

AusAID Indonesia Country Program
Evaluation Capacity Building Program

Standards for Monitoring and Evaluation

Developed for the Indonesia Program by the Indonesia M&E Help Desk

Dr Susan Dawson

Version 2 (first annual revision): 11th November, 2010

- Draft Version: pilot testing not completed -

Along our own pilot testing of the standards we are collating feedback and views from others. All information will be used when we revise the standards towards the end of 2011. Please email any feedback to: sofia.ericsson@ausaid.gov.au

Table of Contents

Introduction	3
Standard 1: Initiative Design (required features for M&E)	6
Standard 2: Initiative Monitoring and Evaluation System	10
Standard 3: Initiative Progress Reporting	15
Standard 4: Terms of Reference for Independent Evaluations	19
Standard 5: Independent Evaluation Plans (Methodology)	24
Standard 6: Independent Evaluation Reports	29
Standard 7: Initiative Supervision Visits	34

Introduction to the Standards

Why Do We Need Standards?

AusAID's Indonesia Program is currently implementing an institutional change program to enhance the quality and utilisation of Monitoring and Evaluation (M&E) systems and deliverables and to integrate evaluative thinking into everyday work. This is called the "Evaluation Capacity Building Program" and it considers "capacity" in the broadest sense.

By the end of the ECB program in 2015, we expect to see the following outcomes:

- Information that is generated by M&E systems and activities have a credible basis;
- Quality information informs and influences decision making at the initiative, sectoral, thematic and country program levels;
- Effective mechanisms are operating for integrating lessons from the ECB program - and integrating corporate developments into the ECB program;
- The Performance and Quality Unit in Jakarta is fully institutionalised; and
- The Performance and Quality Unit is harmonised with IET Performance Unit in Canberra.

Analysis of the constraints to delivering quality M&E systems and products across the Indonesia and other Country Programs has shown 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 system. This lack of agreement has been observed across AusAID, and within the development and evaluation industry. Technical advice on M&E has not been consistent, and this has created unnecessary constraints for AusAID's Indonesia Country Program in their efforts to bring about institutional change. In addition, where M&E products are appraised by independent or AusAID personnel, feedback provided is often not in line with international standards, or agreed expectations communicated by the Indonesia Performance and Quality Unit in Jakarta and Canberra.

Therefore, the development of standards is a *critical* step in the design and management of institutional change across the program. Standards provide the basis for:

- Articulating clear and consistent messages to staff, contractors and consultants about the expectations of the quality of M&E systems and deliverables;
- Providing the basis to measure the performance of the Indonesia Program with regard to improving the quality of M&E systems and deliverables;
- A credible and transparent evidence base to communicate credible information on achievements and challenges in the effectiveness of the institutional change program.

These standards are the first step toward a clear articulation of what quality is expected from a range of M&E products at the *initiative level* across the Indonesia Country Program. In due course, standards at the sectoral and country program level will be added to this series of standards.

The ECB Baseline and Follow-Up Studies

As mentioned, the standards provide the basis by which the baseline and follow-up studies are carried out. For each set of standards, baseline measures have been developed. For several of the standards multiple measures are required, but all measures are directly related to the concepts in each standard.

For each set of standards, the Indonesia Performance and Quality (P&Q) Unit will carry out a systematic assessment of the current quality of M&E products reflected in the standards. This will not only provide credible evidence of the current situation, but allow the P&Q Unit to communicate information about improvements with sufficient confidence to inform decision making.

What is the Basis of these Standards?

Although the international standards for M&E are well described in the literature, and are generally quite consistent, these standards are focused specifically on the needs of AusAID, and in recognition of the current practices and resources made available by AusAID for M&E. These standards are not developed for use in long-term or resource-intensive evaluations. Although these standards are applicable, AusAID would expect a higher level of performance against many of these standards. This institutional change program has taken current practices in program delivery as the starting point so that change can be incremental and institutionalised before higher standards can be resourced and/or expected. These standards are, however, fully consistent with international standards.

The standards have been developed by the M&E Help Desk Consultant over a three year period and are based on a careful assessment of AusAID 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 AusAID staff in the *purchasing and oversight* of M&E systems and deliverables;
- the tasks carried out by contractors and consultants who *design, implement and produce* M&E systems and deliverables;
- the capacity of AusAID staff to absorb new information and perform new tasks within their already busy work demands; and
- the expected end-of-program outcomes of the Evaluation Capacity Building program itself.

What Standards Have Been Developed?

The standards presented in this series are:

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

Standard 2: Initiative M&E Systems

Standard 3: Initiative Progress Reporting

Standard 4: Terms of Reference for Independent Evaluations

Standard 5: Independent Evaluation Plan (Methodology)

Standard 6: Independent Evaluation Report

Standard 7: Initiative Monitoring Visit

In the standards that follow, a summary table or single-page checklist is followed by an elaboration of the standards on the checklist to provide more clarity.

What is the Difference Between Guidelines and Standards?

There could easily be confusion between a standard and a guideline such as AusAID “Rules and Tools” or AusGUIDE. In general, guidance provided corporately focuses on *processes*. The standards referred to in this institutional change program focus on the *quality of the M&E deliverable and utilisation of the findings*. As such these standards do not duplicate corporate work, rather provide an additional resource to assist in the *articulation and measurement of quality*. They are expected to be shared with consultants and contractors working with AusAID.

How will these Standards be Trialled?

The standards will be tested through the conduct and analysis of at least one case study per standard. Under the ECB program, case studies are conducted for each M&E process and related product such as an Independent Evaluation Report. The intention is to study what resources and processes are required to bring about the required standard, and what are the constraints or facilitating factors that account for the standard achieved.

After the completion of the case study, the standard will be systematically reviewed for:

- Appropriateness and completeness of the technical content;
- The extent to which each standard remains relevant to users;
- Usefulness from a learning perspective; and
- Clarity of language in the context of English as a second language for many users.

More than one case study may be conducted for each standard. After each case study is conducted the standards will be reviewed over the life of the ECB program.

Note: In this document the term “end-of-program outcome” is used to refer to end-of-initiative, facility, project or other delivery mechanism outcomes.

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

Trialled in AusAID Indonesia Country Program

	Feature of the Design	Further Guidance Available
Initiative Design Planning		
1.1	Terms of reference reflect the requirements for the M&E Specialist/Designer	Example TOR for M&E Specialist on Design Mission
1.2	M&E Specialist/Designer has demonstrated skills in logic modeling and theory of change	Qualification requirements for M&E Specialist on Design Mission
1.3	Design plan allows for sufficient opportunity to gather relevant contextual information to inform the design	
1.4	Design plan allows for sufficient opportunity to meaningfully engage national partners in the design process	
Design Features		
Program Logic		
1.5	The goal (beyond the life of the initiative) is aligned with relevant Country Program Strategy Objectives, national development priorities, and/or, MDGs.	
1.6	The proposed mechanism by which the entire initiative and the broader goals are linked is made explicit (theory of change)	
1.7	The target groups, beneficiaries or individuals expected to change any behaviours for key outcomes (and related interventions) are identified	
1.8	The end-of-program outcomes are expressed in terms of performance outcomes not open-ended capacity outcomes	
1.9	The end-of-program outcomes are pitched at the correct level for the time, effort and resources applied	
1.10	Immediate and intermediate outcomes are described for key interventions	
1.11	There is a robust logic and theory of change that links the interventions with immediate, intermediate, and end-of-program outcomes	
1.12	The extent to which the context will likely influence the initiative outcomes is described	
Enabling Quality M&E Systems (these features are expected to be carried over into any RFT, Contract or MOU)		
1.13	The design articulates that international standards for evaluation are expected to be met for the design, implementation and products of the M&E system.	http://www.jcsee.org/program-evaluation-standards
1.14	The design articulates the importance of partner government mutual accountability by making provision for the mutual assessment of progress	
1.15	For initiatives over \$3m, adequate provisions have been made for design and oversight of M&E systems for the life of the initiative by an adequately qualified M&E Specialist.	Example TOR and qualifications for Initiative M&E Specialist
1.16	Important evaluation questions emerging during the design are highlighted	
1.17	Adequate resources are available for M&E- approximately 5-7% of total budget	

Detailed Description for Standards for Design

Initiative Design Planning

1.1. Terms of Reference reflect the requirements for the M&E Specialist/Designer

Terms of reference for the M&E Specialist/designer for the design mission adequately articulate the required tasks to meet the following standards. The terms of reference also indicate the qualifications and experience required of the individual responsible for the development of the program logic.

1.2. M&E Specialist/Designer has demonstrated skills in logic modelling and theory of change

The individual recruited to develop, or facilitate the development of the initiative program logic has demonstrated experience in logic modelling and/or theory of change. They can provide evidence of program logic experience if requested. If an AusAID member of staff has been tasked to design the initiative 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 contextual information to inform the design

Contextual analyses are carried out in advance of the final design process to ensure that the design logic is feasible and addresses the right development issues and priorities. Adequate provisions are made to inform the design of: a) current status of developments in the sector, and any relevant existing or emerging issues; b) how the initiative expects to contribute to the broader development needs (at the level of the goal statement); c) how the institutional context informs the selection of partners and interventions; and d) how the context is likely to impact on the likelihood of achieving the end-of-program outcomes (including identification of risks and their management).

1.4. The design plan allows sufficient opportunity to meaningfully engage 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 enable full 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.

Program Logic

1.5. The goal (contribution to outcomes beyond the life of the initiative) is aligned with relevant Country Program Strategy Objectives, national development priorities, and/or, MDGs.

The goal for the initiative is directly aligned with AusAID Country Program Strategy (CPS), national partner development priorities, and/or the Millennium Development Goals. For AusAID CPS, the goal is best situated at the level of the outcomes articulated in the sectoral delivery strategies that support the Country Program Strategy.

1.6. The proposed mechanism by which the entire initiative and the broader goals are linked is made explicit (theory of change)

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 result in 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.

1.7. The target groups, beneficiaries or individuals expected to change any behaviours for key outcomes (and related interventions) are identified

Where appropriate, key actors in the initiative that are expected to adopt new behaviours or practices are clearly identified. The expected changes in behaviour or practices by the last day of the initiative are described for each key actor.

1.8. The end-of-program outcomes are expressed in terms of performance outcomes not open-ended capacity 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 initiative. Open-ended capacity outcomes are not used. 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 contextual analyses early in the design; and b) there are clear plans for systematically gathering the required information, and a future date planned for further development of the outcome statements.

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 initiative.

1.10. Immediate and intermediate outcomes are described for key interventions

Immediate and intermediate outcomes are described. 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. 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 theory of change is fully described.

1.11. There is a robust logic and theory of change that links the interventions with immediate, intermediate, and end-of-program outcomes

The design explains how the interventions, immediate and intermediate outcomes are expected to be achieved¹. 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 of change or a lack of consensus on how interventions are expected to bring about the required change, then these are highlighted and brought into considerations for M&E during the life of the initiative.

¹ It is beyond the scope of these standards to explain logic modelling and/or theory of change. It is expected that the authors of designs have this expertise.

1.12. The extent to which the context will likely influence the initiative outcomes is described

The relevant contextual analyses conducted before or during the design are fully explored. See Standard 1.3 for more details.

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

1.13. The design articulates that international standards for evaluation are expected to be met for the design, implementation and products of the M&E system.

The design document makes it clear that international standards for M&E are expected to be met. The most universally accepted international standards can be found at: <http://www.jcsee.org/program-evaluation-standards>. In addition any prevailing AusAID rules and tools, policies or other standards are reflected in the requirements for M&E.

1.14. The design articulates the importance of partner government mutual accountability by making provision for the mutual assessment of progress

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 government systems, but where this is not available or appropriate, interventions are posed that enable national partners to participate in the generation and utilisation of findings and recommendations relevant for their needs. Mutual accountability for a program 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.

If capacity building in M&E is proposed for national partners then this is integrated into the design as an intervention, and not an add-on to the M&E system.

1.15. For initiatives over \$3m, adequate provisions have been made for design and oversight of M&E systems for the life of the initiative by an adequately qualified M&E Specialist.

Resources have been made available for a short term M&E Specialist to design and provide oversight of initiative M&E systems for the life of the initiative. A single task of designing an M&E system without on-going technical support is not proposed. Regular visits (at least twice a year) have been resourced. For larger initiatives (over \$10m), a terms of reference for the M&E Specialist has been attached to the design document and requested to be a named individual during the RFT process. Qualification and professional experience requirements for this specialist reflect the magnitude or sensitivity of the initiative. (Note: sample TORs for a Significant Initiative M&E Specialist are available).

1.16. Important evaluation questions emerging during the design are highlighted

Important evaluation questions that emerged during the design process are highlighted for 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 theory of change that require exploration during implementation. Evaluation questions have clear relevance to improvement of initiative implementation, or of interest to a wider sectoral, national partner or AusAID Country Program audience.

1.17. Adequate resources are available for M&E- approximately 5-7% of total initiative budget

For small initiatives this is likely to be a higher proportion and a lower proportion for larger initiatives over approximately \$25m. Consider also IPR, ICR, Technical oversight group and Activity manager monitoring visits.

Standard 2: Initiative Monitoring and Evaluation Systems

Trialled in AusAID Indonesia Country Program

Features of the M&E System		Further Guidance Available
Design Phase to Initiative Mobilisation		
2.1	Adequate resources available 5-7% of total initiatives budget (consider also IPR, ICR, Technical Oversight Group, Activity Manager monitoring Visits)	
2.2	TORs for the M&E Specialist adequately reflect the key tasks of system design, oversight and outcome evaluation with adequate days allocated for the life of the initiative. Adherence to international evaluation standards has been required.	Example TOR
2.3	Adequately qualified M&E Specialist is recruited	Example TOR
M&E System Planning (in time to inform the first annual plan)		
2.4	There is an adequate basis for the development of the M&E Plan (e.g. Evaluability Assessment)	Full description of expected content
2.5	The M&E Plan provides a summary of the overarching system design including key M&E approaches and activities	
2.6	The M&E Plan reflects current international standards for evaluation practice (e.g. Joint Committee on Standards for Educational Evaluation – Program Evaluation Standards)	http://www.jcsee.org/program-evaluation-standards
2.7	End-of-program outcomes are clearly articulated.	
2.8	Key evaluation questions are posed such as exploratory questions that answer important questions about the program.	
2.9	The quality and reach/coverage of key deliverables are monitored and evaluated	
2.10	The methodologies are fully elaborated.	
2.11	Baselines are considered, and if not appropriate then a rationale is provided	
2.12	Responsibility is allocated to specific individuals (not organizations) for all M&E activities	
2.13	Mutual accountability and joint assessment by national partners is provided for	
2.14	Individuals responsible for implementing the M&E plan have the capacity to do so (time, resources and skills)	
2.15	A complete schedule of M&E activities shows when all key M&E activities will be carried out and information available	
2.16	The M&E plan can be easily understood by the non-specialists and key stakeholders	
2.17	The M&E plan is resourced and costed	
During Implementation (Note: see more detailed standard for progress reporting)		
2.18	M&E reports are available upon request and report against planned M&E activities a) Summarizing information collected, analysed and interpreted in the reporting period b) For all major evaluation activities (e.g. surveys, studies)	
3.19	Progress reports meet stakeholder needs, report against M&E plan, have a credible basis for claims, and recommend actions to improve performance.	Guidance to Implementation Partners
3.20	Information has informed initiative learning, decision making and actions	

Detailed Description of Standards for Initiative M&E Systems

Design Phase to Mobilisation

2.1. Adequate resources are available

Between 5 and 7% of the total initiative budget has been allocated to monitoring and evaluation. This includes considerations of IPR, ICR, any technical oversight group, and Activity Manager monitoring visits. For small initiatives this is likely to be a higher proportion and a lower proportion for larger initiatives over approximately \$25m. This allocation remains intact through the RFT and contracting phase.

2.2. Terms of reference for the M&E Specialist reflect requirements of 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 complex M&E tasks relating to outcome evaluation. The TORs also state the expectation that the standards of M&E will reflect international evaluation practice standards such as those described in the Joint Committee for Standards of Educational Evaluation – Program Evaluation Standards. An example TOR for a typical initiative is available on request. The TORs also describe the required standards of qualifications and experience for the position. Example of suitable qualifications and experience are included in the example TOR.

2.3. An adequately qualified M&E Specialist has been recruited

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

Monitoring and Evaluation System Planning

2.4. There is an adequate basis for planning the M&E system

The M&E Specialist is provided with the time and resources to conduct an Evaluability Assessment (or similar exercise). It typically takes 10 days to conduct this assessment. The M&E Specialist has made the full report of the Evaluability Assessment available. The Evaluability addresses such areas as:

- a) Consultation 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) Consultation 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) Identify 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, 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 (QAI) Reporting and Annual Program Performance Reporting. There should be a clear recognition of how the M&E plan is expected to provide evidence for reporting against the Country Program Performance Assessment Framework.
- h) A review of cross-cutting policy areas that will need to be included in the M&E plan;

- i) Clear identification of issues and/or constraints that will affect the design of the M&E plan.

2.5. 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 system. This describes the key M&E approaches and activities. The reader should be able to quickly see the broad design features, such as what the major evaluations are, and a broad description of the associated methods. The 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).

2.6. The plan reflects international standards for program evaluation

The M&E design and plan reflects international standards for program evaluation. For example, those described in the Joint Committee for Educational Evaluation – Program Evaluation Standards. If adherence to this standard is being formally assessed, it should be done so by a suitably qualified evaluator.

2.7. End-of-Program outcomes are clearly articulated.

Early in the M&E plan, end-of-program outcomes are clearly articulated. They 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 program and reflect good program logic. If intermediate outcomes are identified, then these are presented in a logical sequence from immediate to intermediate, to end-of-program outcome (if this is appropriate). If the Evaluability Assessment shows that end-of-program outcomes are not well set in the design, then these have been adjusted to be more appropriate for the program intent.

2.8. Key evaluation questions are posed

The evaluator has the freedom to pose whatever evaluation questions seem appropriate after the Evaluability Assessment. At a minimum, there are questions relating to important contextual (and risk) information, progress toward intermediate and end-of-program outcomes (including cross-cutting outcomes), and exploratory questions about the program, especially questions that assess or test the proposed theory of change. These questions have a related method to ensure that they are answered.

The list of questions should be as focussed as possible to allow primary information users to make appropriate decisions about the program, and future programming in the country program, or inform corporate decisions where this need has been made explicit by AusAID. It is readily apparent how information generated from these questions will be utilised.

If only indicators are presented (did we achieve our expected intermediate or end-of-program outcomes), consider whether or not these will provide sufficient information for program improvement.

2.9. The quality and reach/coverage of key deliverables are monitored and evaluated

The quality of key program deliverables is monitored and evaluated. Key program deliverables are defined as those accounting for large effort and/or expenditure, those who are likely to have a major influence on the extent to which the program 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 for quality. The coverage and reach is monitored to allow output reporting on beneficiaries exposed to the program.

2.10. The methodologies are fully elaborated

Proposed methods are fully described. Major studies are fully designed, or provisions made for the future design of major studies by the M&E Specialist at a future date. Implementation teams are not expected to design evaluation studies. Methods are fully described for sampling, data collection, and data analysis or processing. Tools should be annexed if appropriate. 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 during the Evaluability Assessment.

[Note: If this standard is being formally appraised, it should be done by a suitably qualified evaluator].

2.11. Baselines are considered

A baseline is considered in the design of the M&E system. A baseline is not a broad needs assessment; rather a baseline is tightly focused on the actual interventions of the program. As a baseline is not always appropriate, a full rationale is provided where a baseline has not been included in the system design.

2.12. Responsibility is allocated to specific individuals (not organizations) for all M&E activities

For all tasks required to implement the M&E system, the responsible individuals are identified. This is not naming organizations, 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.

2.13 Mutual accountability and joint assessment by national partners is provided for

Mutual accountability by national partners is a feature of the M&E system. Joint assessments are planned. Mechanisms for engaging national partners for mutual accountability are explained. If this is not considered appropriate, a full description of the rationale is provided.

2.14. 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 Specialist is not expected to 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 M&E system.

2.15. 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 25th of each month; or June annually). The schedule makes it explicit when information will be available to primary users and is compatible with national reporting or QAI reporting cycles.

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

The 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 in developing an accessible plan.

2.17. The M&E Plan is resourced and costed

The M&E plan has been fully resourced and costed. For resourcing, the number of person days (by position) is described, personnel associated costs, M&E materials, software or hardware required, any respondent costs, and all other resources such as transport are described. Simply stating the person days of the M&E Specialist is unlikely to be sufficient resources to complete all tasks to an adequate standard. A costing of the resource schedule should be presented. See 4.1 for adequate resourcing standards.

During Implementation (see also Standard3: Progress Reporting)

2.18. M&E reports 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 AusAID).

2.19. 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 their expected information needs. They are not too long, and they are written for managers and non-specialists. Progress reports provide information promised in the M&E plan, they include a credible basis for claims made, and they include recommended

actions for program improvement. The latter include expected dates for management responses to be completed. Refer to further guidance on recommended content for implementation partner progress reports.

2.20. The M&E system has informed learning, decision-making and action

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

DRAFT

Standard 3: Initiative Progress Reporting

Trialled in AusAID Indonesia Country Program

	Feature of Progress Report	Further Guidance Available
3.1	There is an executive summary that communicates the key information required for QAI reporting.	
3.2	There is a reflection on the continuing relevance of the expected end-of-program outcomes	
3.3	An assessment of the adequacy of progress toward sustained end-of-program outcomes is described	
3.4	Factors that have accounted for the adequacy of progress toward sustained end-of-program outcomes are explored	
3.5	The implications of factors identified are discussed with respect to the likely achievement of end-of-program outcomes	
3.6	Management responses taken to date or proposed for the future are described in terms of the likely achievement of end-of-program outcomes	
3.7	The quality, reach and coverage of key outputs or deliverables for the reporting period are described	
3.8	The adequacy of progress against the annual plan is described	
3.9	A reflection on the adequacy of planned inputs to meet the expected end-of-program outcomes is provided	
3.10	The adequacy of progress against the budget is assessed	
3.11	Factors that have accounted for the progress against the annual plan and budget are explored	
3.12	The implications of factors identified are discussed with respect to the likely achievement of outputs against the annual plan and budget	
3.13	Management responses to issues taken to date or proposed for the future are described in terms of progress against the annual plan and budget	
3.14	The efficiency and effectiveness of key management or implementation systems is assessed or demonstrated	
3.15	Factors that have accounted for the strengths or weaknesses in management or implementation systems is described	
3.16	The implications of management systems factors identified are described	
3.17	Management responses taken to date or proposed for the future are described in terms of management or implementation systems	
3.18	The relevant aspects of the context are adequately described	
3.19	The report achieves a fair balance between reporting of positive or negative issues or achievements	
3.20	The report provides credible evidence of claims made	
3.21	Important lessons are summarized	

Detailed Description of Standards for Initiative Progress Reporting

Further Guidance Available: "Suggested Report Content for Implementation Partners"

3.1. There is an executive summary that communicates the key information required for QAI reporting

An executive summary provides easy access to information to inform the Quality at Implementation (QAI) report. 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 gender outcomes); 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 key management responses proposed. Any issues with the management or implementation systems are described.

3.2. There is a reflection on the continuing relevance of the expected end-of-program outcomes

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. A reflection of the continuing relevance of these current expected end-of-program outcomes is provided with supportive rationale for any proposed changes.

3.3. 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 program. A firm statement of the adequacy of progress toward these outcomes is provided. A simple list of immediate and/or intermediate outcomes or achievements is not sufficient. A judgement of adequacy of this progress must be explicit.

3.4. Factors that have accounted for the adequacy of progress toward sustained end-of-program outcomes are explored

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.

3.5. The implications of factors identified are discussed with respect to the likely achievement of end-of-program outcomes

Particularly where problems or challenges have been identified, there is a full analysis of the implications the situations will have on the likely achievement of end-of-program outcomes. It is clear from the report the extent to which AusAID 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.

3.6. Management responses taken to date or proposed for the future are described in terms of the likely achievement 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 stakeholders are required to take what actions. The cost implications of these responses are discussed.

3.7. The quality, reach and coverage of 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 geographical coverage of outputs is provided, the reach or number of outputs or beneficiaries of those outputs enumerated, and the quality of key outputs is assessed. An assessment of the quality of minor outputs is not required.

3.8. The adequacy of progress against the annual plan is described

Self explanatory

3.9. A reflection on the adequacy of planned inputs to meet the expected end-of-program outcomes is provided

The adequacy of the planned inputs (and outputs) to meet the expected end-of-program outcomes is assessed. Where there are anticipated shortfalls, the report provides a well supported argument for any proposals for additional inputs.

3.10. 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.

3.11. Factors that have accounted for the progress against the annual plan and budget are explored

For any issues identified in 3.8 and 3.10 above, a full exploration of the supportive and inhibiting factors that account for this situation is provided.

3.12. The implications of issues identified are discussed with respect to the likely achievement of outputs against the annual plan and budget

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 AusAID and other stakeholders will need to track the relevant issues.

3.13. Management responses to issues taken to date or proposed for the future are described in terms of progress against the annual plan and budget

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 stakeholders are required to take what actions. The cost implications of these responses are discussed.

3.14. The efficiency and effectiveness of key management or implementation systems is assessed or demonstrated

Key management or implementation systems are described where there are issues to consider. This includes: a) strategic and annual planning; b) governance and oversight c) monitoring, evaluation and risk management; d) financial management and value for money; and e) staffing or human resource systems.

3.15. Factors that have accounted for the strengths or weaknesses in management or implementation systems is described

For any issues identified 3.14 above, a full exploration of the supportive and inhibiting factors that account for the situation is provided.

3.16. The implications of management or implementation system factors identified are described

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

3.17. Management responses taken to date or proposed for the future are described in terms of management or implementation systems

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 stakeholders are required to take what actions. The cost implications of these responses are discussed.

3.18. The relevant aspects of the context are adequately described

The relevant aspects of the context are described. Information is provided on: the progress of the broader national sector or sub-sector in relation to the program area of work; any relevant areas in the context which impact on the relevant needs and priorities of the program actors; any factors which are impacting on the achievement of sustained outcomes or quality, reach and coverage of the program deliverables (discussed above under factors); extent to which the program may be contributing to changes in the sector. Findings on key risks being monitored are presented and continuing risk assessed.

3.19. The report achieves a fair balance between reporting of positive or negative issues or achievements

The report appears to be 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 complexity of human development.

3.20. The report provides credible evidence of claims made

For major claims of achievement (both in terms of achievement of outcomes, and quality of outputs or deliverables), credible supportive evidence is provided. This may not always be methodologically defensible evidence, but 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)..

3.21. Important lessons are summarized

Only important lessons are summarized. These have relevance to the on-going implementation of the program; national partner development strategies; the AusAID sectoral delivery strategy; the AusAID Country Program, or AusAID corporately. Minor, generic development lessons are not listed.

Standard 4: Terms of Reference for Independent Evaluations

Trialled in AusAID Indonesia Country Program

No.	Standard	Further Guidance Available
Background and Orientation to the Evaluation		
4.1	A brief orientation to the initiative is provided which includes: total value; time frame; expected end-of-program outcomes; a short summary of the key approaches employed in the initiative; a brief explanation of the initiative's expected contribution to the Country Program Strategy and national development planning.	
4.2	The rationale for the request for the evaluation: including the overall purpose and primary users of the information are described	
4.3	Any key management decisions that will be addressed as a result of the evaluation are explained	
4.4	The key issues that lead to the evaluation questions are described in neutral language	
Evaluation Questions/Scope		
4.5	The DAC criteria are evident in the evaluation questions where appropriate	
4.6	The questions are consistent with the overall purpose of the evaluation	
4.7	It is clear which questions are considered to be the priority questions for AusAID that must be answered with a good degree of rigor	
4.8	The scope of the questions is suitable for the time and resources available to the team	
4.9	There is adequate specificity in the questions to guide the development of evaluation plan	
Evaluation Process		
4.10	A verbal briefing of the key issues and priority information is planned	
4.11	Adequate time has been allocated to document reviewing and document <i>appraisal</i>	
4.12	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	Evaluation Plan Standards
4.13	Proposed scheduling allows for adequate data collection opportunities to meet the evaluation questions, but doesn't distract the team from priority information requirements	
4.14	A feed-back session to relevant information users are planned 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 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 AusAID to track progress of the evaluation without distracting the team from carrying out important evaluation activities	

Detailed Description of Standards for Evaluation Terms of Reference

Background and Orientation to the Evaluation

4.1. A brief orientation to the initiative is provided

As the terms of reference are used to explore with proposed consultants whether or not they are interested in, or to comment on the proposed evaluation, 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; a short summary of the key approaches employed (such as training, technical advisers, secondments, provision of infrastructure, equipment and so on). The context in which the initiative is situated is described such as the country program strategy that the initiative aims to address as well as the national development plans of relevance. The delivery mechanism is described (contracted, multi-lateral development partner, NGO) and whether or not the initiative is a project, program or facility. Any information which can guide the reader in quickly understanding the scope/reach of the program is provided.

4.2. The purpose of the evaluation is described

Although Independent Progress Reports and Completion Reports are generally addressing accountability, program improvement (or future programming improvement) and/or knowledge development, the TOR should clearly identify the overall purpose(s) and show which purposes are of most importance. 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, "AusAID" is made up of senior executive, desk officers, senior post managers and posted activity managers. "The Contractor" is made up of head office personnel and implementation managers and advisers.

Although the *audience* is a wider group of stakeholders than the *primary users*, if the report is to be published on the AusAID web-site then this is indicated – in this case there is clear instructions on how sensitive information is to be communicated.

4.3. The key management decisions are identified

Any important management decisions that the primary users are expected to make are identified and described.

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.

Evaluation Questions or Scope

4.5. The DAC criteria are evident in the evaluation questions where appropriate

Although the DAC criteria do not have to be mechanically replicated, due consideration has been given to the DAC criteria (relevance, effectiveness, efficiency, sustainability). If the evaluation is designed to address a specific question, then a brief rationale for why the DAC criteria are not reflected in the TORs is provided.

4.6. Evaluation questions are consistent with the overall purpose

Each of the evaluation questions is clearly related to the stated purpose(s) of the evaluation. There are an adequate range of questions to meet all the stated purposes. There are no additional questions unrelated to the stated purpose.

4.7. 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. These priorities are consistent with the overall purpose of the evaluation.

4.8. 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 large number of sub-questions which show what specific information will be required to answer the broader questions. In addition, for a typical interview with a respondent without the need for translation, only four or five topics can be addressed with depth. 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.

4.9. There is adequate specificity in the questions to guide the development of the evaluation plan

Evaluation questions are not broad or vague or open to a wide range of interpretations. The evaluator needs to break down these broad 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. It is appropriate to provide supportive information with questions to guide interpretation.

Evaluation Process

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

4.10. 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 (up to half day)

*4.11. Adequate time has been allocated to document **reviewing** and document **appraisal***

Time has been allocated to *reviewing* initiative documentation (approx 2 days) as well as time to *appraise* any key documents such as strategies or the M&E system plan (often a day per document for full appraisal).

4.12. 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 an IPR two days is required to develop an evaluation plan which includes a fully elaborated methodology.

4.13. Proposed scheduling allows for adequate data collection opportunities to meet the evaluation questions, but doesn't distract the team from priority information requirements

The proposed schedule in the TOR is not too detailed as this is developed after the evaluation plan identifies suitable respondents 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 are planned together or separately depending on the sensitivity of findings (e.g Aide Memoire, discussion or presentation)

There is adequate time to provide adequate detail in evaluation findings to allow contestability of those findings, and 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.

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 has been given adequate time to process information from interviews, document reviews and appraisals, observations etc. to provide a credible evidence base to support findings. Typically, processing of data for a 12 day in-country mission would take about three days. This is additional time to actual report writing.

4.16 Adequate time is made available to complete the draft report

The number of days allocated to completing to 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 on the final draft. It is recommended to allow the evaluation team sufficient time to rest after the mission and reflect on the mission. For example, five days allocated to report writing could require a two 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 final report

The process for commenting on the draft report is described and is efficient. Only those relevant individuals are invited to comment, and the focus of their comments is identified. 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 AusAID will either provide comments in a consolidated form to the evaluation team, or, allows additional time to respond to a large number of comments from all stakeholders.

The comment processes also ensures that the Independent evaluation team are allowed to make final decisions on what to include in the report. It may be necessary to provide comments from different stakeholders (national partner, implementation team and AusAID) 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 range of comments generated; and b) the extent to which these comments require significant structural change in the final report.

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

Although it will be up to the team leader to provide detailed direction on tasks in the evaluation plan, the TORs show AusAID 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 AusAID expects international and national consultants to participate and assume responsibility for any particular tasks.

4.20. Skill sets of evaluation team reflect priority questions

Unless there is a compelling reason provided, the team leader of the evaluation is the monitoring and evaluation specialist. 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, and as such is the Monitoring and Evaluation specialist. 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 program management systems, then the number of days allocated to technical specialists from other sectoral areas reflects this focus.

4.21. The reporting requirements allow AusAID 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 with allowing the team to focus on important evaluation activities. The evaluation plan is a critical document for AusAID to ensure that the evaluator has correctly interpreted the TORs and has made suitable plans to conduct the evaluation to meet the TORs. 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 the evidence base of processed data. The TORs indicate that where the findings are unexpected, or there are significant challenges from key stakeholders, the processed data (evidence base) will be made available. The evidence base may be requested for other reasons. *Note: If additional days are not allocated to processing data, then this requirement is not appropriate.*

Standard 5: Independent Evaluation Plans

Trialled in AusAID Indonesia Country Program

No.	Standard	Further Guidance Available
5.1	The evaluation design 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	Limitations or constraints on the evaluation are described (e.g. time frame; resources; available data; political sensitivities).	
5.4	The purpose and/or objectives of the evaluation are stated.	
5.5	A broad investigatory framework is provided to orient the reader to the overall evaluation design.	
5.6	More detailed evaluation questions are posed. These are based on the terms of reference, but provide AusAID with greater clarity in how the terms of reference a) have been interpreted; and b) will be met.	
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	The design is flexible enough to allow unexpected issues to 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 approach to data processing is described and is consistent with the time and resources available.	
5.14	Ethical considerations have been addressed where relevant (e.g. privacy and confidentiality).	
5.15	It is clear who will be making the judgments.	
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).	

Detailed Description of Standards for Evaluation Plans

5.1. Evaluation design is based on a collaborative approach

The evaluator has consulted AusAID, and stakeholders identified as important by AusAID, 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 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 vague group of people that may be interested in, or may be affected by any decisions that result from the evaluation.

5.3. 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 AusAID 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. Political sensitivities are highlighted where appropriate.

5.4. The purpose and/or objectives of the evaluation are stated.

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

5.5. A broad investigatory framework 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. The evaluation plan does not go straight into detailed descriptions of methodology without this general orientation.

5.6. Detailed evaluation questions are posed

Although the terms of reference is where AusAID 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 questions are generated. Because many of the questions presented in a terms of reference are usually very broad, the more detailed questions allow AusAID 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 AusAID evaluation manager to see the implications of the scope of the evaluation described in the terms of reference. This breakdown of questions allows the reader to assess whether or not the original scope was realistic. The evaluation manager needs to pay careful attention to this aspect of the evaluation plan.

5.7. High priority questions are identified

AusAID evaluations often have a very large number of evaluation questions that cover a very wide number of aspects of the program to be evaluated. Some of these questions will be more important than others. The evaluation plan reflects where the emphasis will be placed. 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. Flexible design to allow unexpected issues to emerge

This flexibility is usually built in to the questioning technique employed during an interview. However, the evaluation plan provides for sufficient time during individual interviews and across the schedule as a whole to allow new issues to emerge and be responded to if they are important.

5.9. Methods for each evaluation questions are described

The 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 AusAID reviews 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 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. The design of major evaluation activities/studies should be annexed and include tools such as interview guides or questionnaires. If these are to be developed later, then a due date for completion of the design, and competent persons responsible for designing them are identified.

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 program'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 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 AusAID 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 use a number of methods to examine the same issue.

5.12. Sampling strategy is clear and appropriate

Most evaluations will require some sort of sampling strategy. For example, if case studies are going to be used to investigate certain aspects of a program in more detail, then the basis for the selection needs to be transparent. For most reviews and evaluations that employ qualitative methods then purposeful sampling is a common approach. In this case the criteria for selection of the cases should be described.

It is clear that the criteria are suitable for the evaluation questions posed. Sampling may also be important where there are a large number of respondents that could be interviewed such as beneficiaries. For quantitative studies that are attempting to generalize findings across a wider population of activities or populations (e.g. household survey), random sampling strategies would be appropriate – these surveys are not often resourced for typical AusAID reviews and initiative level evaluations.

5.13. Data processing is described

A common problem with AusAID reviews and evaluations is that data is often collected, but not systematically processed. This does not mean to infer that complex analytical techniques will be required, rather that the findings in evaluation reports are based on a systematic treatment of the information collected. The evaluation plan describes how the evaluator plans to process the data (it is fully transparent) - no matter how simple the procedures.

5.14. Ethical considerations have been described

The evaluation design identifies where there are any special ethical considerations. For most of the evaluations and reviews conducted by AusAID, this will mostly be around privacy and confidentiality issues. There is no straight forward answer to what is most appropriate. In some cases where information is not likely to be especially sensitive, it will be sufficient to simply tell the respondent or persons involved that the information will be available to relevant stakeholders as a basis for the report findings. However it is not acceptable to publish processed data identifying individuals or organizations where the respondent was not made aware of this. In cases where the evaluator has informed the respondent, but the information is still considered to be sensitive, or could cause difficulties for others, then the information should not be made available unless there could be no identification with the respondent. Any potential untoward outcomes should be anticipated and managed for in the evaluation design. Ethical considerations are not a sufficient reason for an evaluator to not make the evidentiary basis of an evaluation available to AusAID as the commissioner of the evaluation. In cases of sensitivity, privacy or confidentiality can be preserved, while still making the evidentiary basis transparent.

5.15. Who will be making the judgments is clear

Often there is confusion about who will be making judgments about a program. This has been communicated to the evaluator and is reflected in the evaluation design. In some cases AusAID may want an evaluator to report neutrally on facts and leave AusAID to make the final judgments. More commonly, AusAID would require the evaluator to make a professional judgment based on the evidence gathered. 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.

5.17. Scheduling guidance is provided

The schedule is developed by AusAID **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. Check to see that enough time is available to meet with the implementation team – this usually requires a lot longer than what is allocated. As part of reviewing the methodology AusAID negotiates the proposed list of respondents before final scheduling. The evaluator scheduling guidance is realistic for the time frame.

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.

Standard 6: Independent Evaluation Reports

Trialled in AusAID Indonesia Country Program

No.	Standard	Further Guidance Available
Introductions		
6.1	A background to the evaluation summarizes: the total value of the initiative; the number of years of the initiative; the stage of initiative delivery; key outcomes of the program; and the key issues identified in the terms of reference	
6.2	A brief summary of the methodology employed is provided	
6.3	Key limitations of the methodology 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.	
Findings and Analysis		
6.5	The evaluation report clearly addresses all questions in the Terms of Reference	
6.6	There is a full description of each of the issues identified so that the reader feels they have been given the full picture	
6.7	The relative importance of the issues communicated is clear to the reader	
6.8	There is a good balance between operational and strategic issues	
6.9	The text clearly establishes that the evidence supports the arguments posed	
6.10	Alternative points of view are considered where appropriate	
6.11	Complex issues are fully explored and not oversimplified	
6.12	The role of the context in program performance is explored	
6.13	The text uses appropriate methods/language to convince the reader of the findings and conclusions	
6.14	There is an adequate exploration of the factors that have influenced the issues identified and conclusions drawn	
6.15	The implications of key findings are fully explored	
6.16	The overall position of the author is clear and their professional judgments are unambiguous.	
Conclusions and Recommendations		
6.17	The conclusions and recommendations logically flow from the presentation of findings and any associated analyses.	
6.18	Individuals have been allocated responsibility for responding to recommendations.	
6.19	Where there are significant cost implications of recommendations, these have been estimated (financial, human and materials costs).	
6.20	The recommendations are feasible	

Detailed Description of Standards for Evaluation Reports

Introductions

6.1. The background provides adequate information for individuals not familiar with the initiative

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

6.2. A brief summary of the methodology 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. For important evaluations the full methodology is annexed.

6.3. Key limitations of the methodology 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. 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 complex reports).

Findings and Analyses

6.5. The evaluation report clearly addresses all questions in the Terms of Reference

It should be relatively easy to identify where each of the questions in the terms of reference 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 terms of reference.

6.6. There is a full description of each of the issues identified

There is a full description of each of the issues identified so that the reader feels they have been given the full picture. The description of the issues allows the reader to understand the issues from a range of perspectives and not only the perspective of a single individual. The issues are described specifically, with no broad motherhood statements that make practical conclusions difficult to generate.

6.7. 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.

6.8. There is a good balance between operational and strategic issues

The report must address the full range of issues identified in response to the TOR and other critical issues that have emerged. There will be some technical, managerial or operational issues that are very important to consider and are often at the core of many important challenges. How the initiative is relating to its context, how the context is impacting on the initiative need to be reflected in the report. Strategic direction, approaches employed and progress toward end-of-program outcomes are important considerations. The strategic direction or any higher order issues of the initiative have been given adequate space, and minor technical issues are treated in a more limited fashion.

6.9. The text clearly establishes that the evidence supports the arguments posed

The author does not assume the reader accepts the evidence is of sufficient quality. For key findings, the basis of the findings and related conclusions is communicated clearly. The range or number informants that shared a view should be communicated – this does not have to be systematically counting the number of respondents, but the degree to which views are shared across respondents should be communicated in some way. The information is brought together from a range of sources, but communicated as a coherent whole. Evaluator opinions that are not supported by evidence or some basis are not the dominant communication style in the report.

6.10. Alternative points of view are 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, AusAID or other important stakeholder views.

6.11. Complex issues are fully explored and not oversimplified

The text acknowledges the complexity of issues if they are complex. The text does not present solutions to complex problems as simple, correct answers. For example, the findings must be presented fairly so that specific stakeholders are not held fully accountable for problems when multiple factors are involved. Human development is complex, and the report recognises that implementation teams and national partners are often working in complex environments and facing multiple challenges. This complexity must be recognised. Although the author takes a clear position, the reader is able to draw their own conclusions as the findings are presented comprehensively.

6.12. The role of the context in program performance is explored

The context can have two major effects on a program. The first is how the context has affected the achievement of outcomes (both supportive and inhibiting); and the second concerns the extent to which the program has had any effect on the context. Sufficient information is presented to allow the reader to understand the relationship between the initiative and its context. Evaluation reports do not describe initiatives as if they are isolated activities.

6.13. 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. As discussed above, the method 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 in the text.

6.14. 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 issues.

6.15. The implications of key findings are fully explored

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

6.16. 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 not required the evaluator to do this. The report does not simply state findings and expect AusAID to interpret them and draw their own conclusions. The report presents the authors view unambiguously. Has the initiative 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.

Conclusions and Recommendations

6.17. 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 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, findings presented, analysis conducted, interpretation carried out and reported. If questions in the methodology have not been addressed then an explanation has been given.

6.18. Individuals have been allocated responsibility for responding to recommendations.

Where appropriate, job titles, rather than organisations, have been allocated responsibility for all recommendations for action. If some recommendations are for broader partner government, or AusAID sectoral or corporate learning then these are identified separately.

6.19. 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.

DRAFT

Standard 7: Supervision Visits

Trialled in AusAID Indonesia Country Program

No.	Standards	Further Guidance
The Visit Plan		
7.1	The number of the monitoring visit for the initiative is recorded	Plan template
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 (interviews, observations, informal interaction, reading documents)	
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. The date sent to host is recorded.	
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	
Recording and Managing Information		
7.10	Key points or issues are recorded after each activity in line with visit questions (content and impressions of supervisor)	
Visit Report and QAI Update		
7.11	A background to the visit is presented which describes: a) the purpose of the visit; b) the visit number; c) the visit questions; d) the dates of the visit; e) the participating visit members (AusAID and joint visit with other partners); e) the cost (expenditure and person days)	Report template
7.12	A brief summary of findings for each visit question posed is provided. The four analytical questions are integrated into the questions where appropriate	
7.13	Important lessons learned or insights of general interest to the sector, the country program, or AusAID corporately are listed	
7.14	Information relevant to updating the QAI is presented in language suitable for pasting into a QAI report <i>[further discussion is required about the process for updating the QAI]</i>	
7.15	Management responses are summarised according to who is required to take responsibility for the action (Implementation Partner, Activity Manager, Unit Manager, or Counsellor). Management responses are given a time frame for action, steps required, and resource implications highlighted where relevant.	
7.16	The date the report is sent to Unit Manager for oversight of management responses is recorded.	
7.17	The date the report is sent to the Performance and Quality Unit for aggregation of findings and generation of lessons learned is recorded.	
7.18	Follow Up section is completed as each management response is completed. Report is closed when last management response is taken. If a management response will not be taken, then the reason is recorded.	

Detailed Description of Standards for Supervision Visits

Note: If several AusAID staff travel together on a Supervision Visit, it is only necessary to compile one, joint supervision 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.

The Visit Plan

The visit plan is sent to the Performance and Quality Unit on completion.

7.1. The number of the supervision visit is recorded

Over the life of each initiative there will be a number of visits. The visit plan shows what number of visits this is, and the number of months the initiative has been under implementation.

7.2. The broad purpose of the visit is described

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

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.

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. 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.

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. The proposed schedule of activities and/or meetings is presented

Once the ways of collecting information are fully described, a schedule of activities (such as observations of initiative 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 is described

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 activity 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 AusAID 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 AusAID staff member.

Visit Report and QAI Update

7.11. A background to the visit is presented

This describes: a) the purpose of the visit; b) the number of the supervision visit for the initiative; c) the visit questions; d) the dates of the visit; e) participating visit team members (this could be AusAID 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 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 initiative success; d) what are the responses that have been taken thus far (and their success), and what future responses are still required.

7.13. Important lessons learned or insights are listed

Lessons or insights that are important to the sector more broadly, the country program, or AusAID corporately are recorded. This section can be left blank if there is nothing of note.

7.14. Important information for QAI updates is recorded

This information is presented in very brief language in a format that can be cut and pasted into a QAI report. This information is arranged in accordance with QAI heading.

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, Activity Manager, Unit Manager, Counsellor, and/or National 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. The date the report is sent to the Unit Manager is recorded

The Unit Manager will need to be aware of the management responses that are generated from each initiative they are responsible for. They are responsible for supervising the follow-up required from each visit.

7.17. The date the report is sent to the Performance and Quality Unit is recorded

The report is sent to the P&Q Unit when it is completed. The P&Q Unit aggregates findings, lessons learned are generated, quality of the supervision visits are monitored, the impact on QAI reporting and management responses is assessed, and value for money of the visits considered.

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 P&Q Unit. If a management response is not taken then this is explained before closing the report.