



Department
for International
Development



International Development Evaluation Policy

May 2013



Contents

Preface	ii
Chapter 1: Introduction.....	1
• The changing context of development assistance	1
• The contribution of evaluation	2
• The purpose of the policy	3
Chapter 2: Guiding Concepts and Principles.....	5
• Concept of evaluation definition	5
• Principles	6
• Evaluation in fragile and conflict affected environments	7
• Complementary tools of enquiry	7
Chapter 3: Requirements and Standards.....	9
• Framework for evaluation	9
• Planning, design and implementation	9
• Types of evaluation	11
• Approaches to evaluation	12
• Quality assurance	14
Chapter 4: Working in Partnership	15
• A more inclusive approach to partnership working	15
• A stronger role for partner countries	15
• Partnerships with others	16
Chapter 5: Using Evaluations	17
• Closing the information loop	17
Chapter 6: Evaluation Policy Implementation within DFID	19
• Senior management	19
• Operational units	19
• Evaluation cadre	19
• Evaluation department	20

Preface

This document sets out the UK Department for International Development's (DFID) policy for the evaluation of its Official Development Assistance (ODA). It comes at a critical time as the external scrutiny of UK assistance, how it is delivered and what the money is spent on, is increasing. The public, as taxpayers, and the media take a strong interest in the aid budget and its impact, and thus it is vital that DFID has a clear, transparent framework for the independent evaluation of its work.

This policy reflects assistance provided primarily through DFID, but may also include assistance provided through other UK Government Departments such as the Department of Energy & Climate Change, the Department of Environment, Food & Rural Affairs and the Foreign and Commonwealth Office. It succeeds DFID's 2009 evaluation policy and reflects the transformation in the breadth and depth of evaluation over the past four years. This policy therefore enables consistency in the application of principles and standards in evaluation across DFID.

The intended audience of this policy is those responsible for the delivery of UK assistance. This includes officials and staff of the UK Government, partner Governments, the private sector and non-Governmental agencies. While evaluation may be the preserve of a smaller group of individuals within these institutions, this policy is intended to illustrate its importance to all those involved in delivering UK assistance.

The evaluation of UK assistance has changed radically over the past two years. In 2011, the Independent Commission for Aid Impact (ICAI) was established, reporting directly to Parliament on the results and value for money of all ODA. ICAI sets a new benchmark for independence and transparency in the scrutiny of all ODA. This has been matched by the embedding of evaluation across DFID to ensure that lessons are learned during the development process and resources shifted to where they are most effective. Evaluation is now considered at the design phase of all DFID's programmes and evaluations themselves are being designed and delivered to a higher standard.

This policy provides the framework under which all DFID evaluation work will be carried out and sets the standard for all evaluation of UK assistance. In some cases these high standards are already being met; in other cases more work needs to be done. DFID is committed to improvement and to support those with whom it works to ensure quality.

Progress will be captured in the Annual Evaluation Report. The effectiveness of the policy, and the direction it spells out, will be assessed annually through reviews of the coverage, quality and utility of DFID financed evaluations and will be subject to a formal review after five years of implementation. I am committed to these changes to deliver better evaluation and encourage you to act on this policy to ensure that we continue to improve.

Mark Lowcock
Permanent Secretary, UK Department for International Development

Chapter 1: Introduction

This chapter provides an overview of the context within which DFID's evaluation work is conducted, its relevance and contribution. The purpose of the policy is also outlined.

The changing context of development assistance

Global changes

1. To be valuable, evaluation must be shaped by, and respond to, change. The global landscape is increasingly affected by population growth, urbanisation, resource discoveries and scarcities, climate change and governance issues. These provide opportunities, for example where there are high working-age populations and few dependents, or economic growth, but also challenges. Climate effects are likely to be harshest in developing countries and disproportionately affecting the most vulnerable, including the poor, women and girls.

Role of the state

2. The role of the state is changing in many contexts with less focus on direct state delivery and more on enabling delivery with, and through, others. In Asia and sub-Saharan Africa, expressions of popular demands have ranged from opposition to authoritarian leadership to informal citizen networks that are monitoring elections, budgets, service delivery and state violence. However, while there is often strong citizen demand for improved governance, it remains a major challenge to convert this demand into real improvements in the quality of public institutions.

Fragile states

3. The majority of countries in which DFID assistance is focused are fragile and conflict-affected states with complex governance problems. The underlying causes of conflict and fragility need to be addressed by strengthening bonds and engagement across the state-society divide and by implementing peace-building and state-building priorities. Making the transition from conflict and insecurity to peace and prosperity hinges on governance reforms that transform weak institutions into inclusive agents that promote growth and are sustainable. Increasing the size of DFID programmes to support these countries and tackle the causes of instability is challenging. DFID is prepared to take calculated risks to better understand what works, how it works, and to measure progress on the ground.

Changing roles

4. The aid system is becoming more complex with new actors and new rules. Emerging donors and philanthropic agencies are gaining in significance and their influence on partner countries is increasing. At the same time, Official Development Assistance (ODA), as a percentage of financial flows to developing countries, is decreasing. Funding is increasingly derived from domestic and international resource flows with increased oversight by respective parliaments and other bodies.

5. These changes require new approaches, relationships and ways of working based on new models of global development cooperation. DFID assistance will need to become increasingly flexible and dynamic to adapt appropriately while retaining the core values of transparency and value for money.

UK direction

6. The UK Government has set out a clear direction for international development through its policies and reviews including improving the effectiveness and efficiency of the multilateral system. DFID assistance is increasingly focused on support for issues that are transformational, including an emphasis on girls and women, the shift to empowering poor people to drive their own development and the power of the private sector to drive growth. The UK is working with new partners and taking different approaches to tackling global public goods, including climate change, and to how it responds to humanitarian and emergency situations. This increasingly involves working more closely across UK Government Departments and with partners across the globe.

Evaluation's role in this context

7. This policy recognises that development is complex and non-linear and thus evaluation must be designed to reflect this in terms of its ambition, design and application. In many of the states in which DFID works, events are fluid and influenced by multiple internal and external factors. Contexts and programmes are likely to be subject to considerable change and it is essential to regularly test and question the assumptions behind the theories for change.

8. In terms of delivery, more than half of DFID spending is channelled through multilateral and civil society partners. This means that DFID requires a clear line of sight on whether these channels are delivering the intended results or not, and evaluation provides one way of determining this.

9. It is therefore crucial that DFID both recognises and responds to these contextual influences if its assistance is to remain relevant and effective. This is why evaluation will increasingly be embedded in DFID programmes and promoted in its development partnerships.

The contribution of evaluation

10. DFID's embedding evaluation agenda aims to ensure that the Department is world class in using evidence to drive value for money and impact and to influence other partners to do the same; that programme design is driven by rigorous evaluation of what works, allowing DFID to test, innovate and scale up; and that DFID takes measured risks using high quality evidence of impact on poverty. This includes developing staff capability to do more, and better quality, evaluations.

11. It is important to bear in mind that evaluation is a significant part (but only one part) of DFID's strategy for using evidence more effectively in order to ensure the Department is achieving the maximum impact from its development assistance. It overlaps with development research and complements the strengthened focus on development results across the organisation.

What evaluation contributes

12. When well done, evaluation can serve a number of functions. Evaluation has a key role in generating evidence and learning about what is working in development and what is not; it can identify better ways of doing things; allow for course corrections of programmes to improve effectiveness; ensure that lessons are learned during the development process and resources shifted to where they are most effective; and improve the ability to respond to change.

Chapter 1: Introduction

13. Evaluation also provides evidence to the UK taxpayer, and to the citizens of the countries in which assistance is being delivered, of how their contribution to international development has been utilised. Done well and transparently, evaluation can help in promoting greater accountability as well as helping to strengthen the relationship between the citizen and the state.

14. In addition, evaluation contributes a unique knowledge of which courses of action can make a difference to people's lives. It differs from other forms of inquiry by its steadfast attachment to understanding whether change has occurred as a consequence of a particular course of action, the nature and implications of the change and whether there have been any unintended effects.

How it can be used

15. UK assistance, and hence evaluation, is a contribution to national and regional development in the developing world. At scale, evaluation of development assistance is primarily about whether aid affects or leverages overall development impacts; whether this is through evaluating the impact of a successful pilot intervention which will be scaled-up by national stakeholders, or through evaluating a contribution of financial and non-financial support to a development process. In the context where aid dependency is reducing, and with the contingent rise of domestic and foreign investment, evaluation will also be required to shift focus and capture these changing influences if it is to remain relevant.

The purpose of the policy

Purpose

16. This policy sets out DFID's principles, standards and approach to evaluation of international development. The policy aims to:

- raise awareness of the importance of evaluating development spending to determine whether the desired effects are being realised
- provide clarity and consistency in the design, conduct and use of evaluation of all UK assistance, irrespective of origin, funding modality, partner or implementing agency
- set high standards, rigour and ethical practice in evaluation to ensure its quality and utility.

What it is not

17. The policy is not intended to be a detailed guide on procedures and practices for evaluation of international development. These are covered in DFID's Evaluation Handbook which reflects this policy and the drive towards better evaluation. Another source of guidance is The Magenta Book which is HM Treasury guidance on evaluation for Central Government, but will also be useful for all policy makers, including in local government, charities and the voluntary sectors.

Chapter 1: Introduction

Embedding evaluation

18. The UK Government has made results, value for money and evaluation central to its approach to tackling poverty and promoting growth. This commitment has led to fundamental and transformational shifts in the evaluation architecture for international development programmes.

19. Since the start of 2011, DFID has been transforming its business processes to ensure evaluation considerations are embedded at all stages of the policy and programme cycle. This initiative contributes directly to better understanding of the impact of DFID's work and will assist in the development of a strong evidence base for decision making and lesson learning.

20. This represents a significant step change. DFID aims to ensure that credible evaluation arrangements are in place for the more strategically important, resource intensive and/or most innovative programmes; that a wide range of new and high quality evaluations are commissioned with partners each year, including impact evaluations; and that a culture is forged within DFID, and amongst its partners, where rigorous evaluation is a routine and accepted part of the policy and programme cycle. This is an ambitious approach which includes promoting change and building capacity of partners, as well as within DFID.

Independent scrutiny

21. Independent scrutiny is vital in determining what results are being achieved. The Independent Commission for Aid Impact (ICAI) was launched in 2011 and reports directly to the International Development Select Committee in Parliament. ICAI has been established to scrutinise the impact and value for money of UK ODA, and provides a greater level of independent, external scrutiny than ever before of ODA managed across Government Departments.

22. ICAI operates with the following core values:

- Independence: undertaking its work without fear or favour and reporting the facts as found
- Professional rigour: using the highest professional standards to gather and evaluate evidence
- Transparency: publishing all its reports and information on ICAI's costs and activities on its website
- Responsiveness: taking account of public and Parliamentary opinion in determining its work programme and in undertaking that work
- Innovation: making the most of ICAI's status to experiment with new ways of working, reporting and interacting with its stakeholders
- Integrity: ensuring that ICAI's operations are characterised by value for money, high ethical standards, transparency and accountability to Parliament and to the public.

23. Evaluation is just one element of the overall accountability and performance management framework for UK ODA. Independent scrutiny is vital for determining what results are being achieved and, in addition to ICAI, comes from the UK's National Audit Office and from Parliament, in particular the International Development Committee.

Chapter 2: Guiding Concepts and Principles

This chapter outlines how DFID defines evaluation of development assistance, how evaluation relates to other forms of investigation and the principles on which evaluation practice are based.

Concept of evaluation definition

24. Evaluation is not one single tool but a body of approaches to assessing change, its causes and impact. DFID subscribes to the Organisation for Economic Co-operation and Development (OECD) Development Assistance Committee's (DAC) internationally agreed definition of evaluation as a '*systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results in relation to specified evaluation criteria*'.

25. In this definition, 'systematic' implies following a pre-determined, replicable and standards-driven process; while 'objective' means avoiding intentional or unintentional bias in the selection and study of the subject.

26. For an evaluation to be so classified within DFID, it must fulfil the three conditions of independence (those conducting the evaluation must be objective and not connected with the intervention under study), transparency (results must be publicly available) and robust methodology (such that, if replicated, similar results would be obtained).

27. Evaluation focuses, in particular, on whether planned changes have taken place, how changes have impacted, or not, on different groups of people and investigates the theory behind the change. The theory of change includes hypotheses and assumptions, the processes, contextual factors and causality to understand better the results, or the lack thereof.

28. An evaluation should provide evidence-based information that is credible, reliable and useful, enabling the timely incorporation of findings, recommendations and lessons into the decision-making processes of the primary stakeholders. It should service the need to account for resources provided and to understand what difference they have made.

Relationship with the subject

29. The other important characteristic of an evaluation is its relationship with the subject of the evaluation. This relationship is determined primarily by the intention of the evaluation process and product. There are broadly three categories of relationship:

- **External and independent:** this is when the independence and autonomy of the evaluation is at its highest level and the evaluation is designed to serve primarily an external accountability function. The involvement of those being evaluated will usually be restricted to responding to information requests and to the findings and recommendations of the report.
- **Internal and independent:** this is when evaluations are contracted out to evaluation consultants but where the evaluations are financed and managed by the same institution that manages the initiative. Independent quality assurance mechanisms are used to ensure that the design, approach and methods

Chapter 2: Guiding Concepts and Principles

used are appropriate and have been applied correctly. The emphasis here is on the quality of the evaluations and thus their independence and credibility are primarily a factor of the design.

- **Internal and contingent:** this is when the evaluation is designed as a reflexive process, focusing on learning through the evaluation, and so there is a strong emphasis on the participation of the subject and other stakeholders in its design, implementation and use. This type of evaluation is intentionally non-independent, and the credibility is determined through the choice of design, approach and methods and as verified experientially by those participating in the exercise.

30. DFID recognises the value in all three of these evaluation categories: the first two as independent evaluation and the third as self-evaluation. This policy focuses on independent evaluation.

Principles

31. The evaluation of DFID assistance is guided by the core principles of independence, transparency, quality, utility and ethics.

Independence

32. High quality evaluation depends on evidence that is objective, trustworthy and credible. Evaluations should be carried out with oversight from specialists who are independent of those responsible for the design and implementation of the development intervention being evaluated. In contrast, a monitoring review might be undertaken by people involved in the design, implementation or management of a policy or programme. If full independence is not possible, then steps should be taken to maximise impartiality (for example, through setting up strong governance arrangements or asking another donor to carry out the evaluation). Governance arrangements for management and quality assurance of the evaluation are key to safeguarding independence.

Transparency

33. Transparency requires that an evaluation is undertaken with a view to sharing lessons learned publicly and enabling accountability, for example, through follow-up of recommendations. This means publication of data, process and findings of evaluations. Allowing for any legal obligations and the need to protect confidentiality, data should be available, accessible and published on a website after an evaluation has been completed. It should be possible to revisit and repeat analysis of primary data that has informed evaluation findings.

Quality

34. Quality pertains to personnel, process and product in evaluation. DFID adheres to the OECD DAC's international quality standards for evaluation. Evaluation strategies and plans should be clear, costed and actionable. Evaluation designs, approaches and methods should reflect the best available in the profession, drawing on the latest research and focusing on what will give the most reliable and useful answers to the evaluation questions. Professionalism and intellectual integrity of the evaluators is crucial. Evaluation findings and recommendations should reflect the needs of the users in form and content.

Utility

35. Utility requires that the evaluation process and product is designed with 'use' in mind. This requires ownership by the relevant users in the selection of the evaluation topic, involvement in designing the evaluation questions and establishing a clear mechanism for responding to the findings and recommendations.

Chapter 2: Guiding Concepts and Principles

DFID evaluations are expected to be designed to be of use to all those affected directly and indirectly by the findings and recommendations.

36. Dimensions of utility include:

- **Timeliness:** ensuring that evaluation reports are produced at relevant points in the policy and programme cycle to inform decision-making
- **Relevance:** ensuring that evaluations ask questions to which the users are keen to know the answers
- **Quality:** ensuring that the evaluation is credible and reliable and thus can legitimately be used for decision making.

Ethics

37. Evaluators and evaluations must adhere to DFID's ethics principles for research and evaluation. It is the responsibility of the evaluators undertaking a study to identify the need for and seek the necessary ethics approval prior to a study commencing. Approval should normally be obtained from both the institution to which the evaluators are affiliated and the relevant local and national authorities of where the study will take place. Where there are no appropriate national or local ethics committees in the study country, the ethics standards of the institution to which the evaluators are affiliated may suffice and, in these instances, further guidance should be sought from DFID.

38. The planning and design phases of evaluations need to consider what possible ethical challenges the evaluation may raise and seek to address these issues. Evaluators should avoid harm to participants in studies and their participation should be voluntary and free from external pressure. There should be assurances of the confidentiality of information, privacy and anonymity of participants.

39. Evaluators need to take account of differences in culture, local behaviour and norms, religious beliefs and practices, sexual orientation, gender roles, disability, age and ethnicity and other social differences such as class when planning evaluations and communicating findings.

Evaluation in fragile and conflict affected environments

40. The same principles apply across all contexts within which UK assistance is provided. However, working in fragile states requires particular emphasis on building transparency and partnership in the evaluation process given the additional sensitivity of the environment within which assistance is provided. This requires a particular focus on seeking beneficiary views in deciding what should be evaluated, and when, and consulting on what questions should be asked. It ensures that evaluations are designed so that the views of under-represented groups are considered properly and particularly so that issues relevant to women and girls are considered at every stage. It also requires that ethical approaches to evaluation are applied rigorously and results are shared. DFID adheres to the OECD DAC guidance 'Evaluating Peacebuilding Activities in Settings of Conflict and Fragility'.

Complementary tools of enquiry

41. Evaluation falls into a spectrum of tools for enquiry along a range from audit at one end to research at the other. This spectrum reflects intent, from internal control of systems and processes to learning about what

Chapter 2: Guiding Concepts and Principles

works, what does not work, and the reasons. The most common forms of complementary tools include audit, monitoring and review.

Audit

42. Audit is an objective assurance activity to systematically review the policies, procedures and operations established to ensure achievement of objectives; assessment and management of risk; reliability of internal and external information; reporting and accountability processes; safeguarding of assets and efficient and economic use of resources; compliance with applicable laws and regulations, behavioural and ethical standards.

Value-for-money studies

43. Within audit, a branch focusing on value-for-money is perhaps the closest cousin to some forms of evaluation, focusing on a systematic and objective examination of activities and spending. The scope includes the examination of economy, efficiency, cost-effectiveness and environmental effects of activities; procedures to measure effectiveness; accountability relationships; protection of public assets; and compliance with authorities. The scope and procedure for value-for-money audits are typically prescribed in detail, making them narrower in focus than other forms of evaluation, and less focused on outcomes.

Monitoring

44. Monitoring is a continuous process of systematic data collection to inform managers and key stakeholders on progress in relation to planned inputs, activities and results, as well as the use of allocated resources. Monitoring is structured around indicators, which are the measures of performance of the input, activity or results (output or outcome). Indicator targets provide the benchmarks against which progress is monitored. Monitoring takes place to enable managers to rapidly identify problems and make necessary corrections to ensure proper implementation. Information from systematic monitoring serves as a critical input to evaluation but does not look in-depth at the causes behind why results are being delivered or not.

Review

45. Review is a periodic or ad hoc assessment of performance or progress of a policy, sector, institution, programme or project. Unlike evaluation, which assesses the results (outcomes and impacts) of initiatives, reviews tend to emphasise operational aspects and are therefore closely linked to the monitoring function.

46. A systematic review is a literature review focused on a research question that tries to identify, appraise, select and synthesise all high quality research evidence relevant to that question.

Research

47. Evaluation has increasingly turned to research methods and practices to help define and assure what it delivers. Research, in its fundamental form, aims at expanding knowledge rather than for direct utility. Research may generate additional questions that require further research to lead to public or commercial utility. Applied research is focused on solving practical problems some of which may be part of an initiative that is being evaluated. In this sense, research designs are used in evaluations to address specific issues and questions.

Chapter 3: Requirements and Standards

This chapter addresses DFID's requirements for its evaluations and the standards upon which these will be assessed. As noted above, DFID follows the OECD DAC international quality standards for evaluation.

Framework for evaluation

48. The commissioning of an evaluation should be guided by a clear identification of need and a strategy that provides the justification for what is to be evaluated and why. Relevant systematic reviews of the available evidence should be considered prior to commissioning evaluations to ensure that the required knowledge is not already available.

Coverage

49. Key decisions about what to evaluate, and how, is reflected in evaluation strategies developed by the appropriate DFID operational unit levels, including all country offices, regional programmes, policy and spending departments. Evaluation strategies should demonstrate how they plan to address gaps in the evidence base, the demand from key stakeholders, the evaluation of risk and innovation where it is evident in the programme portfolio, testing pilots for potential replicability or up-scaling and ensure that the coverage is proportionate to the size, spend and scale of the results to be delivered. Not all projects and programmes can be evaluated so the strategies are an important tool for prioritisation.

Resources

50. A significant and appropriate level of resource should be allocated specifically to evaluation. The allocation of resources should be based on the evaluation strategy and the type and scope of evaluations planned. Obtaining value-for-money implies that evaluation strategies should be proportionate to the portfolio of initiatives that they cover.

Planning, design and implementation

51. The planning and design of an evaluation plays a fundamental role in its likely quality and utility. Considerable investment should be made in terms of time and effort to ensure that the design addresses the needs of those who will use the findings, and that the evaluation has the highest probability of being able to deliver answers to the evaluation questions posed. It is essential that evaluation is considered at the design stage of development interventions so that planning and budgeting are appropriate. A communications plan should also be developed early in the planning process.

Relevance and use

52. Evaluation should be designed to meet the information and decision-making needs of the intended users and other stakeholders. The evaluation must illustrate how it has addressed issues of consistency with the country and/or organisation's overall development strategy. The design must also identify how the evaluation plans to be relevant to national stakeholders, either through focusing on whether the intervention leverages broader development impacts or how the evaluation will cover programmes that include public or private investment beyond DFID assistance. A clear mechanism should be specified for keeping the evaluation on track, engaging the users throughout the study to keep them informed and to ensure that there are no surprises as the best way of ensuring uptake of the findings.

Chapter 3: Requirements and Standards

53. The consideration of evaluation should determine the most essential questions and appropriate methods and look carefully at the existing evidence base and data availability before commissioning expensive collection exercises. Key evaluation criteria help to define specific evaluation questions which in turn guide the selection of the evaluation design, its approach and the methods to use for gathering and analysing data.

Criteria

54. For DFID, this is framed around the five OECD DAC criteria of relevance, effectiveness, efficiency, sustainability and impact:

- Relevance is the extent to which the objectives of a development initiative are consistent with the needs of those it intends to benefit
- Effectiveness is the extent to which the planned results were achieved, or are expected to be achieved
- Efficiency is the measure of how economically resources/inputs (funds, expertise, time) are converted into results
- Sustainability is the actual or probable continuation of benefits from an initiative after major development assistance has been completed
- Impact is positive and negative, primary and secondary effects produced by an initiative, directly or indirectly, intended or unintended.

55. Other criteria have since been added to address the humanitarian context, including coherence, coordination and coverage, all of which can also be applied more generally.

56. This policy does not prescribe which criteria must be applied in all and every case, given the varying objectives of evaluation, but does expect each study to define which criteria were chosen, and why, as well as the methods used to gather and analyse data.

Design

57. Evaluation design is the overarching framework which provides the fundamental basis for how the evaluation will be conducted. The design should include a well-defined and consistently stated rationale, objectives, scope and context of the evaluation. It should have focused evaluation questions related to specific evaluation criteria and which are clearly linked to the purpose of the evaluation.

58. Critical in the formulation of the design are an understanding of the context, its degree of shifting influence over the intervention and details on the nature of the intervention, its complexity and trajectory. The narrower the evaluation questions, such as in the case of impact evaluation, the more careful the evaluator needs to be to understand the likely trajectory and possible influences. The more complex the underlying logic, the more important it will be to account for other factors that might affect the outcome. Hence, the selection of the type of evaluation is based not only on the kinds of questions posed but also on the programme attributes and context. The timing of the evaluation is also important given the speed of change that can occur in these environments.

Approach

59. Supporting the design are the evaluation approaches which provide more detailed frameworks for analysis. The choice of approach depends on the particular evaluation questions; any one evaluation may use different approaches depending on the variety of questions.

Methods and tools

60. The methods are the procedures to be used for collecting and analysing data. Robust methodologies must be outlined which enable plausible judgments to be made. Explicit and due attention should be given to the issues of impartiality, propriety, inclusion and ethics.

61. The methods selected should provide evidence that the study will lead to clear evaluation findings, evidence-based conclusions, lessons and recommendations that can be implemented; with realistic expectations given time and resources that will be allocated; identifies challenges and risks, and makes appropriate proposals for dealing with these.

62. Methods are supported by specific instruments or tools such as surveys and focus groups. Data sources should be clearly established and allow for multiple lines of enquiry and diversity. Techniques for data collection should be appropriate to the data type and allow for disaggregation.

63. Designing the evaluation at the same time as the programme will help inform choices on what type of data can be collected to produce a good baseline, explore methods to use and, if appropriate, allow for the creation of a control or comparison group. Where this is not desirable or feasible, focused participatory monitoring can be used to compile regular, specific information both on the context and the programme which can later be used to inform a theory-based evaluation design.

Management and governance

64. The management arrangements should be clearly spelled out and be appropriate to the focus and scope of the evaluation. Similarly, the requirements for the evaluation team should be commensurate to the scope of the evaluation and should ideally have a mix of disciplines, gender and nationalities.

65. The governance and oversight arrangements for the evaluation must be clearly specified with a view to promoting quality and independence. Ideally, there should be a management team for the evaluation and an independent reference group to promote transparency and impartiality and to provide advice on content and technical quality issues.

Implementation and procurement

66. The evaluation should be undertaken by evaluators who are not connected with the development intervention. This independence does not imply isolation and it is envisaged that evaluators and commissioners will interact throughout the study.

67. Evaluators will be selected in an open and transparent process. The same evaluators may be used for both the design work and implementation of the evaluation but this is not essential. It is the responsibility of the contracted evaluators to ensure the integrity of the evaluation, including that DFID's ethics principles are upheld. In DFID, all evaluation Terms of Reference must be independently quality assured before proceeding to procurement.

Types of evaluation

68. A variety of types of evaluation are possible depending on the requirements of those commissioning the evaluation. The OECD DAC identifies a number of different types of evaluation which can be clustered into four dimensions relating to the **unit of analysis** (policy, country programme, meta, theme, cluster, sector,

Chapter 3: Requirements and Standards

project), **timing** (formative, mid-term, summative and ex-post), **approach** (impact, participatory, process) and **relationship to the subject** (independent, self-evaluation).

69. DFID recognises all of these dimensions of evaluation but places a premium on independent evaluation rather than self-evaluation.

70. Evaluations at the formative and mid-stages of an initiative (whether policy, programme or project) are used to guide or course-correct; summative and ex-post at the end, or after, the initiative are used to determine what was achieved and why. The selection of when to evaluate should also be guided by an understanding of the likely impact trajectory of the initiative. Rarely is impact linear, particularly when targeting empowerment, improved participation of particular social groups, awareness raising or such like. The theory of change guiding the initiative, and instructing the evaluation, should therefore be clear on when particular impacts are expected in order to guide the timing and focus of the evaluation.

71. Formative evaluations can help in situations where there is uncertainty about the direction and scope of an initiative caused either by changes in the environment within which the initiative is operating, delays in the period between design and implementation of the initiative, evidential weaknesses in the original design or changes in the implementing partners. The evaluation design and methods used may vary. Formative evaluations that seek to illicit a stronger interpretation of the initiative from the beneficiaries should use appropriate participatory methods to discuss, gather and analyse information.

72. Summative evaluations can be useful in large and complex programmes, including portfolios and country programmes, where there are several strands and changes that have taken place during the lifecycle. These evaluations should emphasise a strong theory-based design, focusing on the context within which particular decisions were made and courses of action taken, outlining assumptions and intended results. Revealing these can provide key findings and recommendations around programme and organisation change and lead to useful restructuring and re-programming.

73. DFID distinguishes approaches from the methods used to implement them. It does not therefore consider participation and process as approaches as such in independent evaluation, but as methods used to address particular questions.

Approaches to evaluation

74. The choice of evaluation approach is principally determined by four factors: the context within which the initiative is operating, the attributes of the programme, the evaluation questions posed and the criteria being used to judge these components. It is expected that a robust evaluation will utilise more than one approach in its design, reflective of the different questions being posed. However, it is also acknowledged that the context and programme attributes, which remain constant across the questions, will have a determining factor on the choice of approaches. DFID commonly uses theory and impact-based approaches, which are detailed below.

Theory-based approaches

75. Theory-based approaches to evaluation focus on understanding the characteristics of development interventions, answering the question of what it is about a particular approach that works, rather than whether or not it works at the aggregate level. This kind of approach is suitable for a complex initiative where there

are multiple assumptions and lines of inquiry, and also potentially in fragile or conflict environments where wide scale data collection may not always be feasible.

76. Theory-based approaches can use a variety of methods implying the collection and use of qualitative and quantitative data. However, the strong emphasis on theoretical and contextual research typically requires more use of qualitative data.

Impact approaches

77. Impact evaluation approaches focus on establishing cause and effect. They aim to demonstrate, through a recognised design, approach and method, that one or more impacts (positive or negative) have been caused by an intervention, the size of these effects and the reasons why they have taken place.

78. These approaches are best suited to single lines of inquiry, where the intended cause and effect is hypothesised in advance. In complex programmes, strong analysis of the theory of change is required to unpick hypothesised cause-and-effect chains as the basis for impact evaluation.

79. Impact evaluations can be conducted for pilot initiatives where it is important to identify as early as possible, based on an understanding of the impact trajectory, whether it appears to be working or not and hence should be replicated or scaled up, or not. They are also useful and feasible where a programme has a phased roll-out or where there is expected over-subscription of beneficiaries (and hence not all will receive the programme) as an opportunity is presented to establish a comparison group of non-beneficiaries.

DFID-funded impact evaluations of development assistance

80. DFID-funded impact evaluations use an appropriate design which is able to measure impact credibly given the type and context of the intervention. This must include a counterfactual or similar analysis to address attribution and establish causality. It should be noted that where a randomised trial or use of a control group is not appropriate or feasible, alternative approaches can and should be used.

81. The prospective approach taken by DFID should be used so that impact measurement and data requirements are built in from the outset. The design should allow for a disaggregation of impacts by different groups, by gender, wealth and such like, as appropriate. The design should also capture information on the context which affects impact; this is important to draw conclusions about using the intervention in other settings.

82. The best available methods should be used which will stand up to scrutiny and ensure that findings are valid, replicable and minimise bias. Specialist advice from impact evaluation experts is essential at all stages on design and methods.

83. Qualitative methods and analysis are essential for understanding how and why impact is achieved (or not) and how impact is affected by the context of the intervention. This implies that qualitative data collection and analysis approaches are required in all cases.

84. Experimental methods are uniquely capable of establishing causality with a high level of statistical confidence, dealing with selection bias and providing rigorous estimates of the size of impact. However, they cannot always be used. Typically, they are best suited for well-defined interventions where bespoke survey data can feasibly be collected or where high quality administrative data are available.

Chapter 3: Requirements and Standards

85. Quasi-experimental and statistical methods make efficient use of existing household surveys and administrative datasets. They can also use commissioned survey data and are well suited for measuring impact for programmes where a comparison has to be established retrospectively or where randomisation is not possible.

86. Appropriate and statistically valid use of data should be made, including surveys at individual, household or community level, to provide primary data on impact wherever possible. Secondary data may be used to provide a basis for statistical matching.

87. Design, methods and analysis should be open and transparent. The evaluation should be subject to peer review against quality assurance standards appropriate to its design at various stages, particularly in the design phase and when publishing findings. Where appropriate, the design protocol should be published in advance and results published in a peer reviewed journal. Results should also be fed back to beneficiaries in an appropriate and timely manner. Data, and necessary documentation, should be made publicly available in a suitable format so that other researchers can replicate and test the findings.

Quality assurance

88. DFID has introduced mandatory quality assurance requirements to promote and deliver high quality evaluations. In this context, quality assurance is a formal review against a predetermined set of standards and is distinct from support or advice to those planning an evaluation. The aim of quality assurance is to ensure that a high standard of design, methodology, management, product quality and use of evaluation findings is achieved consistently across DFID.

89. It is mandatory for all evaluations to be independently quality assured during the design phase (entry) and at the draft final report (exit) stages. While quality assurance can be undertaken by a number of different groups, the process itself is centrally managed. Evaluation specialists within DFID are expected to help bring the product up to, at least, minimum standards before formal quality assurance takes place. Joint evaluations with partners are not required to submit to DFID's own quality assurance process if the partner institution has a comparable robust quality assurance process in place for both entry and exit stages.

90. In addition to the quality assurance requirements, there must be a formal management response to all findings, conclusions and recommendations from each evaluation. This response must be prepared promptly after the evaluation is concluded and published alongside the report to which it refers.

91. All evaluation products should be made available on the DFID external website and all impact evaluations should be formally registered. Data sets from primary data collection, clear of personal identifiers, must be made available for re-analysis. This requirement is both a public good and a point of transparency.

Chapter 4: Working in Partnership

This chapter outlines DFID's commitment to working with a range of partners and improving beneficiary engagement. For real change to occur, the demand for evaluation, its practice and use must be generated by those whom DFID is aiming to support. These are both the citizens and end-users of the assistance and the intermediary agencies of governments, private sector organisations, local civil society actors, bilateral and multilateral institutions with whom DFID works directly.

A more inclusive approach to partnership working

92. In line with its commitments under the Paris Declaration, Accra and Busan Agreements, DFID will continue to look to partners to take the lead in evaluations where appropriate and will work with others to promote high standards of quality and independence.

93. DFID will support partners, where necessary, to achieve these standards whilst recognising that it takes time to develop skills and capacities, including the establishment of sustainable institutions able to undertake quality evaluations. This implies a stronger role for developing countries including their governments, parliaments and civil societies.

94. The Department will also seek to improve its evaluation work with international partners, including multilateral institutions, global and regional funds and partnerships and international civil society organisations. It will continue to work with the UK Evaluation Society and other national and international evaluation bodies to develop capacity and contribute to emerging thinking on standards and methodologies in the field of evaluation.

95. DFID will also work closely with other UK Government Departments on evaluations which affect all of the UK's development work.

A stronger role for partner countries

Fostering demand

96. Evaluation demand and supply capacity are central to the strengthening of evaluation use in partner countries. Evaluation demand will be addressed through implementing the principles of utility, transparency and quality ensuring that the selection of what, and how, to evaluate is done as much as possible with development partners. Improving the use of the evaluation systems and evaluators from the partner countries will be achieved through encouraging international evaluation firms and individual partners to work with local firms and individuals in the countries where evaluations are taking place. This will increase knowledge and exposure to methods and practices and will increase the demand for use. Beyond this 'learning-by-doing' strategy, DFID is committed to directly building the capacity of evaluation systems and evaluators, addressing the 'demand-side' environment, the institutions who coordinate and manage evaluations, and the individual evaluators themselves. All three – environment, institutions and individuals – need to be addressed if evaluation culture and practice is to strengthen.

Chapter 4: Working in Partnership

Supporting supply

97. DFID works at international, regional and national level to strengthen capacity. It will continue to work with, and through, international evaluation associations, regional associations and with international bodies which set standards and provide guidance that is helping to professionalise the evaluation industry.

98. Capacity development relates to supporting systems and processes and therefore does not lend itself to short-term, project-type assistance. Along with other partners, DFID supports regional initiatives based at existing academic or research institutions around the globe which enables a particular focus on strengthening the performance measurement and evaluation capacity of partner country governments and agencies. In an effort to improve specific skills sets and practical experience of evaluators in impact evaluation, DFID is also providing support through specialist and multilateral agencies that run training courses for national partners and provide technical assistance and financing to conduct impact evaluations on a demand-led basis. In related areas of social science research, the UK Government has extensive programmes of support to national research institutions across Africa and, to a lesser extent, Asia and is committed to improving science and technology capacity over the medium and long-term.

Partnerships with others

99. Currently, DFID channels more than half of its programme expenditure through multilateral agencies, joint funds and civil society organisations at international and at national level. As much as possible, and to promote efficiency, DFID uses the evaluation systems of partner organisations and accesses the evidence they generate. DFID supports this relationship by working with partners to ensure that their systems meet international standards for independence, credibility and utility by, for example, working with professional evaluation bodies such as the DAC Network on Development Evaluation, the United National Evaluation Group and the Evaluation Coordination Group of the Development Banks.

Multilaterals

100. The multilateral system is a crucial complement to what DFID can do alone. To assist with driving transformation in key multilateral organisations DFID is striving to secure improvements in the effectiveness of their evaluation functions, focussing on continuous improvement, transparency and greater system-wide working. Where high-level strategic evaluations of multilateral institutions or joint funds are required, DFID aims to undertake them jointly with the relevant multilateral, or Global Fund, and other partners to maximise ownership and influence, share costs, improve quality and minimise reporting burdens. DFID also seeks to evaluate its role in working through multilateral organisations and is developing an evaluation strategy to guide the work with multilateral institutions.

Civil society organisations

101. DFID channels significant funding to civil society organisations through bilateral country programmes and centrally managed funds. DFID is committed to a partnership approach with civil society with respect to evaluation and evidence generation while recognising that not all organisations are in a position to conduct complex studies, including impact evaluations. DFID is therefore committed to engaging with civil society organisations to jointly identify ways to ensure active participation in evaluation activities including the promotion of evaluation quality, utilisation, and learning.

Chapter 5: Using Evaluations

This chapter outlines how evaluations commissioned by DFID will be used internally and by partners. DFID is committed to promoting and improving the use of evaluation both through the process itself and in the application of evaluation findings.

Closing the information loop

Use

102. Well-designed evaluations have a strong chance of leading to changes in individual thinking, organisational procedures, programme and policy design. The use of findings can be categorised in three ways:

- **Instrumental use:** this is where someone has used evaluation knowledge directly (for example, to inform an on-going programme or to design a follow-on phase)
- **Conceptual use:** this is when no direct action has been taken but where people's understanding has been affected (for example, as part of a body of evidence to guide an understanding of what works)
- **Symbolic use:** this is when a person uses the mere existence of the evaluation, rather than any aspect of its results, to persuade or to convince.

103. DFID also recognises that evaluation can influence and affect by indirect and intangible means. Potential indirect effects, both positive and negative, should be considered during the evaluation design.

Stakeholder involvement

104. Stakeholder involvement is critical to facilitating those aspects of an evaluation's process or setting that leads to greater use. Engagement, interaction and communication between the potential users of an evaluation and the evaluators are central to maximising the use of the evaluation in the long run. Effective stakeholder involvement, from the outset, is also critical to improve the chances that the findings will be accepted and acted upon.

Communications plan

105. The timing of an evaluation is critical to its utility. If the findings are produced too late, they may not be able to be used to influence decision or policy making; if done too early, there may not be sufficient data to ensure robust findings.

106. The reporting requirements should be identified at the design stage and a communications plan produced. An evaluation should be designed with a specific primary use in mind, such as, whether to influence a specific policy process, the design of a new or phased initiative, to determine whether a pilot should be scaled-up, or to provide accountability for the use of resources. If an intervention is being designed with a clear break point after an initial phase with a decision point as to whether to continue or not, then this provides a clear opportunity for evaluation to inform this process and fulfil its accountability function.

107. In line with both the UK Government's Freedom of Information Act and the International Aid Transparency Initiative, all DFID-funded evaluations will be made publically available and actively disseminated. This should include publication and ensuring open access to underlying data sets where

Chapter 5: Using Evaluations

appropriate. A proactive approach to creating appropriate fora for discussions on findings, which will feed into policy and funding decisions, is encouraged to provide transparency and promote uptake and follow-up. Learning will be shared for wider benefit encouraging citizens of developing countries to use evidence, when appropriate, for accountability purposes.

108. All evaluation reports should have a limited number of well-focused conclusions and recommendations that flow logically from the analysis and findings. The recommendations must be realistic, relevant and targeted to the intended users and to specific audiences where possible. They should be logically derived from the evidence and presented in a clear and concise form, laid out in priority order and making clear the essential and obvious recommendations as distinct from those that are more difficult and/or controversial. Where possible, multiple options for courses of action should be proposed.

109. As recommended by the House of Commons Science and Technology Committee, all evaluation reports should include a contact address for raising any formal concerns regarding the study.

110. Evaluations must be accessible to the intended users. Ensuring accessibility requires that the reporting products are readable for those who are colour-blind, low-vision, or reliant on an audio reader. Written, visual and verbal products that represent the initiative and its evaluation should be produced in accordance with the communication plan. Graphic design and data visualisation can be applied to emphasise key pieces of content and increase primary intended user engagement.

Management response

111. A formal management response is required for all DFID evaluations. The form of management response should set out agreement or disagreement with each key finding, conclusion and recommendation; state the reasons why and, where agreed, express what action will be taken to address the recommendation. Where the evaluation is targeting one or more particular groups, each is required to respond to the findings, conclusions and recommendations in a timely manner.

112. DFID is committed to ensuring that management responses are completed, published and followed-up through the annual Directors' Annual Statements of Assurance exercise.

Knowledge management

113. DFID will work with its partners to encourage greater openness and transparency of evaluation findings and data, including better use of evaluation databases. DFID will support the creation of shared databases to register impact evaluations and to capture and improve access to primary data generated through evaluation. This will enable data to be available for re-analysis and for longer-term evaluation time-frames, as well as improving access to baseline data.

114. DFID's Evaluation Department (EvD) will produce Annual Evaluation Reports to provide an overview of DFID's evaluation work. The report will focus on evaluation learning and evidence, evaluation results and evaluation capability.

115. The findings of single evaluations are often not applicable more widely. Evaluations need to be subjected to meta-analysis and/or systematic review to enable broader lessons to be identified and shared. DFID is committed to conducting systematic reviews and meta-evaluations drawing on its own evaluations and other sources of evaluation and research.

Chapter 6: Evaluation Policy Implementation within DFID

This chapter sets out the different roles and responsibilities for evaluation within DFID which is the UK Government Department disbursing the highest percentage of ODA.

Senior management

Leadership

116. DFID has established an institutional mechanism to provide strong leadership to keep evaluation high on the organisation's agenda and strengthen incentives for staff to ensure that evaluations are prioritised. This mechanism includes senior management Committees which meet regularly to discuss evidence generated by the organisation and to consider the overall shape and pattern of evaluation practice across DFID. These Committees also follow up on the utilisation of evaluation results within the Department. It is the responsibility of these senior Committees to ensure that adequate resources and capacity are in place and that robust evaluation activity is being undertaken.

The Independent Commission for Aid Impact and DFID

117. A DFID Senior Civil Servant is appointed to act as the lead contact for each report produced by ICAI. This lead person is responsible for ensuring that DFID provides information directly to ICAI in an efficient manner; check facts and information security in draft reports within agreed timescales; provides clear and timely management responses to ICAI; and implements agreed recommendations effectively.

118. Where an ICAI report covers more than one Government Department, the DFID lead person will normally act as the UK Government's coordinator during the review and provide the joint UK management response. ICAI, however, will engage directly with other Government Departments on specific aspects of the reviews.

Operational units

119. DFID's operational units include all central and country based offices and departments that design and manage programmes and policies. These units' responsibilities for evaluation have grown substantially since the decision was taken to embed evaluation more firmly within DFID's work once EvD stopped commissioning evaluations.

120. Operational units across DFID are now responsible for determining what should be evaluated, allocating funds, procuring, overseeing and responding to the evaluations. They are accountable for using the formal quality assurance mechanisms established to ensure product relevance and quality.

Evaluation cadre

121. Evaluation is recognised in DFID as a professional discipline and a cadre of evaluation specialists has been established. Evaluation is also one of the shared technical competences for all DFID Advisers. A technical competency framework has been defined which outlines the knowledge, skills and experience expected of

Chapter 6: Evaluation Policy Implementation within DFID

specialist evaluators and those involved in commissioning, managing and using evaluations. This framework is used for external recruitment and internal assessment and promotion.

122. Evaluation Adviser posts have been created and embedded throughout DFID to provide technical advice and support in developing evaluation strategies, commissioning and procuring evaluations and to ensure that evaluations meet required quality standards. In addition, they help non-specialists to develop their awareness and skills of using and commissioning evaluations as well as working with external partners to build capacity and promote the use of evaluation.

Evaluation department

123. EvD sets DFID's evaluation standards and is the principal source of advice and capacity building for evaluation practice across the organisation. One of EvD's roles is to provide a strategic focus for DFID evaluation to ensure alignment with key policy and expenditure priorities and addressing key evidence gaps. EvD guides and supports policy, regional and country programme teams to embed evaluation into the Department's operations. It does this by providing specialist technical assistance in the design and implementation of evaluations across DFID. EvD also promotes the uptake and use of evaluation evidence in key thematic areas by acting as a focal point for synergies and learning across the organisation.

124. Another role for EvD is to build evaluation skills and capability to commission and deliver relevant and high quality evaluations. It does this through this policy, maintaining guidance and training on evaluation, establishing standards and implementing quality assurance mechanisms, providing a conduit for sharing information on evaluation, and direct development of evaluation skills and capacity within the organisation.

125. EvD also works to strengthen evaluation capacity of DFID's partners and contributes to global leadership through engagement with a wide range of international partners in evaluation.

The Department for International Development:
leading the UK Government's fight against world poverty.

Department for International Development
22 Whitehall
London
SW1A 2EG
UK

Tel: +44 (0)20 7023 0000

and at:

Abercrombie House
Eaglesham Road, East Kilbride
Glasgow
G75 8EA
UK

Tel: +44 (0)2355 84 4000

Website: www.gov.uk/government/organisations/department-for-international-development

Facebook: www.facebook.com/ukdfid

Email: enquiry@dfid.gov.uk

Public enquiry point: 0845 3004100 or +44 1355 84 3132 (if you are calling from abroad)

© Crown copyright 2013

Copyright in the typographical arrangement and design rests with the Crown. This publication (excluding the logo) may be reproduced free of charge in any format or medium, provided that it is reproduced accurately and not used in a misleading context. The material must be acknowledged as Crown copyright with the title and source of the publication specified.

Published by the Department for International Development, May 2013