



THE PRESIDENCY
REPUBLIC OF SOUTH AFRICA

DEPARTMENT: PERFORMANCE MONITORING AND EVALUATION

National Evaluation Policy Framework

Draft 3

31 August 2011

Foreword

In our Green Paper on Improving Government Performance we said:

'The transition to democracy fostered hope for a society where all citizens would have the opportunity to realise their full intellectual, physical, social and spiritual potential. This vision was captured in the Constitution, which spells out each citizen's entitlement to adequate housing, basic education, health care, food and water and social security. Although the rights are to be realised progressively over time within the available resources, the gap between vision and reality remains large. ... to improve service delivery standards ...we must do more with less. The focus has to be on value for money. Wasteful and unproductive expenditure and corruption cannot be afforded.... This part of the process is about improving our efficiency, it is about reducing the unit cost of the service we provide. Ensuring that the outputs deliver the outcomes that have been politically chosen, is a measure on whether government is being effective. Genuine change based on critical self-reflection is required. That means changes in how we behave, not just superficial adjustments to existing processes, systems and formats.'

If we are to improve our performance we have to reflect on what we are doing, what we are achieving against what we set out to achieve, and why deviations are occurring, or unexpected results occurring. We cannot advance without making mistakes on the way, but we must evaluate and learn from our successes and our mistakes. Without this we cannot improve.

We have moved to establish plans for our priority outcomes, to deliver them and to monitor them. This policy framework provides the next essential part of the jigsaw, setting out the basis for a government-wide evaluation system to be applied across the public sector, but initially focusing on our priority areas. It should help to provide a marked step-up in performance of the public sector and contribute to the establishment of a culture of continuous improvement.

Collins Chabane

Minister of Performance Monitoring, Evaluation and Administration

Policy summary

The National Evaluation Policy Framework (NEPF) provides the basis for a minimum system of evaluation across government. Its main purpose is to improve the effectiveness and impact of government, by reflecting on what is working and what is not working and revising its programmes and policies accordingly. It seeks to ensure that evidence from evaluation is used in planning, budgeting and on-going project management. It provides a common language and minimum standards, and promotes the utilisation of evaluation findings to improve performance. It targets the public sector, evaluators outside the public sector, and training institutions who must ensure that people have the requisite skills and competences.

Evaluation is defined in the NEPF as:

The systematic collection and objective analysis of evidence on public policies, programmes, projects, functions and organizations to assess issues such as relevance, performance (effectiveness and efficiency) and value for money, and recommend ways forward.

The focus is on evaluation of policies, sector and cross-sectoral plans, programmes and projects. Six specific types of evaluation are defined: Diagnosis, Synthesis, Design evaluation, Implementation evaluation, Impact evaluation. These evaluations are at different stages – prior to an evaluation, during implementation, and after implementation.

The key elements of the framework are:

1. All large or strategic programmes, or those of significant public interest or of concern must be **periodically evaluated**.
2. **This requires minimum standards for planning** so that effective M&E can be undertaken. Key for this evaluation policy framework are sectoral and cross-sectoral plans (such as delivery agreements for outcomes), programme and project plans. The plans require effective diagnosis, standardised and good quality theories of change (how the outcomes will be achieved) and good quality measurable indicators. An immediate priority for evaluation is agreeing minimum standards for programme plans.
3. Three year and annual **national and provincial evaluation plans** will be developed.
4. All evaluations in the evaluation plan must be in the **public domain**, on the departmental and DPME website. They must be peer reviewed and results communicated to different audiences, including the relevant Parliamentary Committees.
5. All evaluations must include recommendations, and **plans** to address the recommendations must be produced by departments and their implementation must then be monitored.
6. DPME and Offices of the Premier will provide **technical support and quality control** for evaluations.
7. Appropriate **training courses** will be provided by PALAMA and universities to build evaluation capacity in the country.

Contents

Foreword	ii
Policy summary	iii
Glossary	v
Part A Introduction	1
1 Background	1
1.1 Introduction to the policy framework.....	1
1.2 Focus of this policy framework	2
2 Why evaluate?	3
3 Approach to evaluation	4
3.1 What do we mean by evaluation	4
3.2 How do we compare evaluation to related activities	5
3.3 Overview of evaluation terms	6
3.4 Need for good quality plans to evaluate against	7
Part B How we do evaluation?.....	9
4 Uses and types of evaluations.....	9
4.1 Introduction.....	9
4.2 Evaluative questions	9
4.3 The types/categories of evaluation we will promote across government	10
4.4 When should evaluations be done	11
4.5 Approach and methodologies.....	11
4.6 Calendar for evaluations in relation to new programmes	12
4.7 Evaluations of existing programmes	12
5 Ensuring quality plans at different levels to evaluate against.....	12
6 Assuring credible and quality evaluations	14
7 The process of evaluation	15
7.1 Pre-design and design	15
7.2 Implementation.....	16
7.3 Peer review process.....	16
7.4 Communicating results.....	16
8 Assuring follow-up.....	17
Part C How we make this happen?	18
9 Institutionalising evaluation in the Government system.....	18
9.1 Legal framework.....	18
9.2 Evaluation plan.....	18
9.3 Role and responsibilities	18
9.4 Planning and budgeting for evaluation	19
9.5 Standardised systems.....	20
9.6 Donor evaluation as part of the Government system	20
9.7 Optimising limited capacity	20
10 Management and coordination of evaluation across government.....	21
10.1 Evaluation Working Group	21
10.2 Implementation of these evaluation systems across government.....	21
Annexes	23
Annex 1 References and further sources of information	23
Annex 2: Evaluation Principles.....	24
Annex 3: Annotated glossary	25

Annex 4:	Approaches to evaluation.....	27
Annex 5:	Set of methods that can be used in different types of evaluations.....	29
Annex 5:	Related guidance notes/guidelines to be developed.....	30

Glossary

DPME	Department of Performance Monitoring and Evaluation
DPSA	Department of Public Service and Administration
ECD	Early Childhood Development
EU	European Union
GWM&ES	Government-wide M&E System
MFMA	Municipal Finance Management Act
MPAT	Management Performance Assessment Tool
NEPF	National Evaluation Policy Framework
OPSC	Office of the Public Service Commission
PETS	Public expenditure tracking system
PFMA	Public Finance Management Act
PMDS	Performance Management and Development System
RCT	Randomised control trial
SAMEA	South African M&E Association
SASQAF	South African Statistics Agency Quality Assurance Framework
TORs	Terms of reference
UN	United Nations

Part A Introduction

1 Background

1.1 Introduction to the policy framework

The Constitution (section 195) mandates that in the principles of public administration:

- Efficient, economic and effective use of resources must be promoted;
- Public administration must be development-oriented;
- Public administration must be accountable;
- Transparency must be fostered by providing the public with timely, accessible and accurate information.

While one of the ways government seeks to become effective is through focus on strategic outcomes, another is to promote better understanding of how its programmes and services are working, to monitor and evaluate how they are working in practice, and to feed back this information into better planning, revised budget allocations, improved management practices and implementation. In this way government can better achieve the political mandates that are set for it, and have a stronger impact on the huge development challenges facing the country. It is particularly important that evaluation is applied to these priority outcomes to ensure that they are being planned and implemented appropriately.

According to the Policy Framework for the Government-wide M&E System, 'M&E processes can assist the public sector in evaluating its performance and identifying the factors which contribute to its service delivery outcomes. M&E is uniquely oriented towards providing its users with the ability to draw causal connections between the choice of policy priorities, the resourcing of those policy objectives, the programmes designed to implement them, the services actually delivered and their ultimate impact on communities. M&E helps to provide an evidence base for public resource allocation decisions and helps identify how challenges should be addressed and successes replicated.'

Monitoring is the continuous and systematic collection, recording and reporting of information in order to track progress towards the achievement of the objectives of an intervention, and identify the need for corrective action. While monitoring asks whether things are being done right, evaluation is saying are we doing the right things, are we effective, efficient and providing value for money.

Evaluations contribute to the improvement of public policy interventions and expenditure programmes by providing evidence-based assessments of their relevance and performance. As such they are a critical component of effective policy and programming and budgeting cycles. 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. In applying its evaluation system across government, Chile found that 51% of programmes needed substantial revision. Failure to do so results in impaired implementation and effectiveness.

Effective monitoring and evaluation requires a systematic and rigorous approach to defining what you want to do, how you intend to get there (ie planning) and how you would know if you got there (good indicators). Much of South Africa's planning system is unregulated and has developed in an ad-hoc manner. The plans need to be monitored and evaluated at a wide range of levels. In most cases plans of some sort exist – but there is a lack of consistency in the quality of plans.

In South Africa the need for more systematic evaluation of policy interventions and expenditure programmes by Government is urgent. Public role players need to have better information on whether government is doing the right things in the right way to achieve its political mandate and objectives, and to understand what goes wrong when the results of policy interventions and public expenditure are below expectation.

Some of the challenges we face at present around evaluation include:

- Lack of clear policy and strategic direction around evaluation;
- A need to promote the use of knowledge from both evaluation and research – improving the knowledge base;
- Confusion on what is evaluation, performance auditing, research etc;
- Evaluation work exists but is not necessarily known, either within departments or externally;
- Lack of coordination between organisations and fragmentation of approaches;
- Poor quality plans making evaluation difficult;
- Inadequate use of evaluation, leading to a perception that it is a luxury;
- A lack of institutionalisation of evaluation in the government system.

The **problem statement** that this framework is trying to address is:

Evaluation is applied sporadically and not informing planning, policy-making and budgeting sufficiently, so we are missing the opportunity to improve Government's effectiveness, efficiency and impact.

This document aims to address many of these issues providing a framework with a clear champion (DPME). It completes the set of frameworks that underpin different dimensions of monitoring and evaluation in the public sector.

1.2 Focus of this policy framework

In this policy framework we are aiming to develop:

- A common language and conceptual base for evaluation in Government;
- An institutionalised system across Government linking to the planning and budget processes;
- Clear roles and responsibilities;
- Improved quality of evaluations;
- Utilisation of evaluation findings to improve performance.

The **purpose of this policy framework** is to:

- Define evaluation in the public sector;
- Establish an institutional framework for evaluations;
- Provide guidance on the approach to be adopted when conducting evaluations; and
- Provide for the publication of the results of evaluations.

The **target group** for this policy framework are:

- Political principals, managers and staff in the public sector who must improve their performance and incorporate evaluation into what they do;

- Other actors who need to be involved in the evaluation process, such as potential evaluators (including academics and other service providers);
- Training institutions, who will have to ensure that public servants understand evaluation and that we have a wider cadre of potential evaluators with the required skills and competences.

Evaluations can be undertaken internally by public institutions, or by an external agency. They can look at different units of analysis or plans, albeit an institution or a group of institutions, a policy intervention or an expenditure programme or sub-programme. This Framework will inform evaluations undertaken by the executive at national, provincial and local government level. It seeks to include evaluations conducted by independent oversight bodies such as the Public Service Commission and Auditor-General, but also recognises their independence as part of their Constitutional Mandates.

For the purposes of evaluation in the South African public sector no one evaluation methodology is favoured. This document acknowledges that evaluation can use a wide range of research techniques and data sources, depending on the evaluation field, the evaluation object and the evaluation questions at hand. However, evaluations must be systematic, structured and objective and must use defensible techniques for data collection and reliable data sources. They must also be available for use by policy-makers and not kept secret because the findings are uncomfortable.

This document clarifies the role of evaluations in relation to other performance management instruments and frames the evaluation function in terms of its scope, institutionalization, standards, process requirements, skill requirements, governance, financing and oversight.

The framework is in 3 parts – A which sets the approach and context, B which describes the evaluation system itself, and C, which is how we make this system happen.

2 Why evaluate?

This section aims to clarify why it is worth doing evaluation, and what you might be doing it for.

Evaluation can be undertaken in Government for a number of **purposes**. Four primary uses/purposes of evaluation and some of the questions one may ask in these different uses could be:

- **Judge merit or worth of something:** Was the programme successful, or is the proposed policy or programme likely to be successful? Was it effective? Did the intended beneficiaries receive the intervention? Did it impact on their lives or what is the likely impact of the policy/programme? Will the policy/programme impact differentially on different sectors of the population, etc.
- Improving policy or programme **performance** (evaluation for learning): this aims to provide feedback to programme managers. Questions could be: was this the right intervention for this objective, was it the right mix of outputs, what is the most effective way to do X?
- Evaluation for improving **accountability**: where is public spending going? Is this spending making a difference?
- Evaluation for **generating knowledge** (for research): increasing knowledge about what works and what does not with regards to a public policy, programme, function or organization, which allows governments to build up an evidence base for future policy development.

In this framework our primary purpose is around **improving performance**, but this also involves questions of judgement.

Some of the potential **benefits** which can be obtained include:

- Improved learning and its feedback into policy and implementation;
- Ensuring policy and management decisions are based on evidence;
- Better understanding on which programmes are cost-effective;
- Understanding and so better able to overcome institutional bottlenecks to improved impact (adapting programmes);
- Better understanding and management of risks.

Ultimately these should result in:

- Strengthening the culture of use of evidence to improve performance;
- Better policy and planning;
- Better resource allocation;
- Minimising negative unintended consequences of policy;
- The public being aware of what government does (public accountability).

And finally these should result in better service delivery and the achievement of government objectives.

3 Approach to evaluation

3.1 What do we mean by evaluation

Evaluation is a branch of applied research that attempts to identify cause-effect relationships within a specific context (PALAMA, Basic M&E Course Manual, p17). Evaluation can be defined in different ways eg:

- Evaluation is an objective process of understanding how a policy or other intervention was implemented, what effects it had, for whom, how and why (Magenta Book, UK);
- The systematic and objective assessment of an on-going or completed project, programme, or policy, its design, implementation and results (OECD);

For this Evaluation Policy Framework **evaluation** is defined as:

The systematic collection and objective analysis of evidence on public policies, programmes, projects, functions and organizations to assess issues such as relevance, performance (effectiveness and efficiency) and value for money, and recommend ways forward.

It is differentiated from **monitoring**:

Monitoring involves the continuous collecting, analysing and reporting of data in a way that supports effective management. Monitoring aims to provide managers with regular (and real-time) feedback on progress in implementation and results and early indicators of problems that need to be corrected. It usually reports on actual performance against what was planned or expected (adapted from the Policy Framework on GWM&ES)

So while monitoring continually tracks what is happening, evaluations are periodic and seek to see what has been achieved, and to understand why. Evaluation has the element of judgement, and must be against objectives or criteria (Fournier in Rossi et al p70). The **characteristics** of good evaluation are (EU, 2007):

- evaluations should be **analytical** - they should be based on recognised research techniques;
- evaluations should be **systematic** - they require careful planning and consistent use of the chosen techniques;
- evaluations should be **reliable** - the findings of an evaluation should be reproducible by a different evaluator with access to the same data and using the same methods of data analysis;
- evaluations should be **issue-oriented** - evaluations should seek to address important issues (such as) relevance, efficiency and effectiveness;
- evaluations should be **user-driven** - this just means that successful evaluations should be designed and implemented in ways that provide useful information to decision-makers, given the political circumstances, programme constraints and available resources.

While evaluation is often seen as only occurring at the end of an intervention, different forms of evaluation can be undertaken at different phases, from prior to an intervention (sometimes called ex-ante evaluation), during an intervention (eg to check whether the activities are leading to outputs, and outputs to outcomes), and after the intervention has completed. This is discussed further in section 4. Evaluation evidence is seldom self-evident, but needs interpretation and judgement about its relevance and applicability to problem at hand.

The Policy Framework for the Government-wide Monitoring and Evaluation (GWM&E) has 7 guiding principles for monitoring and evaluation. Annex 2 shows how these can be applied.

Table 1: Evaluation Principles

- 1 Evaluation should be development-oriented and should address key development priorities of government and of citizens.
- 2 Evaluation should be undertaken ethically and with integrity.
- 3 Evaluation should be utilisation-oriented.
- 4 Evaluation methods should be sound.
- 5 Advance government's transparency and accountability.
- 6 Inclusion and participation.
- 7 Learning.

3.2 How do we compare evaluation to related activities

There is no hard line between evaluation, monitoring, performance audit etc. Table 1 below suggests the core elements of each in terms of the activity involved, and the objective. Box 1 also provides some clarification on the differences.

Table 2: Spectrum of results-based management activities (adapted from Ajam, 2011)

	Activity	Objective
Inspection/investigation	Detects wrong doing and verifies information	Control and compliance
Performance audit	Oversight through assessing performance	Accountability, control, compliance and learning
Monitoring	On-going tracking of progress and initiating corrective action	Management, accountability, corrective action
Review ¹	Form of evaluation, but less comprehensive, rigorous and in depth than evaluations	Learning, improving performance
Evaluation	Periodic assessment to explain performance	Learning, accountability, improving performance
Research	Testing hypotheses/propositions through observation of reality	Learning/knowledge creation only (no accountability focus)

¹This does not apply to Comprehensive Expenditure Reviews, which are comprehensive and have a specific purpose.

Box 1: Differentiating evaluation (from EU, 1997)

Firstly, evaluations differ from scientific studies. Both should be analytical, systematic and reliable. However, whereas scientists may undertake research in order to expand the sum of human knowledge and frequently confine themselves to one highly specialised discipline, evaluations are undertaken for more practical reasons. They are intended to be of practical use by informing decisions, clarifying options, reducing uncertainties and generally providing information about programmes within their own specific contexts. They also can draw on a wide range of analytical approaches.

Neither is evaluation the same as audit. Audit is primarily concerned with verifying the legality and regularity of the implementation of resources (inputs) in a programme. Evaluation, on the other hand, is necessarily more analytical. It examines the programme from the point of view of society (defined from different possible perspectives). It looks at the validity of the strategy followed and whether objectives are appropriate given the problems to be solved and the benefits to be achieved. Auditors tend to have coercive powers, sometimes defined in legal texts, whereas evaluators must often rely on “good will” and the power of their arguments.

Audit has traditionally covered activities such as the verification of financial records (financial audit). A more recent innovation is known as performance audit, which is conceptually closer to evaluation. Performance audit is strongly concerned with questions of efficiency (of a programme’s direct outputs) and good management. Performance audit and evaluation share the same aim of improving the quality of programmes, but evaluation goes much further. It also looks at issues such as sustainability, relevance and the longer-term consequences of a programme (evaluation seeks to explain).

Finally, evaluation must be distinguished from monitoring. Monitoring examines the delivery of programme outputs (the goods and services produced by the programme) to intended beneficiaries. It is a continual process, carried out during the execution of the programme, with the intention of immediately correcting any deviation from operational objectives. Evaluation, on the other hand, is specifically conducted at a discrete point in the life cycle of a programme, and consists of an in-depth study. Monitoring is of key importance to improving programme performance, and successful evaluation often hinges upon successful monitoring, for example because monitoring often generates data which can be used in evaluation. (Evaluating EU Expenditure Programmes: A Guide)

3.3 Overview of evaluation terms

A wide range of terms are used to describe evaluations. Managers commissioning evaluations will be exposed to some of these terms so these are introduced here, with a fuller set in annex 3.

Cost-benefit-analysis	An analytical procedure for determining the economic efficiency of a programme, expressed as a relationship between costs and outputs, usually measured in monetary terms.
Counterfactual	What would have happened if an intervention did not exist or if some other policy initiative was implemented.
Ex-ante evaluation	Ex ante evaluation is a process that supports the preparation of proposals for interventions. Its purpose is to gather information and carry out analyses that help to define objectives, to ensure that these objectives can be met, that the instruments used are cost-effective and that reliable later evaluation will be possible.
Ex-post (or post-hoc evaluation)	Assessing/evaluating quality after a programme or institution has been in operation in order to establish strengths and weaknesses
Formative evaluation	Evaluation activities undertaken to assist learning and provide information that will guide programme improvement, especially in terms of how, why, and under what conditions a policy will work or has worked.

Meta-analysis	Determining the overall balance of evidence from different sources and studies (usually expressed in non-quantitative, narrative form) (see statistical meta-analysis).
Multi-criteria analysis ²	Multi-criteria analysis allows us to formulate judgements on the basis of multiple criteria, which may not have a common scaling and which may differ in relative importance.
Programme theory (usually linked to the theory of change below)	The set of assumptions about the manner in which programme relates to the social benefits it is expected to produce and the strategy and tactics the programme has adopted to achieve its objectives.
Public Expenditure Tracking Surveys (PETS)	Assesses whether resources reached the intended beneficiaries and whether they result in better services. It can be combined with a quality service delivery survey.
Quasi-experiments	An impact evaluation design in which intervention and control groups are formed by a procedure other than random assignments.
Research synthesis	A way of establishing what is already known about a policy initiative, especially its achieved impact and its implementation challenges in other policy environments.
Statistical meta-analysis	The aggregation and generation of cumulative statistical estimates of impact from combining the results of different comparable studies.
Summative evaluation	Evaluation activities undertaken to render a summary judgement on the impact of the programme's performance, eg specific goals and objectives were met.
Theory of change	A tool that describes a process of planned change, from the assumptions that guide its design, the planned outputs and outcomes to the long-term impacts it seeks to achieve.

3.4 Need for good quality plans to evaluate against

Evaluation asks questions as to what impact there has been, why things have happened, whether certain outputs were needed to achieve those outcomes, how well resources have been used in achieving certain outcomes. If the plan is not clear, and the indicators are not clear, then it is difficult to evaluate. A challenge in South Africa is that the responsibilities for planning are diffuse, the plans are of variable quality, and the plans are often not aligned with each other.

Therefore this framework has a short section on planning, indicating what planning systems must be established. This framework does not go into detail into these, which will be pursued in subsequent policy frameworks.

The planning model must be based on a results based management approach, as shown in Figure 1 from Treasury (2007). This shows a focus on impacts and outcomes, achieved by certain outputs, which in turn require certain activities to be undertaken. This represents what is sometimes called a theory of change or logic model, which is:

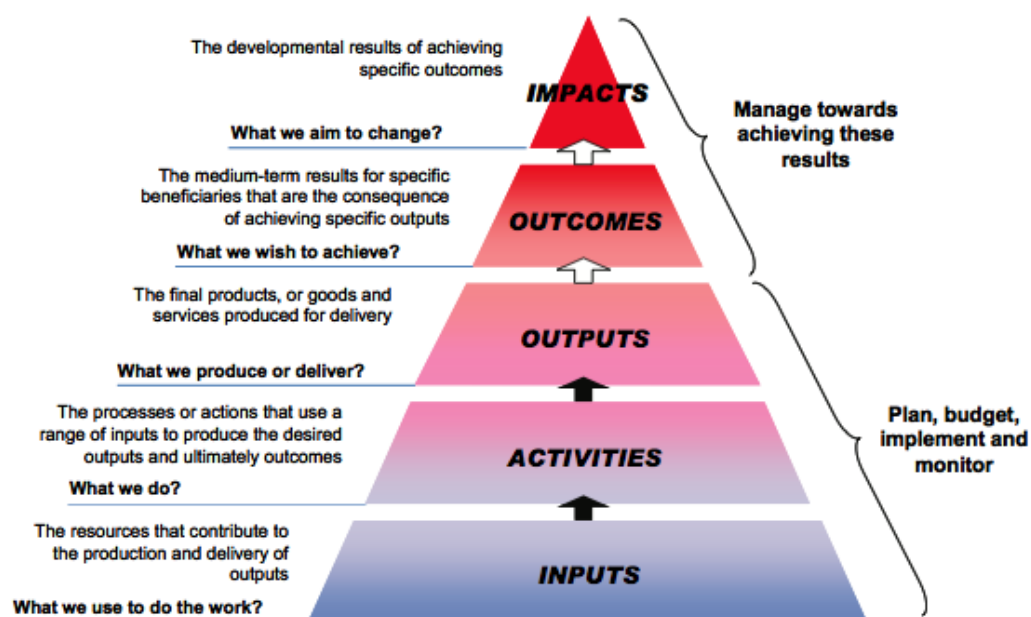
A tool that describes a process of planned change, from the assumptions that guide its design, the planned outputs and outcomes to the long-term impacts it seeks to achieve.

In other words this is a model of cause and effect, if we do this, we should achieve that. That is what a plan is – a hypothesis that if we do what we suggest, we will achieve certain objectives and targets. The plan is only as good as the theory of change. If it is flawed we may do these activities, we may achieve these outputs, but it will not result in the change we desire.

²EU guideline (1997) already cited earlier

Evaluation must be used to test this logic model – were the outcomes achieved and were they correct, where the outputs correct to reach those outcomes, were the activities correct to achieve those outputs. In other words apart from asking what was achieved, we ask why was it achieved, and how was the use of resources in the process.

Figure 1: The results-based management pyramid (National Treasury 2007)



Part B How we do evaluation?

4 Uses and types of evaluations

4.1 Introduction

This section aims to cover the types of evaluations that can be considered, the different approaches to evaluations, and then suggests the types of evaluations that we will concentrate on as part of South Africa's government-wide evaluation system. It also suggests when these evaluations should be conducted as part of the programme or policy cycle.

In considering evaluations a number of factors need to be taken into account:

- The **object** of evaluation – what is to be evaluated;
- The **purpose** of the evaluation;
- The type of **questions** being asked;
- The **approach** to be taken;
- The **methodologies** to consider.

The unit of analysis/**objects** range from: policies; sector plans; programmes; projects; organisations/ institutions; to individuals. Table 3 takes the **purposes** suggested in section 1 and gives some related sub-purposes.

Table 3: Some primary uses/purposes of evaluation studies (adapted from Patton, 1997, in Babbie and Mouton, 2007)

Uses or purposes	Examples of reasons or decisions for use
Judge merit or worth	Assessing impact Quality control Cost-benefit decisions Deciding a programme's future Decisions on accreditation/licensing
Improve programmes	Assessing learning Understanding causal links Identifying strengths and weaknesses Quality enhancement Improving cost-effectiveness Managing more effectively Adapting a model locally
Generate knowledge	Generalisations about effectiveness Extrapolating principles about what works Building new theories and models Informing policy
Improve accountability	Assessing impact Assessing compliance/audit Improve transparency Accountability

4.2 Evaluative questions

Typical evaluative questions may range from:

- Do policies or programmes have clear objectives and theory of change? (later referred to as a design evaluation question);

- Are the steps involved in delivering a service efficient? (later referred to as an implementation evaluation question);
- How have beneficiaries' lives changed as a result of a programme or policy? (impact evaluation question);
- Is this programme providing value for money? (cost-effectiveness question)

Some key issues for evaluation are (EU, 1997):

- a) Relevance** - to what extent are the policy, programme or project's objectives pertinent in relation to the evolving needs and priorities of government?
- b) Efficiency** - how economically have the various resource inputs been converted into tangible goods and services (outputs) and results?
- c) Effectiveness** - how far have the outputs of the policy, programme or project contributed to achieving its intended objectives?
- d) Utility** - how do the results of the policy, programme, or project compare with the needs of the target population(s)?
- e) Sustainability** - to what extent can the positive changes be expected to last after the programme has been terminated?

4.3 The types/categories of evaluation we will promote across government

Different countries use different terms to describe evaluations. Based on the different objects, purposes and questions, a set of standard types of evaluation are proposed to be used across government in South Africa. This will help to develop a common language, and establish standard procedures. Table 4 summarises these types and their uses.

Table 4: Summary of types of evaluation across government

Type of evaluation	Covers
Diagnosis	This is preparatory research to ascertain the current situation prior to an intervention and to inform intervention design. It is important to identify what is already known about the issues at hand, and about the likely effectiveness of different policy options. A research synthesis exercise, such as a rapid evidence assessment, is an appropriate way of doing this.
Synthesis	This is a rapid assessment used to obtain a general and comparative perspective of the design, strategic planning, coverage and targeting mechanisms, operation, beneficiary perception and results. It uses secondary data including existing evaluations and surveys. It may be used to see what areas need more rigorous evaluation. It also assesses the quality of the indicators
Design evaluation	Used to analyse the inner logic and consistency of the programme, either before a programme starts, or in the first year of implementation. This is quick to do and uses only secondary information and should be used for all new programmes. It also assesses the quality of the indicators.
Implementation evaluation³	Aims to evaluate whether an intervention's operational mechanisms support achievement or not and understand why. Looks at activities, outputs, and outcomes, use of resources and the causal links. It builds on existing monitoring systems, and is applied during programme operation to improve the efficiency and efficacy of operational processes. It also assesses the quality of the indicators
Impact evaluation	Measures changes in outcomes – often using experimental or quasi-experimental designs, but should be combined with implementation evaluations to understand the cause effect relationships. The challenge is around control groups, which can be difficult to establish, and may involve ethical questions in refusing a certain group access. These are expensive typically costing around R10-20 million, similar to SA.

³ This is sometimes called a process evaluation

Table 5 shows how these different types of evaluation can be applied to different focuses or objects of evaluation, and some examples of these.

Table 5: Type of evaluation applied to different objects of evaluation (those in green are covered in this framework, light green to be covered later)

Unit/object	Type of evaluation	Example	Covered/not in this framework
Government-wide	Synthesis + research	10 and 15 year reviews, expenditure reviews	Not covered
Policy/sector plans/ thematic plans	Diagnosis/Synthesis/Design /Implementation/Impact	ECD/Evaluation of National Strategic Plan for HIV/AIDS	Covered
Programmes		Child Support Grant impact evaluation	Covered
Projects			Covered
Organisational/ institutional	Combination of performance audits and process evaluation	State of the public service report by OPSC, performance audits by the Auditor General, MPAT	To be covered in later versions of the policy framework
Individuals	Tests	Annual National (learner) Assessment	Not covered
	Self/supervisor assessment	PMDS	

4.4 When should evaluations be done

Evaluations should be part of all stages of a policy or programme cycle. Table 6 shows this prior to an intervention starting, during the intervention and after the intervention is completed.

Table 6: Use of different types of evaluation before, during and after interventions

Before the intervention (pre-programme)	During the intervention	After the intervention is complete (or at major phases)
Diagnostic assessment to understand the problem and inform the design or policy development process	Design evaluation to assess the quality and implementability of the intervention design in practice	Impact evaluations possibly including implementation evaluation to assess why impacts have been achieved or not and implications for the future. Builds on baseline.
Synthesis to review current evidence	Synthesis to review evidence during programme operation, building on monitoring information	
(the intervention is planned, may include a feasibility study)	Implementation evaluation to see whether a programme or policy's operational mechanisms support achievement or not , and understand why.	Synthesis to review evidence during programme operation, building on evaluation data of components
Design evaluation to assess the quality and implementability of the proposed intervention design		
Baseline for impact evaluation		

4.5 Approach and methodologies

A range of different approaches can be applied to evaluations. This will influence the way they are planned and implemented, as well as the methodologies and tools selected. An evaluation may include features of several approaches. These approaches are mentioned in Annex 4.

There are many **methodologies** that can be used to undertake evaluations, including quantitative, qualitative and mixed research methodologies and their use will depend on the approaches selected. Each type of evaluation design can combine different methods of data collection and data analysis informed by the key evaluation questions to be addressed. Annex 5 shows a possible menu of methods. The choice of the methodology should be informed by (MFA, 2007):

- The team's knowledge and experience of the techniques (availability of expertise);
- The need for specific data;
- The prerequisites for using the tool, eg relevance, cost-effectiveness and efficiency;
- The time frame.

4.6 Calendar for evaluations in relation to new programmes

A calendar for new interventions such as policies and programmes is:

Year 0	Do Diagnostic (to inform intervention design), design evaluation to check design, and baseline for impact evaluation (if one is planned)
Year 1	Design evaluation to confirm intervention design is correct
Year 2	Implementation Evaluation to check if intervention is working appropriately
Year 3	Interim impact evaluation , Implementation Evaluation to check if intervention is working appropriately
Year 4	
Year 5	Full impact evaluation

4.7 Evaluations of existing programmes

In principle **all programmes should be evaluated on a 5 year cycle**. In practice due to limited capacity this should start with programmes that are:

- I. **Large** (eg over R500 million) or covering a large proportion of the population, and have not had a major evaluation for 5 years. This figure can be reduced with time;
- II. Of **strategic importance**, and for which it is important that they succeed. If these have not been evaluated for 3 years or more, an implementation evaluation should be undertaken;
- III. **Innovative**, from which learnings are needed – in which case an implementation evaluation should be conducted;
- IV. Of significant **public interest** – eg key front-line services;
- V. Any programme for which there are real **concerns** about its design should have a design evaluation conducted.

In the case of I or II, if no formal impact evaluation has been conducted, one should be planned. DPME will develop a **3 year and annual evaluation plan** which will include categories I-III above.

5 Ensuring quality plans at different levels to evaluate against

A government that **focuses on results or performance**, must know what it wants to achieve, and must plan how it will achieve these results. It must monitor and track progress on how it is doing, and learn the lessons from evaluations as to how to do it better. A performance-based government must also account to the public on the achievement of key results that are of interest to the public, and so evaluations must be made public.

If government strategic priorities are to be achieved then these priorities must be traceable across the levels of government, through the different plans and budgets. This can be called a '**line of sight**'. This requires the following set of plans:

- A **long-term vision and plan** (which the National Planning Commission is taking forward).
- A **plan for a term of office** (5 years), at national, provincial and local level. This must identify the key priority outcomes for that term of office. At national level this is currently provided by the MTSF, although this is not a substantive plan. The Provincial Growth and Development Strategy (PGDS) is an example of such a plan at provincial level, and the IDP at local government level.
- **Sectoral (eg health) and cross-sectoral (eg employment) plans** (as in the outcome delivery agreements).
- Plans for **implementation programmes** (eg maternal health), which may be within one department or cross-sectoral (we differentiate here from budget programmes as identified in a departments budget structure).
- **Project plans** define more limited components, eg training in home-based care, or building a new hospital.
- **Departmental strategic plans/Annual Performance Plans and operational plans** must incorporate both the sectoral/cross-cutting priorities, and the implementation programmes and projects if these are to be implemented, monitored and evaluated, and if the line of sight is to be achieved with the sectoral/cross-sectoral plans and the implementation programmes.

As Table 5 shows, evaluations can be done of all of these plans. Good quality plans require minimum standards, which make the plans more likely to be achieved, and evaluation much easier. Assuming the results-based management (RBM) approach (see Figure 1 on page 8), plans must include:

- A good quality **diagnostic analysis** of the status quo, the current situation and the forces at play, and which are likely to be the main strategic drivers of change. This is the role of the Diagnostic Evaluation.
- A clear **theory of change or logic model** showing the outcomes, the outputs needed to achieve those outcomes, the activities to realise those outputs, the resources required to implement this (see Figure 1), and the assumptions on which this is based. A good diagnosis is needed to understand the situation well and be able to build a robust theory of change. A key function of evaluations is to check if this theory of change is working, because if it isn't then what is being done (activities) will not result in the desired outcomes and the efforts of government are wasted. The Design Evaluation aims to test the theory of change, before the intervention is implemented, and in the first year of implementation. The Implementation Evaluation also tests how these are working in practice, what is working and why.
- Good quality **indicators** showing what you will monitor, and appropriate targets. Poor quality indicators and targets makes monitoring and evaluation very difficult.
- The **roles and responsibilities** of national departments, provinces, municipalities and other actors for achieving these outputs and outcomes. The Implementation Evaluation will also review these roles and responsibilities.
- And for lower level plans, these will include **budgets**.

In many cases evaluations are likely to be of implementation of programmes, services (a form of programme) or projects, which are the key vehicles for implementation. There is no standard definition of the minimum standards of programme plans and this must be developed as an urgent priority.

6 Assuring credible and quality evaluations

This issue relates to the principles of credibility and so perceived independence of findings. There is a trade-off between degree of ownership (highest if conducted internally) and degree of independence and external credibility (highest if external to the organisation, and external to government). The terms internal and external can be interpreted in different ways.

- Internal can mean the programme staff, or the organisation involved;
- External can mean external to programme, external to the organisation, or non-government.

The terms can also refer to initiation – who requests the evaluation in question (programme staff, the organisation itself, or an external body such as the OPSC), as well as who actually undertakes the evaluation. **We will use external to mean outside the organisation in question.**

Table 7 shows the different possibilities of internal or external initiation and undertaking of evaluations. In most cases the evaluations should be internally initiated, or as a partnership with an external body such as DPME. In this way ownership of the evaluation is retained and there is most likelihood of a successful evaluation and one that is followed-up.

Table 7: Internal or external initiation and conducting of evaluation

Undertaken by who	Initiated by who	
	Internally initiated	Externally initiated
Undertaken internally to the institution of study	Undertaken within the institution either by staff of the programme in question, or other relatively independent staff	Evaluation initiated by external body eg Presidency, and institution asked to do internal evaluation to improve performance
Undertaken externally	External service provider or government institution to ensure independence eg impact or implementation evaluation of programme	Evaluation initiated by external body eg Presidency or OPSC, and commissioned to external service provider or government institution

There is also another possibility where an evaluation is initiated externally to the department and undertaken by staff of the evaluating body, eg OPSC or DPME. There can also be a hybrid where part of the evaluation is conducted externally. Note section 6 talks of approaches to assure quality during the evaluation process.

Table 8 discusses who should undertake the types of evaluation we have specified, assuming they are internally initiated.

Table 8: Who should undertake the different types of evaluations

Type of evaluation	Implement internally/externally
Diagnosis	It is important to identify what is already known about the issues at hand, and about the likely effectiveness of different policy options. This is a research-type exercise and will probably need to be contracted out. A research synthesis exercise, such as a rapid evidence assessment, is an appropriate way of doing this. Diagnosis can be quite complex if a large programme is being envisaged, or to design a new policy.
Synthesis	This is a rapid evaluation and can be internal or external. It does require stronger skills and may need to be contracted out.

Type of evaluation	Implement internally/externally
Design evaluation	This is a rapid evaluation and could well be done internally. If a large programme it may be sensible if DPME undertake this as a cross-check. There is no need to contract out unless capacity is stretched.
Implementation evaluation⁴	This can be conducted internally if there is a strong skills base, as internal ownership and learning is very important. However in most cases will need to be undertaken by an external body but with a strong facilitation aspect to maximise learning.
Impact evaluation	External credibility and independence is very important and these will need to be contracted out.

7 The process of evaluation

This section outlines the different stages of the evaluation process.

7.1 Pre-design and design

A necessary, although not sufficient, condition for a quality evaluation is a systematic process that ensures that the evaluation is properly understood by the evaluation client, evaluator and the evaluation object (eg a programme), that it is properly focused, appropriately designed and undertaken and will deliver credible, reliable, valid information in a way which promotes learning.

Pre-design

Preparation for any evaluation requires investment in time and thought. The organisation needs to decide precisely why and when to conduct an evaluation. This needs to begin early in the programming cycle, especially for new programmes. The evaluation head must determine and clarify the object of the evaluation, draft the terms of reference and go through an appropriate process to select an evaluation team and team leader. In addition, more often than not evaluation of results requires the involvement of other partners or stakeholders beyond the line function that manages the programme, the department or entity. So it is advisable that key partners (such as national departments, civil society organisations, UN agencies, private sector, donors and academic institutions) are involved in each step of the evaluation process.

Developing TORs

At a minimum, it is expected that terms of reference for all evaluations will contain the objectives, scope, products/deliverables, methodology or evaluation approach, evaluation team and implementation arrangements. A set of standard TORs will be developed by DPME to provide guidance. The TORs should retain enough **flexibility** for the evaluation team to determine the best approach to collecting and analysing data. The TORs, for example, might suggest a combined approach to questionnaires, field visits and interviews – but the evaluation team should be able to revise this approach. Since the terms of reference involves strategic choices about what to focus on, they should be reviewed by key stakeholders, including the Evaluation Unit in DPME.

Selecting service providers

The rules and regulations governing supply chain management practices in a particular department should be applied. For evaluation, where quality of methodology and research competence is critical, it is important that the selection and choice of the appropriate service provider is not determined only by price but that technical competence is predominant.

⁴ Often referred to as a process evaluation

DPME will establish a national panel of approved evaluators from which service providers can be selected.

7.2 Implementation

Data quality

The errors and shortcomings of the data directly influence the analyst's ability to reflect the real world which the data are meant to describe. The analyst cannot take the data at face value. When examining data quality, whether primary or secondary sources, the department/entity and evaluator need to look at what (if any) appropriate data are available and how useful the data actually are for analysis. There is no point in looking at something for which no data exists. Each data set may have strengths and limitations. To judge the quality of data, Statistics South Africa's SASQAF standards should be used, of relevance, accuracy, timeliness, accessibility, interpretability, coherence, methodological soundness and integrity.

Inception phase

There should be an inception phase which shows the evaluators understand the evaluation object and theory of change, provides for the methodology to be finalised and to reach agreement on the evaluation scope, questions, methodology, process, reports, cost and payments. A finalised document must be approved by the project head and steering committee.

Advisory/steering group

If a big study there must be an advisory or steering group. An advisory group provides technical support, advice and expertise and a steering group that project manages the policy development and its monitoring and evaluation. This should be chaired by the person responsible for the evaluation team in the commissioning institution and include the evaluation team leader as well as other relevant stakeholders that could be affected by the evaluation results/findings, e.g. other government departments, civil society organisations, donors, as well as international organisations or agencies like UN.

Reporting

There need to be regular technical meetings between the evaluation team and the evaluation steering committee or organisational team so that challenges can be addressed.

7.3 Peer review process

This should be established for external evaluations to ensure they are credible. DPME will establish a national panel of evaluators who are methodology experts plus drawing in subject matter specialists for the specific evaluation. Two people should be critical reviewers of each evaluation, which should be budgeted for as part of the evaluation budget. They should give feedback in a session with the department. Note that OPSC does some review of the quality of evaluations. This may be taken further in future.

7.4 Communicating results

Since the evaluation will have different messages for various stakeholders and audiences it is imperative that these messages are drawn out potentially in a range of documents for different audiences – political (very short and punchy), different technical audiences, beneficiaries etc. The messages for political principals and the executives should be very concise and highlight key evaluation questions, findings and recommendations. Some specific things to be done:

- The **1/3/25 rule** applies - a one page policy summary of key policy messages should be produced, a 3 page executive summary, and a 25 summary report from what

might be a very long evaluation report. It is likely that only the 1/3/25 reports will be read;

- The department/entity must ensure that the full evaluation reports are **posted in their websites**. If they have riders to the evaluation results, they should also post those riders on their website;
- A copy of the evaluation must be submitted to **DPME** for lodging on DPME's website, as well as any comments the departments have on the evaluation
- Copies should be **sent to partners** in the evaluation in question.

8 Assuring follow-up

Evaluations are a positive tool for the improvement of institutional performance, policies, programmes and projects. Following up on evaluation findings and recommendations is therefore a crucial stage in the evaluation process. However, knee-jerk responses to evaluation findings and recommendations should be avoided. It is possible that recommendations arising from evaluation findings may be inappropriate, unfeasible or not in line with government priorities, especially if the evaluators are unfamiliar with implementation realities. On the other hand, programme managers or policy makers may find evaluation reports unpalatable.

The following steps should be followed in order to promote the utilisation of evaluations for improving the quality of interventions. Ministers, Heads of institutions, programme or projects managers (where applicable) should:

- a) Analyse the findings and recommendations of an evaluation report.
- b) **Prioritise recommendations** that are clear, important, relevant feasible and implementable.
- c) Be given the right to respond to the findings and recommendations of an evaluation report, if so desired.
- d) Negotiate the findings and recommendations with evaluators, if necessary.
- e) Develop a strategy for the **dissemination** of the evaluation report, including publishing evaluation reports on relevant websites, sharing findings with key stakeholders as well as the media.
- f) Prepare an **improvement plan** in response to the evaluation. This improvement plan should be lodged with Offices of the Premier if provincial in nature, and DPME if national.
- g) Undertake the necessary actions (such as changes in institutional work processes, policy or programme review, amendments to implementation strategies, changes to internal budget allocations) to improve the functioning of an institution or the delivery of a programme or project.
- h) **Monitor the implementation** of these recommendations and report to DPME on a 6 monthly basis on progress on implementing these recommendations.
- i) Treasury should utilise the findings and recommendations of the evaluation report as a source of evidence to support it in **budget allocations**.

Part C How we make this happen?

9 Institutionalising evaluation in the Government system

9.1 Legal framework

The Constitution of the Republic of South Africa (1996), the 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) provide a legal basis for the executive to manage the performance of public policies, programmes and institutions efficiently. These Acts also provide a legal basis for undertaking the different types of evaluation. There is little specific legislation for strategic planning, monitoring and evaluation, apart from around departmental strategic plans, annual performance plans, and IDPs. An immediate priority is to define the minimum standards for programme and project plans.

9.2 Evaluation plan

A national evaluation agenda with a 3 year and annual national evaluation plan will be developed by DPME starting with 2012/13, including large, strategic and innovative programmes and policies suggested by Departments (see 4.6 and 4.7 for categories for new and existing programmes). Offices of the Premier must also draw up evaluation plans in provinces. These plans will inform the evaluation community inside and outside of government. This agenda will lay broad parameters that would be implemented via the annual evaluation plan. The evaluation plan will specify from a national and provincial perspective what needs to be done. Government institutions can choose to do additional evaluations.

9.3 Role and responsibilities

Departments and public institutions have the responsibility to incorporate evaluation into their management functions as a way to continuously improve their performance. They need to:

- Ensure there is an evaluation **budget** in all programmes (see 8.4) and a **plan** over 3-5 years for which evaluations will be undertaken, and the form of evaluation;
- National Treasury needs to ensure that there are additional budgets to support evaluation, eg in Offices of the Premier and DPME;
- Ensure there are **specific structures** within the organisation entrusted with the evaluation role, and with the required skills. This could be an M&E Unit, a research unit, or a policy unit.
- Ensure that the results of evaluations are used to **inform planning and budget** decisions, as well as general decision-making processes. Thus the results of evaluations must be discussed in management forums and used to guide decision-making.

In terms of specialist functions in supporting the evaluation system key players include DPME, Treasury, DPSA, PALAMA.

DPME is the custodian of the evaluation function in Government. This role includes:

- **Standard setting**, with the development and publication of suitable standards and guidelines;
- **Pooling of knowledge** emerging from evaluations, publishing all evaluations in the public sector;
- **Quality assurance** of evaluation processes, and products;

- Capacity building and **technical assistance**, ensuring suitable courses are established and providing technical assistance to the public sector on evaluation;
- Ensuring that evaluation is adding value and that the benefits outweigh the costs;
- **Promotion** of evaluation in government.

Offices of the Premier should play a similar role at provincial level, accessing support from DPME as needed.

National Treasury has to assure value for money when it allocates budgets. To this end it needs to see that:

- Plans and budgets are informed by evidence, including from evaluations;
- Ensure that cost-effectiveness analyses are undertaken, and that suitable methodologies employed.

Provincial Treasuries should play a similar role at provincial level.

DPSA has to see that the results of evaluations that raise questions around the performance or structure of the public service are addressed, and that evaluation is budgeted for.

OPSC has a specific independent role in the evaluation process, reporting directly to Parliament, but is also a source of expertise in helping to build the quality of evaluation and improving the performance of government. It is therefore an important partner.

Auditor-General is also an independent body, but is an important player in its role of performance audit, which is close to evaluation. The differences are explained in section 4.3, but the AG focuses primarily on compliance, and less on explaining why challenges are occurring. However their role is an important part of the evaluative framework of government, and an important spur to improving performance.

PALAMA is responsible for developing capacity-building programmes around M&E across government.

Universities are also an important actor in that they provide the tertiary education including evaluation, and the appropriate skills development to support this framework needs to be provided in these courses. This should not only be for specialised M&E courses, but also courses in public administration, development studies etc. They and other **research service providers** also will supply many of the evaluators, particularly where sophisticated research methodologies are needed, eg around impact evaluations, and undertake research which is closely allied to evaluation, and can help to inform research processes.

The South African M&E Association (**SAMEA**) – is the national association of people and organisations involved in M&E. They need to support the development of systems and capacities, and are an important forum for learning and sharing.

9.4 Planning and budgeting for evaluation

Evaluations will not be realised unless they are budgeted for. Evaluation costs are typically 1-5% of a programme budget, depending on size (large programmes need proportionally less). This needs to be factored into annual budgets and the MTEF. This is particularly important where large budgets are needed, eg for impact evaluations. Typical budgets required are around:

- Design evaluation – can be undertaken internally, if external R50-100 000;
- Synthesis – can be undertaken internally, if external cost of R100-300 000;

- Implementation evaluation – R100 000 to R1 million depending on scale and to what degree primary research is undertaken. If a sophisticated methodology such as PETS is used this would be more;
- Impact evaluation – R5-15 million depending on scale of sampling and sophistication of research design.

These are only indicative, and will depend on complexity and the rates of service providers. The allocation of budgets in conditional grants will need to consider that evaluation budgets may well need to be allocated to provinces.

9.5 Standardised systems

One of the ways to assure quality, particularly when there is limited capacity, is to avoid reinventing tools. DPME will issue specific guidance notes and guidelines setting standards for evaluation to complement this Framework. These will include such elements as:

- Standardised **terms of reference** for different types of evaluation;
- Standard **contract** formats for evaluation by external service providers;
- Models for **programme design** (logic models) which facilitate evaluation;
- Formats for programme **rules of operation** to provide some standardisation of how programmes operate;
- Standardisation of **evaluation processes** to improve quality, such as use of inception reports;
- Guidelines for following-up evaluation recommendations (if evaluations are not followed-up they are a waste of money);
- National panel of evaluators, possible with standardised fee rates.

9.6 Donor evaluation as part of the Government system

Donors have often been the driving force for improving monitoring and evaluation. However they have also often instituted their own parallel systems, which put major strains on government capacity. The Paris Declaration and Accra Agenda for Action commit partners to the following fundamental principles for making aid more effective:

1. **Ownership:** Developing countries set their own strategies for poverty reduction, improve their institutions and tackle corruption.
2. **Alignment:** Donor countries align behind these objectives and use local systems.
3. **Harmonisation:** Donor countries coordinate, simplify procedures and share information to avoid duplication.
4. **Delivering results:** Developing countries and donors shift focus to development results and results get measured. Aid is focused on real and measurable impact on development.
5. **Mutual accountability:** Donors and partners are accountable for development results.
- 6 **Capacity development** - to build the ability of countries to manage their own future.
- 7 **Inclusive partnerships** - All partners - including donors in the OECD Development Assistance Committee and developing countries, as well as other donors, foundations and civil society - participate fully.

This Evaluation Framework must be used by donors working with government partners to build on and strengthen their programme planning and evaluation processes.

9.7 Optimising limited capacity

There is a challenge about limited capacity in government and externally around evaluation. In order to address this:

- Sufficient technical capacity will be established in **DPME** to support departments on methodology and quality;
- Evaluations can be **outsourced** to external evaluations using an accredited panel;
- Donor funds should use **international evaluators** to build capacity and expand the cadre of experienced evaluators beyond a narrow pool;
- **Short courses** are needed from PALAMA, universities, and private consultants. Proactive work will be undertaken to ensure courses address the type of skills needed to address this framework. In addition funds are being sourced to assist with capacity development, in particular to build a cadre of experienced black evaluators, as universities often depend on a limited range of experienced and older white researchers, which can be problematic in securing buy-in by senior managers for implementation;
- **International partnerships** are being built with similar countries (eg Mexico and Colombia), and international organisations, eg 3ie or World Bank.

10 Management and coordination of evaluation across government

10.1 Evaluation Working Group

In order for this policy framework to be implemented it needs a strong champion (DPME) but also needs broad buy-in across government. It also needs to build on the scarce skills which exist around evaluation in government. An Evaluation Working Group will be established to support DPME in taking forward evaluation nationally. This will include the main departments with evaluation capacity, and the OPSC will be invited to participate, because of its strong capacity in this regard.

This group of experienced evaluation professionals will work with DPME to take forward evaluation and meet on a regular basis to review policy documents, technical guidelines, and support the process to take forward evaluation in government.

10.2 Implementation of these evaluation systems across government

The policy framework requires a major upscaling of the use of evaluations, which will have to be addressed in phases. A suggested timeline is:

	Policy and systems	Implementation
2011/2012	• Evaluation Policy Framework adopted by Cabinet • 3 year and annual evaluation plan developed • Practice notes developed on key elements including TORs, Contracting, and the different evaluation types • Evaluation unit created in DPME • PALAMA courses designed to support this approach to evaluation • Panel of evaluators created in DPME • Support agreed with international partners • Minimum standards agreed by Cabinet for programme and project plans, as the key vehicle for implementation	• 4 evaluations commissioned which test out these systems • Audit of all evaluations in the public sector from 2006 • All evaluations hosted on DPME website • Evaluation Working Group starts operation • Capacity development process around evaluation designed

	Policy and systems	Implementation
2012/ 2013	• Systems revised based on experience. • Discussions with university courses to take on this approach to evaluation	• 10 evaluations undertaken or started using standard procedures, of which at least 2 are impact evaluations, and at least 60% of recommendations implemented • Training of 200+ people using PALAMA materials • University M&E courses use adapted materials
2013/ 2014	• Systems revised based on experience • All departments have a budget line for evaluation, and at least 2% of significant programmes have budgets allocated for evaluation	• 20 evaluations undertaken or started using standard procedures, of which at least 4 are impact evaluations, and at least 70% of recommendations implemented • Training of 500+ people using PALAMA materials • All university public administration courses use adapted materials
2014/ 2015	• Systems revised based on experience	• 30 evaluations undertaken or started using standard procedures, of which at least 5 are impact evaluations, and at least 75% of recommendations implemented • Training of 1000+ people using PALAMA materials • Other university courses use adapted materials (eg development studies)

Annexes

Annex 1 References and further sources of information

- Ajam, T (2011): 'Evaluation and Performance Auditing in South Africa; Convergence of reforms?', paper at the 4th International Conference on leadership, Governance and Policy in the Public Sector', School for Public Management and Administration, University of Pretoria 17-18 February 2011, Pretoria
- Barbie, E. & Mouton, J. (2007): 'The Practice of Social Research', Oxford, Oxford University Press.
- DBE (): 'Draft Guidelines for Monitoring and Evaluation of Policy Implementation', Pretoria, Department of Basic Education.
- Kotvojs, F and B Shrimpton (2007): 'Evaluation Journal of Australasia, Vol. 7, No. 1, 2007, pp. 27–35, Australian Evaluation Society.
- OECD (): 'Glossary of key Terms in Evaluation and Results-Based Management', Paris, OECD.
- Presidency (2007): 'Policy Framework for the Government-wide Monitoring and Evaluation System', Pretoria, The Presidency, Republic of South Africa.
- EU (1997): 'Evaluating EU Expenditure Programmes: A Guide', Brussels, European Union.
- Kusek and R Rist (2004): 'Ten Steps to a Results-Based Monitoring and Evaluation System', Washington, World Bank.
- Mayne, J (2001): 'Addressing attribution through contribution analysis: using performance measures sensibly', *The Canadian Journal of Program Evaluation*, vol. 16, no. 1, pp. 1-24, quoted in Kotvoys et al (2007).
- MFA (2007): 'Finland Guidelines for Evaluation', Helsinki, Finnish Ministry of Foreign Affairs.
- OPSC (): 'Basic Concepts in M&E', Pretoria, Office of the Public Service Commission.
- Posavac, E.J. & Carey, R.G. (2007): 'Programme Evaluation Methods and Case Studies', Seventh Edition. New Jersey: Prentice Hall.
- Rossi, P.H, Lipsey, M.W. & Freeman H.E. (2004): 'Evaluation, A Systematic Approach', Seventh Edition. Thousand Oaks: Sage Publications.
- Treasury (2007): 'Framework for Managing Programme Performance Information', Pretoria, National Treasury.
- UK (2011): 'Magenta Book', London, Treasury.

Annex 2: Evaluation Principles

Factor	Description
1 Evaluation should be development-oriented and should address key development priorities of government and of citizens	
Government Priorities	The selection criteria for evaluation projects should take into consideration government priority areas.
Pro-poor orientation	Poverty's causes, effects and dynamics are highlighted and the interests of poor people are prioritized above those of more advantaged groups.
Service delivery and performance	Variables affecting institutional performance and service delivery are analysed and reviewed, links are identified and responsive strategies are formulated.
2 Evaluation should be undertaken ethically and with integrity	
Confidentiality	- Processes ensure the responsible use of personal and sensitive information. - Promises of anonymity and non-identifiability are honoured and relied upon.
Respect	- Dignity and self-esteem is built amongst stakeholders and affected people. - There is skilful and sensitive implementation of evaluation processes.
Representation of competence	Those engaged in monitoring and evaluation fairly represent their competence and the limitations of their reports.
Fair reporting	Reporting provides a fair and balanced account of the findings.
3 Evaluation should be utilisation-oriented	
Defining and meeting expectations	- Evaluation products meet knowledge and strategic needs. - A record of recommendations is maintained and their implementation followed up.
Supporting utilisation	An accessible central repository of evaluation reports and indicators is maintained.
Use of evidence	Promote use of evidence and results-based policy-making, planning, and implementation.
4 Evaluation methods should be sound	
Indicators and measures	Common indicators and data collection methods are used where possible to improve data quality and allow trend analysis.
Data/evidence based	Findings are clearly based on systematic evidence and analysis.
Appropriateness	Methodology matches the questions being asked.
Triangulated	- Multiple sources are used to build more credible findings. - Use a diversity of methods as appropriate, giving weight to both outcomes and processes and not focusing only on one eg quantitative for impacts.
5 Advance government's transparency and accountability	
Availability of findings and recommendations	All findings are publicly available, including in departmental websites
Accountability	Use of resources is open to public scrutiny
6 Inclusion and participation	
Representation of interests	Traditionally excluded interests are represented in the evaluation process. Evaluation needs to ensure that issues of diversity and the differential effects of policy of different socio-economic groups are considered.
Citizens/ Beneficiary participation	The views of the citizens/ beneficiaries are given appropriate weight in an evaluation.
Cooperative governance	An effective evaluation system requires cooperative governance across the three spheres of government.
7 Learning	
Organisational learning	- Evaluations should be learning-oriented to contribute to improved performance. - Knowledge and an appetite for learning are nurtured in institutions and individuals.

Annex 3: Annotated glossary

Adapted from Rossi et al (date) and EU (2007).

Control group	Is the target population that do not receive the programme intervention and that is compared on outcomes measures with one or more groups that do receive the intervention.
Cost-benefit-analysis	An analytical procedure of determining the economic efficiency of a programme, expressed as a relationship between costs and outputs, usually measured in monetary terms
Cost-effectiveness analysis	Analytical procedure for determining the efficacy of a programme in achieving given intervention results in relation to the programme costs. Combines monetary and non-monetary measures.
Counterfactual	What would have happened if an intervention did not exist or if some other policy initiative was implemented.
Efficiency assessment	An evaluation method that answers questions about programme costs in comparison to either the monetary value of its benefits or its effectiveness in terms of the changes brought about in the social conditions is addresses. Cost-benefit and cost-effectiveness analysis can be combined with non-economic methods.
Evaluability assessment	Negotiation and investigation undertaken jointly by the evaluator, the evaluation sponsor, and possibly other stakeholders to determine whether a programme meets the preconditions for evaluation and, if so, how the evaluation should be designed to ensure maximum utility.
Ex-ante evaluation	Ex ante evaluation is a process that supports the preparation of proposals for interventions. Its purpose is to gather information and carry out analyses that help to define objectives, to ensure that these objectives can be met, that the instruments used are cost-effective and that reliable later evaluation will be possible.
Ex-post (or post-hoc evaluation)	Assessing/evaluating quality after a programme or institution has been in operation in order to establish strengths and weaknesses.
Expert panels/opinions	A data collection technique, similar to a survey, which relies on the necessarily subjective views of experts in a particular field. It is not recommended to rely on expert opinion as a sole data source, for example, because of problems with so called "chatty bias".
Experimental design	An impact evaluation design in which intervention and control groups are assigned randomly.
Formative evaluation	Evaluation activities undertaken to assist learning and provide information that will guide programme improvement, especially in terms of how, why, and under what conditions a policy will work or has worked.
Intervention group	Is a target population that receive an intervention and whose outcome measures are compared with those of one or more control or more groups.
Meta-analysis	Determining the overall balance of evidence from different sources and studies (usually expressed in non-quantitative, narrative form) (see statistical meta-analysis).

Multi-criteria analysis	Multi-criteria analysis allows us to formulate judgements on the basis of multiple criteria, which may not have a common scaling and which may differ in relative importance.
Pre-post design	A reflexive control design in which only one measure is taken before and after the intervention.
Programme theory (usually linked to the theory of change below)	The set of assumptions about the manner in which programme relates to the social benefits it is expected to produce and the strategy and tactics the programme has adopted to achieve its objectives. Sub-categories of programme theory include design theory/logic, process theory or impact theory.
Public Expenditure Tracking Surveys	Assesses whether resources reached the intended beneficiaries and whether they result in better services. It can be combined with a quality service delivery survey.
Quasi-experiments	An impact evaluation design in which intervention and control groups are formed by a procedure other than random assignments.
Research synthesis	A way of establishing what is already known about a policy initiative, especially its achieved impact and its implementation challenges in other policy environments.
SMART assessment	This is criteria normally used to assess targets and indicators in terms of whether they are specific, measurable, achievable, relevant and time-bound.
Statistical meta-analysis	The aggregation and generation of cumulative statistical estimates of impact from combining the results of different comparable studies.
Summative evaluation	Evaluation activities undertaken to render a summary judgement on the impact of the programme's performance, eg specific goals and objectives were met.
SWOT analysis	Analysis of strengths, weaknesses, opportunities, and threats related to programme design and implementation. Strengths and weaknesses related to the internal aspects of the program or institutions and opportunities and threats are more external to the programme or institution.
Theory of change	A tool that describes a process of planned change, from the assumptions that guide its design to the to the long-term goals it seeks to achieve.

Annex 4: Approaches to evaluation

Empowerment Evaluation

A participatory/collaborative evaluation involving the evaluator and the intended beneficiaries of the intervention. The evaluator performs the role of the facilitator, coach and trainer in the evaluation process with an aim of fostering improvement, learning, transformation and emancipation amongst the targeted beneficiaries. The assumption is that after the evaluation, the beneficiaries will be equipped with skills to improve their families and communities.

Utilisation-focused evaluation

This approach places emphasis on conducting an evaluation that will be useful. The utility/ultimate use of the evaluation therefore influences decisions on all aspects of the evaluation process, from the beginning to the end. The evaluator involves the target beneficiaries throughout the process, including in decision-making on the use of evaluation.

Naturalistic or Qualitative Model

The purpose of the qualitative model is to obtain a rich and deep understanding of the programme through observation and conversation with the clients and stakeholders. The programme is studied in its natural setting and the insider perspective of the clients is seen as crucial.

Experimental and quasi-experimental

These methods compare the observed changes in a group of people, institutions or states of nature that were exposed to the programme with a similar group that was not exposed. In an experimental design (or randomised control trial approach, RCT) subjects are randomly assigned to different groups at the start of the experiment. In quasi-experimental designs the evaluator would select a comparison group of subjects that are similar to the experimental group, except that they are not part of the programme being evaluated. Experimental and quasi-experimental designs provide high accuracy with regards to the causality between programme and result variables. However they cannot tell programme managers much about why the programme worked or did not work. They may therefore be suited for accountability and reporting purposes, but will not assist much in evaluations that are undertaken to improve programme performance.

Contribution analysis

When experimental or quasi-experimental designs cannot be undertaken on account of cost, the nature of the evaluation object or the evaluation question, the evaluation can use contribution analysis. Rather than attempt to definitively link a programme's contribution to desired results, contribution analysis alternatively seeks to provide plausible evidence that can reduce uncertainty regarding the 'difference' a program is making to observed outcomes (Mayne 2001). Contribution analysis does not work with proof of the counterfactual (ie what would have happened if an intervention did not exist) but aims at providing a 'performance story' to have sufficient certainty that programme activities contributed to observed outcomes. Contribution analysis requires that the existing evidence on results is compiled and examined and that alternative explanations for the results are compiled and explicitly considered.

Theory-based evaluation

Theory-based evaluations examine carefully the 'theory' of how a programme is supposed to work by developing a theory of change model. Theory-based evaluations are based on a careful description of the services to be offered to specified clients, ways the services are expected to change the clients and listing of the outcomes the intervention is expected to

affect (Posavac and Carey, 2007:28). Data is then collected to compare the supposed mechanisms of change with actual events.

Annex 5: Set of methods that can be used in different types of evaluations

Type	Possible data collection methods	Possible analytical methods
Diagnosis	Secondary data, information audits, reviews, evaluability assessments, rapid appraisals, expert panels, focus groups, group methods, surveys, mini-surveys, structured and unstructured interviews	Rapid evidence assessments, document analysis, analysis of secondary data, modelling and simulation, expert panels, meta-evaluation/analysis, case study analysis, problem trees, statistical analysis, content analysis and other qualitative methods.
Synthesis	Secondary data, programme information reviews, literature reviews, meta-evaluation, expert panels.	Rapid evidence assessment, meta-analysis, document analysis, other policy and qualitative analysis tools, statistical analysis, indicator analysis, comparative and benchmarking exercises.
Design evaluation	Secondary data notably programme design documents	Logic model analysis, consistency analysis and indicator analysis
Implementation evaluation	Monitoring data, resource information (budgets, HR and other resources), indicator assessment, inspections, investigations, public expenditure tracking surveys, multi-criteria assessments, interviews, surveys, expert panels, rapid appraisals	Process analysis, program logic chain analysis, cost-benefit analysis, cost-effectiveness analysis, bench-marking, case study analysis, resource analysis, statistical analysis, other qualitative methods, efficiency and effectiveness analysis, retrospective impact analysis.
Impact evaluation	Randomised pre-test and post-test evaluations, experiments, quasi-experiments, ex post comparison or project and non-equivalent control group, rapid assessment ex post impact evaluations, case studies, and other qualitative and participatory methods. RCTs	Advanced statistical analysis, cost benefit analysis, cost-effectiveness analysis, case study analysis, other qualitative methods, meta analysis, consistency analysis, content analysis and econometric analysis tools.

Annex 5: Related guidance notes/guidelines to be developed

Model terms of reference

Model contracts

Process for selection of service providers

Quality standards for evaluations