



**Australian Government**

**Department of Foreign Affairs and Trade**

# **DFAT MONITORING AND EVALUATION STANDARDS**

April 2017

# CONTENTS

INTRODUCTION TO THE MONITORING AND EVALUATION (M&E) STANDARDS	2
STANDARD 1 INVESTMENT DESIGN (REQUIRED ELEMENTS RELATING TO M&E)	6
STANDARD 2 INVESTMENT MONITORING AND EVALUATION SYSTEMS	11
STANDARD 3 INVESTMENT PROGRESS REPORTING	18
STANDARD 4 TERMS OF REFERENCE (TOR) FOR INDEPENDENT EVALUATIONS	23
STANDARD 5 INDEPENDENT EVALUATION PLANS	29
STANDARD 6 INDEPENDENT EVALUATION REPORTS	34
STANDARD 7 MONITORING VISITS	39

# INTRODUCTION TO THE MONITORING AND EVALUATION (M&E) STANDARDS

## WHY DO WE NEED STANDARDS?



### Key messages

High quality monitoring and evaluation products are required to ensure that information generated from investments has a credible basis and is suitable for use to make important programming decisions and for wider learning. In turn, this is expected to lead to improved accountability, and a more effective aid program delivered efficiently.

Monitoring and evaluation products in these Standards refer to:


- Investment design documents
- Investment monitoring and evaluation systems (M&E plans and M&E products generated by the investment management teams)
- Investment progress reports
- Independent evaluation terms of reference
- Independent evaluation plans
- Independent evaluation reports
- Monitoring visit plans and reports

*Note: In line with DFAT guidance and terminology, this document uses the word 'investment' to refer to a commitment of resources and activity of work to achieve defined outcomes. The term investment is generally synonymous with a 'project' but does have a particular meaning in the context of DFAT's aid management system. The term 'program' defines strategies for thematic areas, countries, regions and/or global programs.*

During recent years, DFAT has been reviewing and exploring factors which account for the quality of these products. The quality of monitoring and evaluation products has been found to be highly variable which has important implications for the basis on which programming decisions are made. Some of the factors that have been identified that explain this situation are that key actors in M&E do not share the same understanding of basic concepts of M&E, or have a consistent view of what represents a quality M&E product. This lack of agreement has been observed across DFAT and within the development and evaluation industry more broadly.

DFAT personnel need guidance on: how to articulate their requirements fairly consistently to M&E Practitioners and the industry more broadly; how to assess the quality of the M&E products they receive; assurance that M&E products do meet a reasonable standard of quality; and how to work with implementation teams and M&E practitioners to improve the quality of products where necessary.

Equally, the suppliers of M&E products benefit from this clear articulation of what is required, and the Standards provide a strong basis for the negotiation of the delivery and resourcing of quality products. It also improves the efficiency of monitoring and evaluation processes as both the demand (DFAT) and supply sides



(M&E provider and implementing partners) have a clear idea of the expectations before embarking on M&E tasks.

These Standards have been created in a way that allows M&E Practitioners the flexibility to use whatever evaluation theory, methodology or methods they consider appropriate for the task and in consideration of their individual skill sets. They are not meant to represent a particular theoretical perspective, but at the same time they do reflect current preferences in the development sector such as the inclusion of a results-based management orientation.

## WHAT IS THE BASIS OF THESE STANDARDS?

These Standards were developed and evolved over time by Evaluation Capacity Building Programs operating in the Asia-Pacific Region. This is an institutional change program designed to improve the quality and use of M&E products, and to integrate evaluative thinking into everyday work. The Standards were integrated into DFAT's Department-wide evaluation guidance in 2012 and provide a strong tool for the articulation of DFAT's expectations of the quality expected from a range of M&E products. They are updated annually each November in response to feedback from the full range of users. These changes not only address technical content, but also clarity in the language and explanations of concepts suitable for application in a wide range of country contexts. These Standards were also subjected to a formal peer review in September 2012.

The Standards are based on a careful assessment of DFAT staff roles and functions, and an in-depth assessment of the barriers and facilitating factors that have led to current quality in evaluation practice. The Standards reflect:

- the tasks carried out by the different levels of DFAT staff in the purchasing and oversight of M&E systems and deliverables;
- the tasks carried out by implementation partners and M&E practitioners who design, implement and produce M&E systems and products; and
- the capacity of DFAT staff to absorb new information and perform new tasks within their already busy work demands.

## HOW RIGIDLY SHOULD THE STANDARDS BE APPLIED?

DFAT recognises that change will be incremental and take time, and that different monitoring and evaluation systems may require some flexibility. Time will be required for DFAT personnel to reach a good degree of consistency on how they articulate and provide feedback on requirements, and implementation partners will need time to address supply side issues relating to the provision of competent Monitoring and Evaluation practitioners able to meet the Standards.

Flexibility is required as demand and supply issues are progressively addressed. **These Standards are not meant to represent a checklist of rules that must be adhered to.** They are meant to provide guidance for demand and supply actors as to what a quality product ought to look like. This document attempts to provide guidance on where flexibility is appropriate and how to handle this without compromising on rigour. The two key messages around flexibility are a) the degree of rigour should be proportionate to the importance of the decisions being made; and b) the basis of findings must be transparent.



## WHAT STANDARDS HAVE BEEN DEVELOPED?

The Standards presented in this series are:

Standard 1: Investment Design (required features for M&E)

Standard 2: Investment M&E Systems

Standard 3: Investment Progress Reporting

Standard 4: Terms of Reference for Independent Evaluations

Standard 5: Independent Evaluation Plan

Standard 6: Independent Evaluation Report

Standard 7: Monitoring Visits

## WHAT IS THE DIFFERENCE BETWEEN GUIDELINES AND STANDARDS?

The Standards complement departmental guidance which outlines DFAT's expectations for a given process, by providing an additional resource to assist in the articulation and measurement of quality. They are expected to be shared with M&E Practitioners and implementation partners working with DFAT.

The Standards should be used with the following DFAT Guidelines:

- Aid Programming Guide (2016) – specifically Chapter 3: Aid program management and performance reporting; Chapter 4: Investment management, evaluation and quality reporting; Chapter 5: Investment Design
- DFAT Aid Evaluation Policy (2016).

## TRAINING AND PRACTICAL SUPPORT FOR THE USING THE STANDARDS

These Standards are expected to be delivered within a broad range of support for change. Although the Standards are designed to be a relatively stand-alone document, it is useful for teams to undergo training or be provided with support to apply the Standards appropriately. Such training and support may be provided by Performance and Quality (P&Q) areas and/ or by the Contracting and Aid Management Division (ACD).

The Contracting and Aid Management Division delivers a range of training and workshops in Canberra and at overseas posts, including Managing Design, Monitoring & Evaluation, and Program Logic training.

## A NOTE ON SPECIAL TARGET GROUPS FOR GENDER EQUALITY AND SOCIAL INCLUSION

Throughout the Standards, the terms "special groups" or "special sub-groups" are used to refer to any group that is considered to require tailored interventions or special attention to:

- allow their voices to be heard in the design, implementation and evaluation of the program;
- purposefully reach them, and provide the opportunity to participate as equally in the program as other groups;
- allow them to enjoy program benefits equally with others in that community or population;
- ensure the program addresses any particular needs they may have that may not be shared by the wider population, or may not have been considered in society or by decision-makers.

This could include men, women, boys and girls; people with disabilities; the poor and the near poor; ethnic or religious minorities; indigenous communities; the elderly; the sick or infirm; those with low levels of education; as well as cross-sections of these groups (e.g. women with disabilities).

# STANDARD 1

## INVESTMENT DESIGN (REQUIRED ELEMENTS RELATING TO M&E)

No.	Element
<b>Investment Design Plan</b>	
1.1	Terms of reference reflect the requirements for the M&E Practitioner/Design Team Leader
1.2	M&E Practitioner/Design Team Leader has demonstrated skills in Program Logic/ Theory of Change
1.3	Design plan allows for sufficient opportunity to gather relevant information to inform the design
1.4	Design plan allows for sufficient opportunity to meaningfully engage local partners in the design process
<b>Investment Design Document</b>	
1.5	The goal (beyond the life of the Investment) is aligned with the priority needs of the target population, and is aligned with relevant Program Strategy Objectives and/or SDGs (see IDQS1.a)
1.6	The design document describes a compelling case for how the end-of-program outcomes are expected to contribute to the broader goals (see IDQS1.a)
1.7	The beneficiaries or actors expected to change any behaviours for key outcomes are identified (see IDQS1.c and 3.b)
1.8	The end-of-program outcomes are expressed in terms of performance outcomes where possible rather than capacity, or open-ended outcomes (see IDQS3.a)
1.9	The end-of-program outcomes are pitched at the correct level for the time, effort and resources applied (see IDQS4.c)
1.10	Immediate and intermediate outcomes are described for key interventions where this is reasonable (see IDQS3.a)
1.11	There is a robust logic and theory of change which explains the causal mechanism(s) that are understood to link the interventions with intermediate and end-of-program outcomes (see IDQS3.a)
1.12	The design describes important contextual factors that are likely to influence outcomes (see IDQS2.a)
1.13	The choice of delivery approaches, partnerships and implementation arrangements are appropriate to achieve end-of-program outcomes (see IDQS2.a)
1.14	The design specifies that the design, implementation and products of the M&E system needs to be undertaken with reference to these DFAT Standards 2 and 3 (see IDQS3.d)
1.15	High-level results are articulated (see IDQS3.d)
1.16	The design articulates the importance of partner government mutual accountability by making provision for the mutual assessment of progress where appropriate (see IDQS3.d)
1.17	Adequate provisions have been made for design and oversight of M&E systems for the life of the investment by an adequately qualified M&E Practitioner.

1.18	Important evaluation questions emerging during the design are highlighted
1.19	Adequate resources (personnel and financial) for M&E are included in the design (see IDQS4.c)

## Detailed Description for Standards for Design

**Note:** This standard is a supplement to the Investment Design Quality Standards (IDQS). The IDQS set out DFAT's principal requirements for the content of a completed Investment Design. These M&E Standards build on the IDQS ensuring the design contains sufficient depth to ensure quality M&E during implementation.

### Investment Design Plan

#### *1.1. Terms of Reference reflect the requirements for the M&E Practitioner/Design Team Leader*

Terms of reference for the M&E Practitioner/designer for the design mission adequately articulate the required tasks to meet these Standards and the Investment Design Quality Standards. The terms of reference also indicate the qualifications and experience required of the individual responsible for the development of the program logic. The team leader ought to possess design skills, but can be supported by suitable technical specialists. Terms of reference describe the design approach that will be employed – a facilitated in-house design, a contracted out design, or a hybrid.

#### *1.2. M&E Practitioner/Designer has demonstrated skills in Program Logic/ Theory of Change*

The individual recruited to develop, or facilitate the development of the program logic/theory of change has demonstrated experience in doing so. They can provide evidence of this experience if requested. They need to have the capacity to apply program logic/theory of change in a practical way reflecting the principles of program logic/theory of change, rather than developing intricate models that cannot be easily understood by stakeholders. The ability to deal with simple, complicated and complex designs is required. If an DFAT member of staff has been tasked to design the investment they either can demonstrate the required experience, or are adequately supported by an individual who possesses the required experience.

#### *1.3. Design plan allows sufficient opportunity to gather relevant information to inform the design*

Analyses are carried out in advance of the final design decisions to ensure that the design decisions are based on credible information, that decisions are transparent, and to ensure the design addresses the right development issues and priorities. Adequate time is provided to conduct analyses to inform the design. These analyses could include: a) international definition of key concepts or dimensions of sub-sectoral areas that should be considered; b) international review of factors commonly inhibiting performance in the relevant area; c) analyses of supporting or inhibiting factors encountered in the country; d) important contextual information that may impact on the success of the Investment. Cross-cutting issues, gender equality, disability inclusion, and social inclusion as well as safeguard issues have been addressed (refer to current DFAT policies including DFAT Gender Analysis Approach).

#### *1.4. The design plan allows sufficient opportunity to meaningfully engage local partners in the design process*

The design plan allows sufficient opportunity to meaningfully engage individuals - at the appropriate level of decision-making - in the design process. This is sufficient to maximise national ownership of the final design. The design plan shows the specific mechanism by which the national partner can assume joint decision making and leadership over the design process. For example, through adequate representational work before the final design process is carried out, and regular participation in key decision-making points throughout the design process. It also outlines an effective approach to dialogue during the design process.



Where appropriate, representatives of target special groups are heard in the analysis phase of the design, particularly in terms of identification of their particular needs, or identification of risks and opportunities where different groups gain from program benefits.

## Features in the Design Document

### Program Logic/ Theory of Change

#### *1.5. The goal (beyond the life of the investment) is aligned with the priority needs of the target population*

The goal is aligned with the priority needs of the target population, including any special groups that may be in some way marginalised from the general population or require special interventions. The goal for the investment is directly aligned with DFAT Program Strategy (Delivery Strategy, or agreement with the partner government such as a Statement of Commitment, Partnership for Development, or Record of Understanding), local partner development priorities (including gender equality and social inclusion policies), and/or the Sustainable Development Goals. The goal may be drawn from the level of the outcomes articulated in the sectoral/delivery strategies that support the Program Strategy. The basis for the identification of priority needs is clearly provided.

#### *1.6. The design document describes a compelling case for how the end-of-program outcomes are expected to contribute to the broader goals*

The design clearly states how the end-of-program outcomes are linked to the higher order outcomes (or goals). For example, in an educational program, if the end-of-program outcome is that teachers demonstrate effective, modern teaching practices (against an agreed standard), then the design would argue exactly how improvements in teaching quality are expected to contribute to a higher outcome of increased completion rates for students. The proposed mechanism is fully described and draws on either research or experience in the literature, or another credible basis. This basis must be made explicit.

#### *1.7. The beneficiaries or actors expected to change any behaviours for key outcomes (and related interventions) are identified*

Where appropriate, key actors in the investment that are expected to adopt and sustain new behaviours or practices are clearly identified. The expected changes in behaviour or practices by the last day of the investment (or beyond) are described for each key actor. Beneficiaries and actors are not necessarily treated as a single group. Relevant special sub-groups such as women, the poor or the disabled are identified as having particular behaviour changes if appropriate.

#### *1.8. The end-of-program outcomes are expressed in terms of performance outcomes where possible rather than capacity, or open-ended outcomes*

The end-of-program outcomes are expressed in terms of performance outcomes at either the individual, work group, organisational or institutional level. It is clear what individuals or groups will be doing differently on the last day of the investment (including vis-à-vis women or any special sub-groups). Capacity outcomes are not used (capacity outcomes refer to the potential to change behaviour - clear descriptions of the actual behaviours to be achieved is required).

Open-ended outcome statements such as “improved capacity in human workforce planning” or “improved health information systems”, or “increased resources allocated to X” are not included. The magnitude of the change expected is clearly articulated. If there is a compelling argument for open-ended outcome statements due to a lack of knowledge then: a) this was not due to poor analyses during the design phase; and b) there are clear plans for systematically gathering the required information, and a future date planned for further development of the outcome statements. Outcomes for relevant cross-cutting themes (such as gender and disability) and safeguards (such as child protection, environment, displacement and resettlement) are clearly articulated.

Gender equality and social inclusion outcomes are identified and addressed under the relevant outcomes across the design particularly where they are relevant to technical or process issues in the proposed investment. They are not treated as a separate work program, but fully integrated into the design.

*1.9. The end-of-program outcomes are pitched at the correct level for the time, effort and resources applied*

The end-of-program outcomes can reasonably be expected to be achieved in the time frame, in accordance with the effort exerted, and the degree of resources applied to interventions. There are provisions in the design to allow key actors to first learn, then practice and then apply new skills independently in a supervised environment (consolidate) before the end of the investment.

*1.10. Immediate and intermediate outcomes are described for key interventions where this is reasonable.*

Immediate and intermediate outcomes are described where these outcomes are known or can be predicted with a reasonable degree of certainty. These are interim steps of behaviour change that are pre-requisites to the expected end-of-program outcome. Immediate and intermediate outcomes are provided for the most important interventions, taking into consideration cross-cutting issues. Important interventions are defined as either those that play a key role in bringing about the end-of-program outcome, or account for a significant allocation of effort or resources, or where it is important to ensure the Program Logic/ Theory of Change is fully described. Note: Where a program logic is complicated, complex, or emergent, it may not be feasible to develop intermediate outcomes, or identify all the key interventions expected to trigger the outcome chain or achieve the end-of-program outcomes.

*1.11. There is a robust logic and theory of change which explains the causal mechanism(s) that are understood to link the interventions with intermediate and end-of-program outcomes*

The design explains how the end-of-program outcomes are expected to be achieved. This is based on credible arguments supported by either the literature, or generally accepted experience of professionals with known expertise in the topic area. Where there are competing theories or a lack of consensus on how interventions are expected to bring about the required change, then these are documented and brought into considerations for M&E during the life of the investment.

*1.12. The design describes important contextual factors that are likely to influence outcomes*

Analysis of the important contextual factors likely to support or inhibit achievement of end-of-program outcomes (including outcomes for specific sub-groups) are identified and addressed in the design. Important inhibiting factors have been integrated into the risk monitoring and management section. This requires that Standard 1.3 above has been met.

*1.13. The design provides a clear rationale for how the delivery approaches, partnerships and implementation arrangements are appropriate to achieve end-of-program outcomes*

Choices such as using partner systems, core contributions, specific purpose investments/ funds, pooled funds, projects, technical assistance, scholarships or debt relief may be suitable for the end-of-program outcomes described. Further standards which must be met in regards to the delivery approach and implementation arrangements are outlined in the descriptors for IDQS 4.b.

**Enabling Quality M&E Systems (features carried over into RFT, Contract or MOU)**

*1.14.* The design specifies that the design, implementation and products of the M&E system needs to be undertaken with reference to these DFAT M&E Standards 2 and 3. The design document makes it clear that guidance from the Aid Programming Guide, and M&E standards (such as M&E Standards 2 and 3) are reflected in the requirements for M&E. These Standards reflect international standards such as the OECD DAC and JCSEE Program Evaluation standards.

### 1.15 High-level results are articulated

A small number of high-level results for the investment are identified. These should be at the level of the higher order end-of-program outcomes and are likely to reflect benefits to communities (including any relevant special groups) rather than only changes in systems performance. Where appropriate, these are linked to aggregate development results, any DFAT corporate, thematic or program strategies, or performance assessment framework or to local partner performance assessment frameworks.

### 1.16. *The design articulates the importance of mutual accountability by making provision for the mutual assessment of progress where appropriate*

The design makes provisions for the mutual assessment of progress against the agreed commitments and end-of-program outcomes. Where appropriate, M&E information is generated through partner government and/or local systems, but where this is not available or appropriate, interventions are included that enable local partners to participate in the generation and utilisation of findings and recommendations relevant for their needs. Mutual accountability for an investment is highly desirable and is assumed to be a feature of the M&E system. If, however, this is not considered appropriate then a full discussion of the rationale for this decision is presented.

Note: *If capacity building in M&E is proposed for local partners, then this is integrated into the design as an intervention, and not an add-on to the M&E system or to the TORs for the M&E Practitioner for the investment.*

### 1.17. *Adequate provisions have been made for design and oversight of M&E systems for the life of the investment by an adequately qualified M&E Practitioner.*

Resources have been made available for an M&E Practitioner to design and provide oversight of investment M&E systems for the life of the investment. This requires regular short-term inputs throughout the entire implementation period. There may be special circumstances where a full-time M&E Practitioner is required. A single task of designing an M&E system without on-going technical support is not proposed. Regular visits (at least twice a year, or proportionate to the scale of the investment) have been resourced. For larger, or priority investments, a terms of reference for the M&E Practitioner has been attached to the design document and requested to be a named individual during the tendering process. Qualification and professional experience requirements for this specialist reflect the magnitude or sensitivity of the investment. Requirements should also include at least awareness in relation to gender equality and social inclusion (Note: sample TORs for an M&E Practitioner are available).

### 1.18. *Important evaluation questions emerging during the design are highlighted*

Important evaluation questions that emerged during the design process are highlighted for later inclusion in the final M&E system. These could include questions to guide management and implementation decisions identified by key stakeholders, or questions relating to the mechanisms that link interventions with outcomes that require exploration during implementation. Evaluation questions have clear relevance to improvement of investment implementation, or of interest to a wider sectoral, local partner or DFAT management audience.

### 1.19. *Adequate resources (personnel and financial) for M&E are included in the design*

Adequate budget has been allocated to monitoring and evaluation. For small investments this is likely to be a higher proportion of the total budget than for larger investments. Consider also DFAT-commissioned evaluations, any technical oversight and investment manager monitoring visits. This allocation remains intact through the Request for Tender (RFT) and contracting phase. For pilot or innovative investments, more resources should be available to increase the robustness of M&E for policy implications.

## STANDARD 2

### INVESTMENT MONITORING AND EVALUATION SYSTEMS

No.	Element
<b>Design Phase to Mobilisation</b>	
2.1	Adequate resources are allocated for M&E
2.2	Terms of reference for the M&E Practitioner reflect requirements for a quality system
2.3	Adequately qualified M&E Practitioner is recruited
<b>M&amp;E Plan</b>	
2.4	There is a summary of the investment goals, outcomes, investment size and length and any other relevant information
2.5	There is an adequate basis for the development of the M&E Plan (e.g. Evaluability or Readiness Assessment)
2.6	The M&E Plan provides a summary of the overarching system design including key M&E approaches and activities
2.7	The M&E Plan is consistent with current international standards for evaluation practice (e.g. Joint Committee on Standards for Educational Evaluation – Program Evaluation Standards)
2.8	Goals and End-of-program outcomes are clearly articulated and assessed
2.9	The plan is focused around key performance indicators and evaluation questions linked to specific intended uses of the information
2.10	The reach/coverage, quality, and exposure of participants to key deliverables are monitored and evaluated
2.11	Relevant aspects of the context and key risks are monitored
2.12	Methods are fully described for sampling, data collection, management, analysis and processing
2.13	Baselines are constructed where appropriate
2.14	Responsibility is allocated to specific individuals (not organizations) for all M&E activities
2.15	Mutual accountability and joint assessment by local partners is provided for (using partner systems where appropriate)
2.16	Individuals responsible for implementing the M&E plan have the capacity to do so (time, resources and skills)
2.17	A strategy for the utilization of information is described
2.18	A complete schedule of M&E activities shows when all key M&E activities will be carried out and information available
2.19	The M&E plan can be easily understood by non-specialists and key stakeholders
2.20	The M&E plan is resourced and costed

M&E Products	
2.21	M&E reports or products are available upon request and report against planned M&E activities
2.22	Progress reports meet stakeholder needs, report against M&E plan, have a credible basis for claims, and recommend actions to improve performance
2.23	M&E reports demonstrate how monitoring and evaluation systems have informed learning, decision-making and action

## DETAILED DESCRIPTION FOR INVESTMENT MONITORING AND EVALUATION SYSTEMS

### DESIGN PHASE TO MOBILISATION

#### *2.1. Adequate resources (personnel and financial) are allocated for M&E*

Adequate budget has been allocated to monitoring and evaluation. Budget is required for personnel, and material and financial resources for monitoring and evaluation activities. This budget also includes considerations of DFAT-commissioned evaluations, any technical oversight, and investment manager monitoring visits. For small investments this is likely to be a higher proportion of the overall budget than for larger investments. This allocation remains intact through the RFT and contracting phase or during the negotiation of Agreements for non-commercial implementation partners. For pilot or innovative investments, more resources should be available to increase the robustness of M&E for important policy implications.

#### *2.2. Terms of reference for the M&E Practitioner reflect requirements for a quality system*

The Terms of Reference (TOR) clearly articulate the required tasks including tasks for: M&E system design, oversight of the implementation of the system, and conduct of more advanced M&E tasks relating to outcome evaluation. The TORs also state the expectation that the standard of M&E will reflect these DFAT standards, guidance and international evaluation practice standards such as those described in the [Joint Committee for Standards of Educational Evaluation – Program Evaluation Standards](#). The TORs describe the required standards of qualifications and experience for the position. For investments requiring more robust M&E systems, there is a requirement for experience in the design and conduct of a suitable range of rigorous methods. The TORs provide for an adequate number of days input to cover these M&E needs for the life of the investment.

#### *2.3. An adequately qualified M&E Practitioner has been recruited*

The M&E Practitioner selected has the required qualifications and experience described in the TORs.

### MONITORING AND EVALUATION PLAN

#### *2.4. There is a summary of the investment goals, outcomes, investment size and length, and any other relevant information*

At the beginning of the M&E Plan there is a clear description of the investment. The M&E Plan must be a stand-alone document that makes sense in its entirety. The Plan can be understood by readers who have not been closely involved in the design or delivery of the investment. Important information includes an



overview of what the investment is trying to achieve, general approaches to how this will be achieved (e.g. main interventions and delivery mechanisms), the size of the investment, the start and end times, and any other information that will allow the general reader to grasp the investment itself (the thing that is being monitored and evaluated).

### *2.5. There is an adequate basis for planning the M&E system*

The M&E Practitioner is provided with the time and resources to conduct an Evaluability Assessment (or similar exercise). It typically takes 10 days to conduct this assessment where the program logic does not require major adjustments and is not an extended process. Where the design has not maintained its relevance, or the program logic/theory of change of the design is weak, further time has been allocated. The basis of findings from the EA is made available to DFAT if requested. The Evaluability Assessment report addresses such areas as:

- a. Discussions with key stakeholders to confirm a shared interpretation of the expected long-term and end of program outcomes. Approaches to engage national partners for mutual accountability are explained.
- b. Discussions with key stakeholders to review the program logic and theory of change and describe the extent to which the program can be evaluated including clarity of expression of end-of-program outcomes in the documentation;
- c. Identification of key evaluation questions and information of interest to stakeholders;
- d. An examination of proposed/potential data sources (including partner systems) to ensure that data is of sufficient quality, and is sufficiently disaggregated. In addition, whether it will be collected and analysed as expected, and will be available within the required reporting cycles;
- e. An assessment of the capacity of the implementation team and partners to participate in the design and/or conduct of M&E activities;
- f. A review of the budget/resources available for M&E activities;
- g. Identification of reporting requirements for users of the M&E system-generated information. This includes progress reporting, Quality at Implementation Reporting and Annual Program Performance Reporting. There should be a clear indication of how the M&E plan will provide evidence for reporting against the Performance Assessment Framework and any other higher order results required at a Department level.
- h. A review of cross-cutting policy areas that will need to be included in the M&E plan. Careful assessments of the M&E of gender equality and social inclusion aspects of the investment have been carried out;
- i. Clear identification of issues and/or constraints that will affect the design of the M&E plan.

### *2.6. The M&E Plan provides a summary of the overarching system design including key M&E approaches and activities*

There is a summary of the overall design of the M&E system. This describes the key M&E approaches and activities. The reader should be able to quickly see the broad design features, such as the purpose of the system, the primary users of information, the scope of what will (and will not) be addressed, what the major evaluations are, and a broad description of the associated methods. The M&E plan does not launch into a detailed description of the system without this orientation (i.e. this investigatory framework is similar to the introduction to a methods section in research). Resource implications of the proposed system are summarised.

### *2.7. The M&E plan is consistent with international standards for evaluation*

The M&E plan makes it clear that international standards for M&E will be met. Universally accepted international standards can be found at: <http://www.icsee.org/program-evaluation-standards>, or the OECD-DAC standards are also relevant for the development sector. In particular, consideration is given to how best to balance the requirements of utility, validity, feasibility and ethics (these features are considered as a set of

guiding principles rather than the basis for further quality assessment in addition to these Standards). In addition any guidance in the Aid Programming Guide should be reflected in the plan.

#### *2.8. Goals and End-of-Program outcomes are clearly articulated and assessed*

Early in the M&E plan, goals that are expected to be achieved beyond the life of the investment (the rationale) and end-of-program outcomes are clearly articulated. End-of-program outcomes are pitched at the right level with regard to the resources, effort and time available, and the interventions proposed. These outcomes reflect the intent of the investment and reflect robust program logic. If the Evaluability Assessment shows that end-of-program outcomes are not well articulated in the design, then these have been adjusted to be more appropriate for the investment intent.

End-of-program outcomes related to gender equality and social inclusion are integrated into the relevant areas of M&E system, where they relate to particular technical or process aspects of implementation. They are not treated as a separate M&E system.

Adequacy of progress toward the goal (if quality data can be collected cost effectively) and end-of-program outcomes are assessed through the M&E system.

*Note: For complex or innovative investments, end-of-program outcomes may evolve over the life of the investment. Flexibility is required in these cases. This does not mean that end-of-program outcomes are not articulated, rather that the M&E system is designed to provide important contextual information or analyses that may provide a strong rationale for adjustment of the proposed outcomes. For such investments end-of-program outcomes are reviewed annually.*

#### *2.9. The plan is focused around key performance indicators and evaluation questions linked to specific intended uses of the information*

A limited number of performance indicators are identified that address end-of-program outcomes, and any intermediate outcomes that are important to the underlying program theory (output indicators are addressed under 2.10). Where relevant, indicators for context and risk monitoring and management are identified. Interim and end-of-program targets are set where this is sensible, and the approach to estimating targets is made transparent.

A small number of high level key evaluation questions are posed that relate to specific information needs of stakeholders, and these meet DFAT's information needs in relation to the investment. At a minimum, specific questions address quality of implementation; progress toward end-of-program outcomes (including gender equality and any social inclusion outcomes); the causal mechanisms that link interventions with intended outcomes; important positive or negative unintended outcomes. Methods are identified for each of the evaluation questions. Key evaluation questions should not be phrased as generic program evaluation questions (or indicators phrased as questions), rather very targeted exploratory questions to get a particular issues of concern to stakeholders. The list of questions should be as focussed as possible to allow primary information users to make appropriate decisions about the investment, and future programming in the country program, or inform corporate decisions where this need has been made explicit by DFAT or their national partners. It is readily apparent how information generated from these questions will be utilised.

Some resources are made available to address any emergent questions over the life of the investment.

If only performance indicators are presented in the M&E Plan (these answer the question: did we achieve our expected intermediate or end-of-program outcomes?), consider whether or not these will provide sufficient information for investment improvement (if not, why not – if yes, what were the factors that enabled success, and what were the factors that would be a pre-requisite for scale up or replication), and whether key evaluations questions are required.

#### *2.10. The reach/coverage, quality and exposure of participants to key deliverables are monitored and evaluated*

Key investment deliverables are defined as those accounting for large effort and/or expenditure, those that are likely to have a major influence on the extent to which the investment will meet its expected outcomes, or those that are commonly employed. If a key deliverable is likely to have a major influence on a country or sector, then these are also formally assessed through a limited number of indicators or evaluation questions. The coverage and reach is monitored to allow output reporting on beneficiaries participating in the investment. The adequacy of participation in the investment by target participant groups, including special groups affected by gender inequalities or social exclusion, is assessed. Coverage data is disaggregated according to relevant special groups (with disaggregation by sex necessary under all circumstances). The exposure of particular individuals to the investment deliverables is assessed where this is relevant. The quality of key deliverables is assessed.

#### *2.11 Relevant aspects of the context and key risks are monitored*

Important indicators or evaluation questions about the context are monitored or evaluated. Information is gathered on: the progress of the broader national sector or sub-sector in relation to the investment's area of work; any relevant areas in the context which impact on the relevant needs and priorities of the investment actors (especially gender equality but also any special sub-groups groups); any factors which are impacting on the achievement of sustained outcomes or equitable outcomes for different groups; quality, reach and coverage of the investment deliverables or unintended differential coverage of special sub-groups; or the extent to which the investment may be contributing to changes in the context including any unintended outcomes. Key risks are being monitored and continuing risk assessed. Only highly relevant issues are monitored and this choice represents good value for money.

#### *2.12. Methods are fully described for sampling, data collection, management, analysis and processing*

Proposed methods are fully described, and the degree of rigour is proportionate to the importance of the investment and/or the decisions being made. The evaluation of any pilot activities intended for significant scale-up is treated as rigorous. Major studies are fully designed, or provisions made for the future design of them by the M&E Practitioner. TORs are prepared for major studies that will be designed by others. Implementation teams are not expected to design evaluation studies. Methods are fully described for sampling, data collection, data management, and data analysis or processing. Tools are annexed if appropriate. Where tools have not been developed during the preparation of the M&E Plan, then the expected date of finalisation of tools is identified in the M&E implementation plan and responsibility allocated to a suitably qualified individual. Where tools have been annexed, these adequately address the performance indicators or evaluation questions posed. Any limitations to the methods are highlighted, and advice given on the implications for responding to the findings.

Where national or multi-lateral partner M&E systems are proposed to provide information, then a statement of likely quality and availability is made. This is informed by the Evaluability Assessment (see element 2.5 above).

Consideration is given to the design of data collection methods that are responsive to the needs, rights and security of respondents, with special consideration given to the needs of any special sub-groups. This includes the identification and recruitment of investment participants as respondents, the time of day and location of data collection activities, and the techniques employed during data collection activities. Consideration has been given to the representation of special sub-groups as data collectors, and their preparation to work effectively with these respondents.

*Note: If 2.12 is being formally appraised, it should be done by a suitably qualified M&E Practitioner with methodological expertise*

### *2.13. Baselines are constructed where appropriate*

Baseline data are collected where appropriate. A baseline is not a broad needs assessment, but tightly focussed on the outcomes intended to be affected by the investment. A rationale is provided where a baseline has not been included in the system design.

### *2.14. Responsibility is allocated to specific individuals (not organisations) for all M&E activities*

For all tasks required to implement the M&E system, the responsible individuals are identified. This is not naming organisations, but individuals by position title. For larger studies there will be multiple responsible individuals for design, data collection, analysis, interpretation and reporting. All major sub-tasks should be allocated. Special reference is made to those individuals that will be responsible for the integration of the M&E of gender equality and social inclusion aspects of the investment into the main M&E system.

### *2.15 Mutual accountability and joint assessment by national partners is provided for (using partner systems where appropriate)*

Mutual accountability by local partners is a feature of the M&E system. Partner systems are used where appropriate information is available and of acceptable quality. If alternate systems are proposed, then a suitable rationale is provided. Joint assessments are planned. Mechanisms for engaging local partners for mutual accountability are explained. If this is not considered appropriate, a full description of the rationale is provided. Local partner engagement in the generation and use of information is provided for where appropriate.

### *2.16. Individuals responsible for implementation of the M&E plan have the capacity to do so*

Individuals who are (or will) be recruited to have responsibility for carrying out M&E tasks have the time, resources and skills to do so. The M&E Practitioner should not dedicate significant resources to building implementation team or partner capacity for M&E, rather the system design reflects the resources available. If national partner capacity for M&E is required to be developed, then this is reflected as an *intervention in the design*, rather than as a sub-set of activities in the implementation of the M&E system.

### *2.17 A strategy for the utilisation of information is described*

Information that is generated by the M&E system is expected to be utilised by primary users (and perhaps a wider audience) of the information. There is a clear strategy described that explains: a) how evolving evaluation questions will meet the needs of primary users; b) how information will be presented in a suitable format to primary users; c) what approaches will be employed to ensure information is considered during relevant decision making activities. Consideration is given to providing feedback to representatives of any relevant special groups.

### *2.18. A schedule of M&E activities is presented*

A full schedule (such as a Gantt chart) is provided that indicates when all key M&E activities are to be carried out. Reporting is not described as a frequency (e.g. monthly, annually) rather by date (e.g. the 25<sup>th</sup> of each month; or 25<sup>th</sup> June annually). The schedule makes it explicit when information will be available to primary users and is compatible with national and/or DFAT reporting cycles.

### *2.19. The M&E plan can be easily understood by non-specialist and key stakeholders*

The M&E plan is written so that the primary users can readily understand the M&E system, and can easily see when and how their information needs will be met. The M&E plan should not take more than a couple of hours to fully understand. Orientation sections, detailed annexes, and sign posts to relevant sections to different primary users are all aids used in developing an accessible plan.

#### *2.20. The M&E Plan is resourced and costed*

The resource and cost schedule includes other costs beyond the person days of the M&E Practitioner. These may include person days of other people, personnel associated costs, M&E materials, software and hardware required, any respondent costs, and other resources such as transport.

## **M&E PRODUCTS (SEE ALSO STANDARD 3: INVESTMENT PROGRESS REPORTS)**

#### *2.21. M&E reports or products are available upon request and report against planned M&E activities*

All M&E reports described in the M&E plan are available on request. This may be summary M&E reports to inform progress reports, or reports from specific evaluations or studies. All major evaluations or studies require a separate evaluation report to be available (though not necessarily submitted to DFAT).

#### *2.22. Progress reports meet the needs of primary information users and report against the M&E plan*

Progress reports are available and understandable to primary information users. They meet DFAT's expected information needs. They are not too long, and they are written for managers and non-specialists and present important data in suitable formats to communicate the meaning efficiently and effectively. Progress reports provide information promised in the M&E plan, they include a credible basis for claims made, and they include recommended actions for investment improvement. The latter include expected dates for management responses to be completed. Refer to Standard 3: Progress Reports for a full list of elements required.

#### *2.23 M&E reports demonstrate how monitoring and evaluation systems have informed learning, decision-making and action*

Progressive reports demonstrate how information generated from the M&E system is informing learning, decision-making and action.



## STANDARD 3

### INVESTMENT PROGRESS REPORTING

No.	Element
	Feature of Progress Report
3.1	There is an executive summary that communicates the key information required for routine DFAT reporting and oversight.
3.2	The relevant aspects of the context and any risks are adequately described
3.3	There is a reflection on the continuing relevance of the investment
3.4	An assessment of the adequacy of progress toward sustained end-of-program outcomes is described
3.5	The reach/coverage, quality and exposure of investment participants to relevant key outputs or deliverables for the reporting period are described
3.6	The adequacy of progress implementing the annual plan is described
3.7	An assessment of the likely adequacy of planned inputs to meet the expected end-of-program outcomes is provided
3.8	The adequacy of progress against the budget is assessed
3.9	Key management or implementation systems are described and their performance assessed
3.10	The report provides balanced and fair reporting of positive and negative issues, achievements and challenges
3.11	For claims of achievement or barriers to achievement, credible supportive evidence is provided
3.12	Data or findings are presented in formats that effectively and efficiently communicate important information
3.13	The frequency of reporting is suitable for effective investment management
3.14	The report includes lessons learned from implementation that have potentially important implications more broadly
3.15	Previous and/or proposed management responses or recommendations are summarized

## Detailed Description of Standards for Investment Progress Reporting<sup>1</sup>

### *3.1. There is an executive summary that communicates the key information required for routine DFAT reporting and oversight*

An executive summary provides easy access to information to inform the Quality at Implementation assessment, and meets DFAT's information needs. The main focus for this information is summary information on: the continued relevance of the expected end-of-program outcomes; progress toward achievement of sustained end-of-program outcomes (including any relevant cross-cutting themes, and gender equality and social inclusion outcomes); any differences in outcomes for special sub-groups; quality, reach and coverage of key outputs or deliverables; important factors (including contextual factors) impacting on adequacy of outcomes and outputs; implications of key issues to the achievement of outcomes and outputs, implementation of the annual plan and meeting the budget; and any key management responses proposed. Make special note of any information that provides insight into the achievement or challenges for gender equality and social inclusion aspects of the investment. Any issues with the management or implementation systems are summarised. The DFAT investment manager will explain any updated information requirements for DFAT Program or corporate purposes.

### *3.2. The relevant aspects of the context and any risks are adequately described*

The relevant aspects of the context are described. Information is provided on: the progress of the broader sector or sub-sector in relation to the investment's area of work; any relevant areas in the context which impact on the relevant needs and priorities of the investment actors, especially women and any other relevant special groups; any factors which are impacting on the achievement of sustained outcomes or quality, reach and coverage of the investment deliverables (discussed above under factors) noting how contextual factors may account for any differential outcomes for, or participation in, the investment by relevant special groups; and extent to which the investment may be contributing to changes in the sector. Findings on key risks being monitored are presented and continuing risk assessed. Context does not have to be presented as a stand-alone section, but can be woven into discussions throughout the progress report, commonly as explanations for achievements, or barriers to progress. It can be helpful for risks to be addressed as a standalone section so that senior managers can quickly identify issues that need to be monitored (see element 3.9 below).

### *3.3. There is a reflection on the continuing relevance of the investment*

The end-of-program outcomes from the original design document are listed. Any revisions to these end-of-program outcomes that had previously been agreed are explained. There is a reflection on the extent to which the problem the investment is addressing is improving, worsening or continues the same and the continuing relevance of the current expected end-of-program outcomes in relation to the needs of the target population, including any relevant special sub-groups. Continuing relevance of the choice of interventions and program logic are also discussed where necessary. Any proposed changes are identified and justified.

### *3.4. An assessment of the adequacy of progress toward sustained end-of-program outcomes is described*

Where relevant, a distinction is made between end-of-program outcomes and those outcomes that are expected to be sustained beyond the life of the investment. A firm judgement of the **adequacy of progress toward these outcomes is described**. This also includes outcomes relating to cross-cutting issues such as gender equality, disability and social inclusion (see DFAT policies on the full range of issues). A simple list of

<sup>1</sup>Further guidance is available from participating DFAT ECB programs entitled: "Suggested Report Content for Implementation Partners"

immediate and/or intermediate outcomes or achievements is not sufficient. A judgement of the *adequacy* of this progress must be explicit, and explained with reference to appropriate criteria, standards and evidence.

Where progress has been particularly good, or particularly poor, a full **exploration of the supportive and inhibiting factors** that account for this situation is provided.

Particularly where problems or challenges have been identified, there is a full analysis of **the implications of the situation** on the likely achievement of end-of-program outcomes. It is clear from the report the extent to which DFAT and other stakeholders will need to track the relevant issues, including whether there is a need to reassess the suitability of end-of-program outcomes.

If previous **management responses** have been taken, then their success or failure is described. Any options for future management responses are fully elaborated. This includes the identification of what concrete actions stakeholders are required to take. The cost implications of these responses are discussed.

*3.5. The reach/coverage, quality and exposure of investment participants to relevant key outputs or deliverables for the reporting period are described*

Only the key outputs are described. Key outputs are defined as those outputs that are important to the achievement of expected end-of-program outcomes, or those accounting for a reasonable proportion of the budget or effort. Long lists of minor outputs are not required. The adequacy of the geographical or any relevant special sub-groups coverage of outputs is provided, the adequacy of the reach or number of outputs or beneficiaries of those outputs enumerated, the adequacy of the exposure of investment participants to key relevant outputs, and the quality of key outputs is assessed. An assessment of the quality of minor outputs is not necessarily required.

*3.6. The adequacy of progress implementing the annual plan is described*

Progress implementing the annual plan on time is described. A firm judgement is made on the adequacy of this progress. Where there are significant delays, the factors leading to this have been identified, and the implications for completing the investment on time are outlined. Management responses are proposed. Where extensions are requested these are adequately justified, and other alternatives have been considered.

*3.7. An assessment of the likely adequacy of planned inputs to meet the expected end-of-program outcomes is provided*

A firm judgement on the adequacy of the planned inputs to meet the expected end-of-program outcomes is made. The assessment makes clear the assumptions on which this judgement has been made, particularly regarding potential changes in implementation. Where there are anticipated shortfalls, the report provides a well-supported argument for any proposals for additional inputs.

*3.8. The adequacy of progress against the budget is assessed*

The amount of budget actually expended to date is presented against the planned budget. The proportion of variation is provided. For any issues identified, a full exploration of the supportive and inhibiting factors that account for this situation is provided. Particularly where problems or challenges have been identified, there is a full analysis of the implications the situation will have on the likelihood of delivering on the annual plan and budget. It is clear from the report the extent to which DFAT and other stakeholders will need to track the relevant issues. If previous management responses have been taken, then their success or failure is described. Any options for future management responses are fully elaborated. This includes the identification of which stakeholders are required to take what actions. The cost implications of these responses are discussed.

### *3.9. Key management or implementation systems are described and their performance assessed*

Key management or implementation systems are described where there are issues to consider. This could include: a) strategic and annual planning; b) governance and oversight c) monitoring, evaluation and risk management (including knowledge management systems); d) financial management and value for money; and e) staffing or human resource systems. Other relevant management systems are included. For any issues identified, a full exploration of the supportive and inhibiting factors that account for the situation is provided.

Particularly where problems or challenges have been identified, there is a full analysis of the implications the situations may have on the successful management of the investment. It is clear from the report the extent to which DFAT and other stakeholders will need to track the relevant issues.

If previous management responses have been taken, then their success or failure is described. Any options for future management responses are fully elaborated. This includes the identification of which stakeholders are required to take what actions. The cost implications of these responses are discussed.

### *3.10. The report provides balanced and fair reporting of positive and negative issues, achievements and challenges*

The report provides a balance between statements of achievement and challenges and issues. Statements of limited achievements would be supported by a reasonable discussion of challenges. The report reflects the challenging nature of human development.

### *3.11. For claims of achievement or barriers to achievement, credible supportive information is provided*

For claims of achievement (both in terms of achievement of outcomes, and quality of outputs or deliverables), or the explanation of barriers to achievement, credible supportive evidence is provided. The basis by which the claim is made is articulated. There is not an overemphasis on using examples to demonstrate achievement, rather the emphasis is on *how we know* that these examples have been achieved or outputs are of sufficient quality (the basis of the claim).

### *3.12. Data or findings are presented in formats that effectively and efficiently communicate important information*

Data or findings are presented in a way that allows the general reader to interpret the information appropriately and efficiently. A range of presentation formats have been considered and the report is not presented as text only. The presentation of information is easy to access, and is presented in formats suitable for the primary users.

### *3.13. The frequency of reporting is suitable for effective investment management*

Progress reports are submitted at suitable intervals to enable effective investment management. For example, activity implementation and budgetary information could be provided quarterly, where reports on the reach/coverage, quality and exposure of investment participants to the deliverables and the progress toward to achievement of outcomes could be provided annually. The submission of reports is timed to allow DFAT to meet its own internal reporting requirements and quality processes.

### *3.14. The report includes lessons learned from implementation that have potentially important implications more broadly*

The report includes lessons learned or insights that have been generated from the monitoring and evaluation activities that may have important implications for the investment, local partner development strategies, the DFAT delivery strategy, the DFAT Program, or DFAT corporately. Minor, well established or generic development lessons are not included.



### *3.15 Previous and/or proposed management responses or recommendations are summarized*

The report provides a summary of the important recommendations or management responses proposed. Any recommendations or management responses from previous progress reports are discussed in terms of their implementation and effectiveness.



## STANDARD 4

### TERMS OF REFERENCE (TOR) FOR INDEPENDENT EVALUATIONS

No.	Element
Background and Orientation	
4.1	A brief orientation to the investment is provided which includes: total value; time frame; expected end-of-program outcomes; a short summary of the key approaches employed in the investment; a brief explanation of the investment's expected contribution to the Program Strategy and local development planning.
4.2	The rationale for the evaluation, including the overall purpose and primary users of the information, is described
4.3	The TOR identifies the key decisions (management, operational and/or policy) which the evaluation is intended to inform
4.4	The key issues that lead to the evaluation questions are described in neutral language
Key Evaluation Questions and Scope	
4.5	The key evaluation questions are consistent with the overall purpose and management decisions of the evaluation
4.6	It is clear which questions are considered to be the priority questions for DFAT that must be answered with a good degree of rigor
4.7	The scope of the questions is suitable for the time and resources available to the team
4.8	Sufficient supporting information is provided about Key Evaluation Questions to guide the development of an appropriate evaluation plan
Evaluation Process	
4.9	A verbal briefing of the key issues and priority information is planned
4.10	Adequate time has been allocated for document review and document <i>appraisal</i>
4.11	There is a requirement for an elaborated evaluation plan – the depth of planning required reflecting the importance of the review/evaluation questions and management decisions
4.12	The submission date for the evaluation plan allows sufficient time for data collection activities to be scheduled
4.13	Proposed scheduling allows for adequate data collection and analysis to answer Key Evaluation Questions
4.14	A feed-back session to relevant information users is planned, either together or separately depending on the sensitivity of findings (Aide Memoire, discussion or presentation)
4.15	There is provision for processing the information collected to enable systematic analysis and interpretation, and the development of an evidence base
4.16	Adequate time is made available to complete the draft report

4.17	The process for commenting is efficient and allows independence of the evaluation team's final report
4.18	Adequate time has been allocated to responding to comments
4.19	The roles and functions of each team member are stated
4.20	The skill sets of the evaluators reflect the priority questions of the evaluation
4.21	The reporting requirements allow DFAT to track progress of the evaluation without distracting the team from carrying out important evaluation activities
4.22	The final evaluation report is published on the DFAT website as per the DFAT Aid Evaluation Policy

## Detailed Description of Standards for Evaluation Terms of Reference

*Note: The term evaluation is used in this document to refer to both reviews and evaluations. Evaluation would normally refer to a piece of work with a higher degree of methodological rigour usually requiring longer time frames and additional resources.*

## BACKGROUND AND ORIENTATION TO THE EVALUATION

### 4.1. A brief orientation to the investment is provided

As the terms of reference are used as a basis for discussion with proposed consultants (whether or not they are interested in conducting the evaluation, or to provide comment on it) the orientation must ensure the TORs are a stand-alone document. Important information includes: the total value; the time frame; a summary of the expected end-of-program outcomes including any specific outcomes around gender equality and social inclusion and key approaches to ensure effective gender integration across the program; a short summary of the key approaches employed (such as training, technical advisers, secondments, provision of infrastructure, equipment, and budget support or pooled funding). The context in which the investment is situated is described, such as the program strategy and/or delivery strategy that the investment aims to address and the partner government development plans of relevance. The delivery mechanism is described (contracted, multi-lateral development partner, NGO) and whether or not the investment is a project, program or facility. Any information which can help the reader to quickly understand the scope/reach of the investment is provided.

### 4.2. The purpose of the evaluation is described

The TOR clearly identifies the overall purpose(s) and shows which purposes are of most importance – accountability, investment improvement, knowledge generation, or developmental<sup>2</sup>. This allows the consultant to reflect these priorities in the evaluation plan. The *primary* users of the information are identified so that the consultant can collect relevant information, contribute to deepening an understanding of the findings during the mission, and prepare an appropriate report. Primary users are identified by title not only organization. For example, “DFAT” is made up of senior executive, desk officers, senior managers and investment managers. “The Contractor” is made up of head office personnel, implementation managers and advisers.

<sup>2</sup> Developmental evaluation is used in highly complex situations, or in programs that are in the early stages of innovation. See Gamble (2008) A Developmental Evaluation Primer. JW McConnell Family Foundation.

Although the *audience* is a wider group of stakeholders than the *primary users*, it is clearly articulated that the report will be published on the DFAT web-site in accordance with the Transparency Charter, and there are clear instructions on how sensitive information is to be communicated.

#### *4.3. The TOR identifies the key decisions (management, operational and/or policy) which the evaluation is intended to inform*

Any important management decisions that the primary users are expected to make are identified and described. Management decisions are more specific than the purpose and involve decisions such as whether or not to extend an investment, whether or not to involve a new partner, whether partner systems are ready for use, or whether to consider a new modality for a future investment.

#### *4.4. Key Issues are identified and discussed*

Any important issues that have informed the call for, or design, of the evaluation terms of reference are identified and described. They are described in neutral language and do not infer an expectation of findings. They are described in sufficient detail to enable the evaluator to develop the evaluation plan to adequately explore the issues.

## KEY EVALUATION QUESTIONS OR SCOPE

*Note: Evaluation questions are best generated together with primary information users. Ideally, participation includes representatives from senior management, investment level staff, and national partners and implementation partners where relevant. Senior managers are best placed to provide strategic direction to the evaluation, rather than delegating this to more junior program staff.*

#### *4.5. The key evaluation questions are consistent with the overall purpose and management decisions of the evaluation*

Each of the key evaluation questions is clearly related to the stated purpose(s) of the evaluation (and clearly related to the key management decisions, or a country program level monitoring and evaluation strategy where these have been developed). There are an adequate range of questions to meet all the stated purposes, and to ensure DFAT's information needs are met. There are no additional questions unrelated to the stated purpose. Although the DAC criteria are an important consideration for the evaluation, these have not been cut and pasted into the TOR resulting in broad questions of ambiguous scope. The TOR makes it clear to what extent gender equality and social inclusion issues ought to be investigated as stand-alone topics and/or integral to other questions of effectiveness, efficiency, sustainability and so on. There is a single list of questions in one place in the TOR.

#### *4.6. Priority evaluation questions are identified*

Of the full list of questions, the TORs clearly show what the priority questions for the evaluation are. This will allow the evaluator to make judgments during the evaluation of what questions must be answered in the final report, and what questions would be desirable if there is sufficient time. These priorities are consistent with the overall purpose of the evaluation. Ideally, only priority questions are posed, but in some cases where stakeholders have generated numerous questions which they want to keep in the TORs, prioritising these can be a way of showing the evaluator exactly what are the critical questions.

#### *4.7. The scope of the questions is suitable for the time and resources available for the evaluation*

Typically, a 12-day in-country evaluation can only address four or five broad questions. During the development of the evaluation plan, these are broken down into a larger number of sub-questions or information requirements. In addition, for a typical interview with a respondent without the need for

translation, only a small number of topics can be addressed with depth. Still, this is only possible if both the interviewer is skilled in questioning techniques and the respondents are relatively articulate and experienced in the topic areas. In addition to collecting the information, it also needs to be processed, interpreted and reported on. The questions posed in the TORs reflect this reality.

*4.8. Sufficient supporting information is provided about Key Evaluation Questions to guide the development of an appropriate evaluation plan*

Evaluation questions are not broad or vague or open to a wide range of interpretations. There is clarity in either the Key Evaluation Questions, or the supporting information provided. The evaluator will be able to break down questions and identify the specific information requirements. For this to be successful and for the purpose(s) to be met, the evaluator will need to be able to correctly interpret the expected information from the way the questions are worded. The Key Evaluation Questions (and supportive information) pose questions in a way that the evaluator can select suitable methods for the time and resources available (for example, cause-and-effect questions are difficult to answer in a short review without access to suitable secondary data sources).

## EVALUATION PROCESS

Adequate time and resources are required to enable the evaluation to be completed with an adequate degree of rigour. The following processes are allowed for:

*4.9. A verbal briefing of the key issues and priority information is planned*

A phone or face-to-face briefing is planned to discuss the back ground, issues and priorities for the evaluation with the evaluator before the evaluation plan is developed. Sufficient time is allocated to allow DFAT and the evaluator to work together to clarify scope, priority questions and issues, and general approach to methods. This may require more than one discussion.

*4.10. Adequate time has been allocated for document review and document appraisal*

Time has been allocated to *reviewing* investment documentation (approx. 2 days) as well as time to *appraise* any key documents such as gender equality, disability and social inclusion, or sustainability strategies, or the M&E system (often a day per document for full appraisal).

*4.11. There is a requirement for an elaborated evaluation plan – the depth of planning required reflecting the importance of the review/evaluation questions and management decisions*

The depth of planning required for an evaluation reflects the importance of the related decisions that will be made in response to the evaluation. If important decisions are to be made then more time is allocated. Typically for a DFAT-commissioned evaluation three days is required to develop an evaluation plan which includes fully elaborated methods. See Standard 5: Evaluation Plan for more details.

*4.12 The submission date for the evaluation plan allows sufficient time for data collection activities to be scheduled*

The data collection activities proposed by the evaluation team will be set out in the evaluation plan. This plan needs to be submitted to the evaluation manager well in advance of the in-country visit to allow for data collection activities such as interviews and site visits to be scheduled

#### *4.13. Proposed scheduling allows for adequate data collection, processing and analysis to answer Key Evaluation Questions*

The proposed schedule in the TOR is not too detailed as this is developed after the evaluation plan identifies suitable respondents and activities to address the evaluation questions. There are a sufficient number of days allocated to answer all the evaluation questions, as well as to work together as a team to process and discuss findings and identify further requirements as the mission unfolds.

#### *4.14. A feed-back session to relevant information users is planned, either together or separately depending on the sensitivity of findings (e.g. Aide Memoire, discussion or presentation)*

There is adequate time to provide detailed evaluation findings, to allow contestability of those findings, and test the feasibility of recommendations with key stakeholders. As the uses of information may be different for the different primary users, a suitable range of feedback options are offered. Consideration is given to whether feedback to representatives of different special sub-groups is necessary.

#### *4.15. There is provision for processing the information collected to enable systematic analysis and interpretation, and the development of an evidence base*

The evaluation team have been given adequate time to process information from interviews, document reviews and appraisals, observations or other methods to provide a credible evidence base to support findings. Typically, three days would be required for processing of data for a 12 day in-country mission that relied strongly on interviews. More complicated evaluations (or those with an emergent design) would require more time. This is additional time to actual report writing. Flexibility is balanced with value for money, but final time frames should be negotiated with the evaluator.

#### *4.16 Adequate time is made available to complete the draft report*

The number of days allocated to completing the report reflects: a) the scope of the evaluation questions; b) the complexity of the issues that have emerged; c) the number of people contributing to the writing of the report; d) team reviewing and discussions of the final draft. It is recommended to allow the evaluation team sufficient time to rest after the mission and to reflect on the mission. For example, ten days allocated to report writing could require a three week period to deliver. It is also useful to discuss with evaluators whether or not they expect to be working on other reports and missions during this time.

#### *4.17 The process for commenting is efficient and allows independence of the evaluation team's final report*

The process for commenting on the draft report is described and is efficient. Only relevant individuals are invited to comment, and the focus of their comments is identified. Consideration is given to inviting gender advisors and any relevant thematic group members to comment. Note that those invited to comment on the final report would also be invited to comment on the evaluation plan to ensure their final comments are within the scope and expectations for the evaluation. The TOR explains that DFAT will either provide comments in a consolidated form to the evaluation team, otherwise the TOR should allow additional time to respond to a large number of comments from all stakeholders.

Note: Be aware that where DFAT personnel consolidate comments, there must be transparency of decisions on what comments to include or remove. It may be necessary to provide comments from different stakeholders (national partner, representatives of any special sub-groups, implementation team and DFAT) separately if there are conflicting views.

#### *4.18. Adequate time has been allocated to responding to comments*

The time allocated to the evaluation team to respond to comments reflects a) the likely range of comments generated; and b) the possibility that comments require significant structural change in the final report.



#### *4.19. The roles and functions of each team member are stated*

Although it is the responsibility of the team leader to produce the final report and provide detailed direction on tasks in the evaluation plan, the TORs show DFAT expectations about how each team member will contribute. This is especially important with respect to writing responsibilities. There is clear guidance on the extent to which DFAT expects international and national consultants to participate and assume responsibility for particular tasks. If there is any requirement of the team leader for capacity building of team members, then adequate time has been allocated to carry this out effectively.

#### *4.20. Skill sets of evaluation team reflect priority questions*

Unless there is a compelling reason provided, the team leader is an evaluation expert, not only a technical expert in the relevant sector or thematic area. They are supported by technical specialist(s) who will focus on technical aspects of the evaluation. The development of the evaluation plan is allocated to the team leader who will be responsible for its implementation. The development of the evaluation plan is not allocated to a team member. The tasks and balance of work of technical advisers reflects the evaluation questions. For example, if there is a strong focus on gender, or investment management systems, then the number of days allocated to technical specialists from other sectoral areas reflects this focus. Team members have at least awareness level knowledge of gender equality in relation to the program's focus area(s), even if the evaluation does not specifically focus on questions around gender equality.

#### *4.21. The reporting requirements allow DFAT to track progress of the evaluation without distracting the team from carrying out important evaluation activities*

The requirement for reports during the mission provide for a good balance between monitoring the progress of the evaluation and allowing the team to focus on important evaluation activities. The evaluation plan is a critical document for DFAT to ensure that the evaluator has correctly interpreted the TORs and has made suitable plans to conduct the evaluation to meet the TORs and reasonable standards of rigour. Other reporting requirements to consider are a) the aide memoire (which should be short and only provide anticipated key findings and recommendations); and b) the provision of any processed data. If requiring the provision of processed data, it is important to consider the implications for preserving the confidentiality of respondents.

#### *4.22. The final evaluation report is published on the DFAT website as per the DFAT Evaluation Policy*

The requirement for evaluation reports to be published on the DFAT website should be noted at initial stages and planning put in place to ensure this requirement is met, as per the DFAT Evaluation Policy.

Note: A negotiation of the TORs should be encouraged during the contract negotiations. This allows the team leader to provide professional advice on the feasibility of the TOR in terms of the scope of questions and the resources applied.

## STANDARD 5

### INDEPENDENT EVALUATION PLANS

No.	Element
5.1	The evaluation plan is based on a collaborative approach
5.2	The primary intended users of the evaluation are clearly identified and their evaluation needs are described
5.3	The purpose and/or objectives of the evaluation are stated
5.4	A summary is provided to orient the reader to the overall evaluation design
5.5	Limitations or constraints on the evaluation are described (e.g. time frame; resources; available data; political sensitivities)
5.6	The Key Evaluation Questions are supplemented by detailed descriptions and/or sub questions
5.7	It is clear which questions are considered to be of higher priority and are expected to provide the most important information
5.8	There is sufficient flexibility to be able to address important unexpected issues as they emerge
5.9	The methods to collect data are described for each question (or related questions)
5.10	The proposed data collection methods are appropriate for the questions posed
5.11	Triangulation of data collection methods is proposed to strengthen the confidence in the findings
5.12	The sampling strategy is clear and appropriate for the evaluation questions posed
5.13	The plan describes how data will be processed and analysed
5.14	The plan identifies ethical issues and how they will be addressed
5.15	The process for making judgments is clear
5.16	Approaches to enhance the utilization of findings are outlined (if this has been requested in the terms of reference)
5.17	The evaluation plan provides guidance on scheduling. The final schedule (if attached) reflects adequate time to answer the posed evaluation questions
5.18	The allocation of evaluation tasks to team members is clearly described (i.e. data collection, processing and reporting)
5.19	The plan for publication of the final evaluation report is documented

## DETAILED DESCRIPTION OF STANDARDS FOR EVALUATION PLANS

*Note: The Evaluation plan is developed by the evaluator based on the ToR. It is a negotiated document between the client and the evaluator and should provide more detail and reflect final agreements after that negotiation. The evaluation plan should be submitted as early as possible, to enable scheduling of site visits, interviews and other data collection activities. The agreed Evaluation Plan ought to provide the basis by which evaluator performance is assessed.*

### *5.1. The evaluation plan is based on a collaborative approach*

The evaluator has consulted DFAT, and the stakeholders identified as important by DFAT, to develop the evaluation plan. Consultation may have been in-person, by phone or by email. Important stakeholders have been given the opportunity to comment on the evaluation plan before the evaluation commences. *Note: This ensures that additional information will not be requested after the data collection phase is complete.*

### *5.2. Primary intended users of the evaluation are clearly identified*

An evaluation cannot meet the needs of all interested stakeholders. Individuals (by title) in named organizations should be identified as the *primary users of the evaluation findings*. These are the people who will be using the information to make judgments and decisions. Audience is a different concept and often refers to a broader group of people that may be interested in, or may be affected by any decisions that result from the evaluation.

### *5.3. The purpose and/or objectives of the evaluation are stated.*

These would normally be taken from the terms of reference. The evaluation plan restates these so that it is a stand-alone document.

### *5.4. A summary is provided to orient the reader*

This is an introductory orientation of the overall design of the evaluation. It is short, about one paragraph in length. For example, it could highlight whether the evaluation is predominantly exploratory or descriptive, or whether a cause and effect design is proposed, or whether or not any case studies would feature in the overall design. It would highlight the major methods for data collection and analysis. This is called the investigatory framework in research and evaluation terms. The evaluation plan does not go straight into detailed descriptions of methods without this general orientation.

### *5.5. Limitations or constraints for the evaluation are described*

The time available for the evaluation has implications for the scope of the evaluation. If a large number of questions are posed, but DFAT only wants a cursory look at many of these, then a shorter time frame may be appropriate. The evaluator highlights any important limitations in terms of time available, resources applied, or the expertise of the evaluation team to deliver a credible, defensible evaluation product. Consideration is also given to any potential limitations in addressing gender equality issues and/or reaching representatives of special beneficiary sub-groups. Political sensitivities are highlighted where appropriate. The implications of these limitations are discussed. *Note: A long list of limitations is not considered a substitute for a poorly negotiated TOR.*

### *5.6. The Key Evaluation Questions are supplemented by detailed descriptions and/or sub questions*

Although the terms of reference is where DFAT communicates what the evaluation is to address, the evaluator will still need to give careful consideration to how these larger questions will be addressed. This means that more detailed information requirements and/or sub-questions are generated. Commonly, questions presented in a terms of reference are broad, therefore this more detailed information allows information users to know how the evaluator has interpreted the broader questions, and whether or not the

evaluation will generate sufficient information to meet these broader questions. It also allows the DFAT evaluation manager to see the implications of the scope of the evaluation described in the terms of reference. This breakdown of information requirements or questions allows the reader to assess whether or not the original scope was realistic. Where relevant, the evaluator demonstrates how gender equality and social inclusion issues are investigated as a stand-alone area of enquiry or as integral to other questions of effectiveness, efficiency, sustainability and so on. The evaluation manager needs to pay careful attention to this aspect of the evaluation plan.

#### *5.7. High priority questions are identified*

DFAT evaluations often have a very large number of evaluation questions that cover a very wide number of aspects of the investment to be evaluated. Some of these questions will be more important than others. The evaluation plan reflects where the emphasis will be placed, and it is clear that DFAT's information needs will be met. The evaluation team will not usually be able to answer all the questions listed for all respondents and so will need to make decisions during interviews about what will be dropped and what is essential. The evaluation manager needs to be confident that the evaluator will, at a minimum, deliver information on the priority questions.

#### *5.8. There is sufficient flexibility to be able to address important unexpected issues as they emerge*

This flexibility may be built in to the questioning technique employed during an interview. It may be built into the schedule as a whole to allow new issues to emerge and be responded to through additional data collection if they are important. Where new issues cannot be adequately addressed within the schedule, there are processes to review possible trade-offs to allow them to be addressed.

#### *5.9. Methods for each evaluation question are described*

The evaluation plan shows how each of the evaluation questions will be answered by describing the methods that will be used to collect the information. For most DFAT evaluations this is likely to include in-depth interviews, focus group discussions/interviews, document reviews and in some cases observations of activities. Large workshops are not usually a suitable method to gather substantive, reliable and valid information – however, they may have other important political purposes. For several questions there may be a number of data collection methods proposed to strengthen confidence in the findings.

Consideration is given to the design of data collection methods that are responsive to the needs, rights and security of respondents, with special consideration given to the needs of any special sub-groups. This includes the identification and recruitment of investment participants as respondents, the time of day and location of data collection activities, and the techniques employed during data collection.

The design of major evaluation activities/studies are annexed and include tools such as interview guides or questionnaires. In some cases the evaluator will need to develop these later, or adjust them as the evaluation proceeds, but there is an absolute expectation that the evaluator uses tools to guide each evaluation activity, and does not rely on memory of all the evaluation questions identified in the evaluation plan. Where team members are working in different locations then tools are available ahead of time so that data is collected systematically. If flexibility on this is required, then a compelling rationale is provided. Summary statements of methods that are not linked with specific evaluation questions are not considered adequate.

#### *5.10. Methods are appropriate for the evaluation questions posed*

Although this takes evaluation expertise, it is still worth reviewing the questions posed and consider if the methods described could reasonably answer the questions. For example, a focus group discussion would be most unlikely to answer a sensitive question; a review of a program strategy document (such as gender)

would be unlikely to tell you if the investment's actual gender activities were of a high quality. It would need to be supported by information from other sources.

#### *5.11. Triangulation of methods is proposed*

Triangulation is the use of a range of methods and/or sources of information to come to a conclusion or result. It can develop greater confidence in a finding. Given the short time frame of most DFAT evaluations or reviews, it is difficult to employ a wide range of methods. To deal with this, the evaluation has planned to discuss similar questions across a range of different respondents within and across different organizations or target beneficiary groups (particularly special sub-groups), or use a number of methods to examine the same issue. It is not sufficient to state that triangulation will be used if this is not demonstrated in the evaluation design.

#### *5.12. Sampling strategy is clear and appropriate*

Most evaluations will require some sort of sampling strategy across individuals and special sub-groups, sites or time periods. Appropriate sampling strategies are chosen and justified. For short reviews that rely on analytical rather than statistical inference, purposeful sampling will be appropriate and could include maximum variation, a critical case, or a typical case. Efforts should be made to avoid relying on a convenience sample which is likely to be unrepresentative of the population of interest. Where statistical inference will be used to generalize from the sample, random sampling strategies are appropriate – especially stratified random sampling which reduces the sample size required.

#### *5.13. The plan describes how data will be processed and analysed*

The evaluation plan describes how the data will be processed, including measures to check and correct any errors in data, ensure security of storage and prepare for analysis. The plan also describes how the data will be analysed in order to answer the Key Evaluation Questions. This may not necessarily require advanced analytical methods, but users of the information can determine exactly what is to be done. Consideration is given to the analysis of disaggregated data for gender and other relevant sub-groups where data collection or secondary sources has allowed this.

#### *5.14. The plan identifies ethical issues and how they will be addressed*

For most of the evaluations and reviews conducted by DFAT, this will mostly be around privacy and confidentiality issues. The plan identifies how these will be addressed when data are collected, stored and reported. In particular, assurances about anonymity must be honoured and data stored and reported in ways that do not inadvertently identify informants, including when providing a database of the evidentiary basis to DFAT as part of the deliverables. Other relevant ethical issues are addressed including processes for reporting serious issues if identified during data collection.

#### *5.15. The process for making judgments is clear*

The evaluation plan makes it clear that the evaluator will make professional judgements based on the evidence gathered and on the agreed criteria or standards. DFAT's response to the evaluator's judgement should be provided in the Management Response to an evaluation. In some exceptional cases, DFAT may require an evaluator to report neutrally on facts and leave DFAT to make the final judgements, in which case the plan should make it clear how evaluative judgements will be made and by whom, as this is an important distinction and can affect the way information is collected and presented.

#### *5.16. Approaches to enhance utilization of findings are outlined*

The importance of utilization of findings needs to be communicated to the evaluator. There are a variety of well-tested approaches to utilization that a professional evaluator will be familiar with (e.g. stakeholder engagement strategies for evaluation design or developing acceptance of recommendations before the

report is published). Approaches to utilisation of findings are outlined in the evaluation plan. Utilization begins with the evaluation design stage. Consideration is given to providing feedback to representatives of women's groups and relevant special sub-groups.

#### *5.17. Scheduling guidance is provided*

The schedule is developed by DFAT after the evaluation plan is submitted, and reflects guidance from the evaluator. The most common problem is that the persons recruited for interview are not always the best respondents for the evaluation questions posed. Often there are many donor meetings where respondents cannot provide substantive comment on many of the evaluation questions. Also consider the time for each interview with the associated evaluation questions. Most 60 minute interviews with a respondent cover no more than four or five key topics; less if translation is required. Sufficient time is available to meet with the implementation team. As part of reviewing the methodology DFAT negotiates the proposed list of respondents before final scheduling. The evaluator scheduling guidance is realistic for the time frame. Sufficient time is allocated to other methods proposed. There is sufficient time allocated to evaluation team discussions and early data processing (not late at night). Sufficient time is given to the recruitment of respondents who represent special sub-groups, any time required to travel to the appropriate locations, and setting the tone for interactions that enable respondents to provide their views with confidence and in comfort.

#### *5.18. Evaluation tasks are allocated to team members*

It is very important that each team member knows before the evaluation begins what they will be expected to do. It is not appropriate for the team leader to allocate reporting responsibilities on the last day of the in-country mission. The evaluation plan shows what responsibilities each team member has so they can ensure that adequate data is collected, processed, and interpreted and they can meet a high standard during the reporting stage. It is often useful to show which evaluation questions each team member carries responsibility for.

#### *5.19 The plan for publication of the final evaluation report is documented*

The plan for publication of the final evaluation report should be documented as part of the Evaluation Plan in order to facilitate meeting the requirements of the DFAT Aid Evaluation Policy.



## STANDARD 6

### INDEPENDENT EVALUATION REPORTS

No.	Element
<b>Introductions</b>	
6.1	A background to the evaluation summarizes: the total value of the investment; the number of years of the investment; the stage of investment implementation; key outcomes of the investment; and the key issues identified in the terms of reference
6.2	A brief summary of the methods employed is provided
6.3	Key limitations of the methods are described and any relevant guidance provided to enable appropriate interpretation of the findings
6.4	The executive summary provides all the necessary information to enable primary users to make good quality decisions
<b>Findings and Analysis</b>	
6.5	The evaluation report clearly addresses all questions in the Terms of Reference
6.6	The relative importance of the issues communicated is clear to the reader
6.7	There is a good balance between operational and strategic issues
6.8	The report clearly explains the extent to which the evidence supports the conclusions and judgments made
6.9	Alternative points of view are presented and considered where appropriate
6.10	Complicated and complex aspects of issues are adequately explored and not oversimplified
6.11	The role of context and emergent risks to investment performance are analysed
6.12	The text uses appropriate methods/language to convince the reader of the findings and conclusions
6.13	There is an adequate exploration of the factors that have influenced the issues identified and conclusions drawn
6.14	The implications of key findings are fully explored
6.15	The overall position of the author is clear and their professional judgments are unambiguous.
<b>Conclusions and Recommendations</b>	
6.16	The conclusions and recommendations logically flow from the presentation of findings and any associated analyses
6.17	Individuals have been allocated responsibility for responding to recommendations
6.18	Where there are significant cost implications of recommendations, these have been estimated (financial,

	human and materials costs)
6.19	The recommendations are feasible
6.20	The circumstances under which any important lessons are transferable are described
6.21	The final evaluation report is published within the timeframes outlined in the DFAT Aid Evaluation Policy

## DETAILED DESCRIPTION OF STANDARDS FOR EVALUATION REPORTS

### INTRODUCTIONS

#### *6.1. The background provides adequate information for individuals not familiar with the investment*

The background provides adequate information to enable individuals not fully familiar with the investment to interpret the report. It summarizes: the total value of the investment; the number of years of the investment; the stage of investment implementation; the delivery mechanism; key expected outcomes of the investment; and the key issues identified in the terms of reference.

#### *6.2. A brief summary of the methods employed is provided*

Although a fully elaborated methodology was developed before the evaluation, a summary of the significant details is included. Sufficient information is required to enable the reader to quickly understand the evidentiary basis of the evaluation. The evidentiary base must be convincing and in proportion to the resources invested in the evaluation. The full methodology is annexed. Important aspects of the strategy to ensure findings are utilised are summarised here.

#### *6.3. Key limitations of the methods are described and any relevant guidance provided to enable appropriate interpretation of the findings*

Key limitations are summarised in the evaluation report to enable the reader to make appropriate decisions. Where necessary the author has provided specific guidance of where the reader ought to be cautious about the findings.

#### *6.4. The executive summary provides all the necessary information to enable primary users to make good quality decisions.*

The executive summary provides all the necessary information to enable *primary stakeholders*, especially senior management, to make good quality decisions without reading the entire document. It is not a simple cut and paste of the main body of the report. It summarises the key findings, provides sufficient analyses and arguments, and presents final conclusions and recommendations. Important information about gender equality and social inclusion are included to allow the reader to appreciate important achievements and challenges. Resource implications of recommendations are summarised. The length of the executive summary is proportionate to the length of the report (e.g. two to three pages for short uncomplicated reports, and up to five or six pages for more lengthy reports with complex issues).

### FINDINGS AND ANALYSES

#### *6.5. The evaluation report clearly addresses all questions in the Terms of Reference/Evaluation Plan*

Note: As the Evaluation Plan supersedes the Evaluation Terms of Reference, the Plan is the appropriate document to assess whether or not the evaluation has delivered on expectations. In the absence of an Evaluation Plan, the Terms of Reference should be used.

It is relatively easy to identify where each of the questions in the Evaluation Plan are addressed. The report does not need to be a mechanical presentation of these questions, but it should be relatively easy to negotiate the report and find relevant information about specific questions in the Evaluation Plan. Where there are gaps, these have been explained. DFAT's information needs, as set out in the Terms of Reference and Evaluation Plan, have been met.

*6.6. The relative importance of the issues communicated is clear to the reader*

The report makes it clear what issues are priority issues to consider. Minor issues are not set out mechanically against the terms of reference and given the same depth of treatment as more important issues. The breadth of description, depth of analysis and attention in the recommendations can indicate the degree of priority. The author may simply state the relative importance of issues.

*6.7. There is a good balance between operational and strategic issues*

The report addresses the full range of issues identified in response to the TOR and other critical issues that have emerged. There will be technical, managerial or operational issues that are very important to consider and are often at the core of many important challenges. The strategic direction or any higher order issues of the investment have been given adequate space, and minor technical issues are treated in a more limited fashion. Flexibility is required where the TOR evaluation questions demonstrate that this balance was not required.

*6.8. The report clearly explains the extent to which the evidence supports the conclusions and judgments made*

For key findings, the basis of the findings and related conclusions is communicated clearly. This includes reporting the degree to which views are shared across respondents, and representatives of relevant special sub-groups if appropriate. The information is brought together from a range of sources, but communicated as a coherent whole. Evaluator opinions that are based on limited evidence are made transparent and proposed as suggestive only.

*6.9. Alternative points of view are presented and considered where appropriate*

Alternative views must be presented, especially for important, controversial or disappointing findings. They are not immediately dismissed, but are seriously considered. Key stakeholder views such as those of the implementation team must be given sufficient attention, and balanced by national partners, DFAT or other important stakeholder views.

*6.10. Complicated and complex aspects of issues are adequately explored and not oversimplified*

The report adequately acknowledges complicated aspects of issues, such as multiple contributing factors, or emergent challenges and opportunities. The report does not present simple solutions to these types of situations. The findings are presented fairly so that specific stakeholders are not held fully accountable for problems when multiple factors are involved. Human development is challenging, and the report recognises that implementation teams and national partners are often facing multiple challenges.

*6.11. The role of context and emergent risks to investment performance are analysed*

The report identifies relevant aspects of the context within which activities are implemented. These might include geographic, cultural, gender, political, economic or social context. Sufficient information is presented

to allow the reader to understand the relationship between the investment and its context. The report addresses: a) how the context may have affected the achievement of outcomes (both supportive and inhibiting); and b) the extent to which the investment may have had any effect on the context. Important emergent risks are identified.

*6.12. The text uses appropriate methods/language to convince the reader of the findings and conclusions*

Arguments presented do not use emotive word choices in an effort to appeal to the emotions of the reader. The method used to convince readers is the presentation of evidence or a credible basis for the finding. Using the international literature to build the credibility of the report can be effective. The report handles political issues with sensitivity. A good report considers the expected positions of the important stakeholders – if findings are unexpected then this is carefully communicated and explained in the text.

*6.13. There is an adequate exploration of the factors that have influenced the issues identified and conclusions drawn*

It is not sufficient to simply describe a situation. A full analysis of the likely factors that have led to the situation is necessary. Factors that enable progress or achievement are just as important as factors that inhibit them. These factors should be generated from a range of data sources. A range of causes should be considered rather than regularly offering a single cause for major and/or complex issues.

*6.14. The implications of key findings are fully explored*

DFAT investment managers, senior management and other stakeholders need some direction on the implications of the findings if this is not immediately apparent. Implications for achieving investment objectives, and for meeting time frames, expenditure projections, or sustainability are often important considerations.

## CONCLUSIONS AND RECOMMENDATIONS

*6.15. The overall position of the author is clear and their professional judgments are unambiguous.*


The task of the evaluator is to evaluate. They must make their position clear (and as early as possible in the report) unless the TORs have required the evaluator to report on findings with neutrality. The report does not simply state findings and expect DFAT to interpret them and draw their own conclusions. The report presents the authors view unambiguously. Has the investment made adequate progress or not? Are the factors that have accounted for the limited achievements been unavoidable or are they due to poor management. Unambiguous judgements also present findings and conclusions sensitively and constructively.

*6.16. The conclusions and recommendations logically flow from the presentation of findings and any associated analyses.*

It is possible to trace issues through the text from description, to analysis, to conclusion and recommendation. No recommendation appears at the end that is not supported by descriptive and analytical work in the text. There are no important inferred recommendations buried in the text that have not been drawn into the conclusion or list of recommendations at the end.

The “chain of evidence” is evident. This is where all questions in the methodology have data that has been collected, analysis conducted, findings presented, interpretation carried out and reported. If questions in the methodology have not been addressed then an explanation has been given.

*6.17. Individuals have been allocated responsibility for responding to recommendations.*



Where appropriate, job titles, rather than organisations, have been allocated responsibility for actions against all recommendations. If it is not appropriate or possible to identify the individual, then the relevant work group is identified. If some recommendations are for broader partner government, or DFAT sectoral or corporate learning then these are identified separately.

*6.18. Significant cost implications of recommendations have been estimated*

If recommendations imply human, financial or material costs, these are estimated. If recommendations for additional technical support are made, then the number of days input is estimated. For important technical assistance positions proposed, the key content to consider for the terms of reference is annexed.

*6.19 The recommendations are feasible*

Recommendations, in the most part, are acceptable to relevant stakeholders (recommendations that stakeholders do not agree with rarely get implemented – coming to acceptability is dealt with by the utilisation strategy). Recommendations are feasible from a resourcing and cost perspective. Recommendations are likely to be effective to rectify a situation, or to achieve an expected outcome.

*6.20 The circumstances under which any important lessons are transferable are described*

Where there are important lessons to be learned, the report provides sufficient information to inform the reader about the circumstances under which these lessons can be transferred. This could be at the sector level, the Country Program Level, for the Department as a whole, or for the development sector more broadly.

*6.21 The final evaluation report is published within the timeframes outlined in the DFAT Aid Evaluation Policy*

The final evaluation report must be published within the timeframes outlined in the DFAT Aid Evaluation Policy.

## STANDARD 7

### MONITORING VISITS

*Note: A Monitoring Visit is a visit conducted by DFAT personnel on an investment in their portfolio. It can also be conducted jointly with implementation partners. There is no DFAT-wide approach to how monitoring visits should be undertaken – this should be determined in accordance with the context and information needs of the investment manager. This standard provides one possible format for monitoring visits.*

No.	Elements
<b>The Monitoring Visit Plan</b>	
7.1	The number of the monitoring visit for the investment is recorded
7.2	The broad purpose of the visit is described
7.3	The key questions for the visit are listed and are related to the purpose of the visit
7.4	The scope of the questions is suitable for the time and skills available for the visit
7.5	How information will be collected is described
7.6	The proposed ways of collecting information are suitable for the questions posed
7.7	A proposed schedule of expected activities and persons to meet is prepared
7.8	The proposed schedule allows sufficient time to address the visit questions
7.9	Roles and responsibilities for each team member are clearly described (including joint missions)
<b>Recording and Managing Information</b>	
7.10	Key points or issues are recorded after each activity
<b>Visit Report and Response</b>	
7.11	A background to the visit is presented
7.12	A brief summary of findings for each visit question posed is provided.
7.13	Important lessons learned or insights of general interest to the sector, the program, or DFAT corporately are identified
7.14	Information relevant to updating DFAT reporting is presented in language suitable for pasting into relevant reports
7.15	Management responses are summarized
7.16	Feedback to respondents or participants is provided where appropriate
7.17	The date the report is sent to the management team for oversight of management responses is recorded.
7.18	Follow Up section is completed



## DETAILED DESCRIPTION OF STANDARDS FOR MONITORING VISITS

*Note: If several DFAT staff travel together on a Monitoring Visit, it is only necessary to compile one joint monitoring plan and report. In such cases, one person will be nominated as the visit leader and they will prepare the plan and write the report. For joint missions with other implementation partners, the plan and report are developed together to ensure all information needs are addressed.*

### THE VISIT PLAN

The visit plan is sent to relevant management staff on completion and attached to the travel request.

*7.1. The number of the monitoring visit for the investment is recorded. Over the life of each investment there will be a number of visits. The visit plan shows what number the visit is, and the number of months the investment has been under implementation.*

*7.2. The broad purpose of the visit is described*

Monitoring visits are not expected to meet a list of pre-determined questions or purposes. In the most part, they are useful to fill in gaps in knowledge about the investment that are important for DFAT routine reporting, general management and decision making. They can also be used to learn more about the context in which the investment is operating. Therefore, the broad purpose can address a range of relevant information needs. The purpose is clearly described.

*7.3. The key questions are listed and are related to the purpose*

Like in an evaluation, the visit is guided by key questions that need to be answered. It is clear how the information generated from these questions will be used, and that this is consistent with the purpose. The plan reflects on management responses from earlier DFAT reports and identifies any follow-up that might be required. Consideration is also given to whether gender equality and social inclusion issues will be treated separately or as integral to other questions of the visit.

*7.4. The scope of the questions is suitable for the time and skills available for the visit*

Developing an in-depth understanding of a few focused issues will be more advantageous than trying to cover a wide range of issues poorly. The scope of the questions is quite focused on information needs of DFAT and others involved in the visit. The questions do not require complex methods to collect and process the information. The questions are simple questions that a generalist using good common sense has a reasonable expectation of answering.

*7.5. How the information will be collected is described*

The visit plan shows how the information will be collected for each question. Who will be interviewed, what will be observed, what documents will be reviewed, and where informal interactions are expected. For interviews, short interview guides are developed; for observations, what will be observed is briefly described; for documents, what will be read will be described. It is not a lengthy description, but provides sufficient guidance to ensure that the visit will deliver the expected information without major gaps. Consideration is given about how to reach and engage representatives of different special sub-groups in a way that is sensitive to their needs, rights and security.

*7.6. The proposed ways of collecting the information are suitable for the questions posed*

Consideration is given to the best way of collecting information to answer the questions posed. For example, it may be more effective to observe training, or review a training agenda to discover the quality of training,

than ask the participants if it was good quality. However, interviewing participants or their supervisors is more effective if you want to learn about the application of training.

*7.7. A proposed schedule of expected activities and persons to meet is prepared*

Once the ways of collecting information are fully described, a schedule of activities such as observations of investment activities in progress or meetings is developed. The date is recorded when this was sent to hosts (either government or implementation partners).

*7.8. The proposed schedule allows sufficient time to address the visit questions*

The schedule allows sufficient time to answer the questions to a reasonable degree of depth. The schedule reflects the scope of the visit.

*7.9. Roles and responsibilities of visit team members are described (including for joint missions)*

The designations and roles and responsibilities of the different team members are clear. It is poor practice to allocate roles and responsibilities during the mission. For example, if an investment manager is expected to act as a translator, then they cannot also collect information, take notes and be able to process information after the meeting. Whoever is responsible for writing the report from the visit needs to be able to control the information that is collected so that they can answer the questions posed in the visit plan.

## RECORDING AND MANAGING INFORMATION

*7.10. Key points are recorded after each visit activity*

The DFAT staff member responsible for the interview(s) or other information collection activity ensures that there are dot points recorded against the key questions after each activity. These recorded notes provide the basis for the final report. Dot points address: a) the visit questions in terms of *actual content* provided by respondents, or observed; and b) the *impressions* of the DFAT staff member. Record notes in a manner sensitive to respondent comfort and sense of security.

## VISIT REPORT

*7.11. A background to the visit is presented*

This describes: a) the purpose of the visit; b) the number of the monitoring visit for the investment; c) the visit questions; d) the dates of the visit; e) participating visit team members (this could be DFAT and joint visits with other partners); f) the cost (person days and expenditure). If there are any important issues that led to the visit, these should be very briefly summarized.

*7.12. A brief summary of findings for each visit question posed is provided*

A summary is provided for each question. It is as brief as possible as these reports may be read by a wide range of colleagues. The four analytical questions are reflected in the text where relevant: a) what is the current situation; b) what are the factors that have led to the situation; c) what are the implication to investment success; d) what are the responses that have been taken thus far (and their success), and what future responses are still required. Any differential outcomes or engagement in the investment for subgroups are noted here.

*7.13. Important lessons learned or insights of general interest to the sector, the program, or DFAT corporately are identified*

The report includes lessons learned with potentially important implications for the investment; local partner development strategies; the DFAT delivery strategy; the DFAT Program, or DFAT corporately. Minor, well established or generic development lessons are not included. This section can be left blank if there is nothing of note.

*7.14. Information relevant to updating DFAT reporting is presented in language suitable for pasting into relevant reports*

This information is presented in very brief language in a format that can be cut and pasted into DFAT reports. This information is arranged in accordance with required headings in routine reports.

*7.15. Management responses required are summarised*

All the proposed management responses are listed according to who is required to take responsibility for the recommended action (e.g. Implementation Partner, investment manager and relevant management staff, Counsellor, and/or local partner). Each management response has a time frame for action, a description of the steps required, and whether or not it has been negotiated with partners if they are responsible. Resource implications are highlighted where relevant.

*7.16 Feedback to respondents or participants is provided where appropriate*

Any feedback provided to participants in the monitoring visit is recorded. This feedback can be provided to implementation teams, local partners or beneficiaries from relevant special sub- groups where appropriate. A plan to provide more detailed feedback after the visit could be included.

*7.17. The date the report is sent to the management team for oversight of management responses is recorded.*

Management staff will need to be aware of the management responses that are generated from each investment they are responsible for. They are responsible for supervising the follow-up required from each visit.

*7.18. The follow up section is completed*

All management responses are expected to be followed up, or a reason recorded for non-follow up. As each management response is completed, the date is recorded. When all the management responses for a visit have been completed, then the visit report is “closed” and the updated version forwarded to relevant management staff. If a management response is not taken then this is explained before closing the report.



## FEEDBACK

If you have any comments or feedback about this Tool, or related business processes information, please contact either [qualityreports@dfat.gov.au](mailto:qualityreports@dfat.gov.au) or [opeval@dfat.gov.au](mailto:opeval@dfat.gov.au). A log of all feedback received will be kept and used to feed into a process of annual review.