department of Social Development: A Case Study of Eerste River Office: Stellenbosch University: Cape Town, South Africa.

Uitto, J. and Batra, G. (eds). (2020): Transformational Change for People and the Planet. Evaluating Environment and Development. Cham: Springer.

Umoinyang, I. E., Asim, A. E., Akwa, A. M. & Bassey, S.W. (2004): Principles and Techniques of Educational Assessment and Evaluation. Calabar: Helimo Assosiates.

Van Den Berg, R.D., Magro, C. and Adrien, M-H. (2021): Transformational Evaluation for the Global Crises of Our Times. Exeter: International Development Evaluation Association (IDEAS).

Weiss, Carol. H. (1998): Evaluation: Methods for Studying Programmes and Policies. New Jersey: Prentice Hall.

Weiss, C.H. (1977): Using Social Research on Public Policy Making. Lexington: MA: DC Health.

Windsor, R. A., Baranowski, T., Clark, N, and Cutter, G. (1984): Evaluation of Health Promotion and Education Programs. J Sch Health: 54(8):318. doi:10.1111/j.1746-1561.1984.tb08946.x.

Writing Terms of Reference for an Evaluation: A How To Guide. (2011): Independent Evaluation Group Communication. Strategy, and Learning International Bank for Reconstruction and Development / World Bank. Washington, D.C. 20433, U.S.A.

Wollmann, H. (2017): Policy Evaluation and Evaluation Research. In Handbook of Public Policy Analysis (pp. 419-428). Routledge.