



Australian Government

Department of Foreign Affairs and Trade

Australian
Aid 



EVALUATION OF DFAT INVESTMENT LEVEL MONITORING SYSTEMS

December 2018

OFFICE OF DEVELOPMENT EFFECTIVENESS

dfat.gov.au/ode

© Commonwealth of Australia 2018

ISBN: 978-0-6483308-0-6

Authors: Jim Fremming, Amy Gildea, Melissa Kamp and Neal Forster.

Published by the Department of Foreign Affairs and Trade, Canberra, 2018.

Disclaimer: The views contained in this report do not necessarily represent those of the Australian Government.

With the exception of the Commonwealth Coat of Arms and where otherwise noted all material presented in this document is provided under a Creative Commons Attribution 3.0 Australia (<http://creativecommons.org/licenses/by/3.0/au/>) licence. The details of the relevant licence conditions are available on the Creative Commons website (accessible using the links provided) as is the full legal code for the CC BY 3.0 AU licence (<http://creativecommons.org/licenses/by/3.0/au/legalcode>). The document must be attributed as Office of Development Effectiveness, *Evaluation of DFAT Investment Level Monitoring Systems: Evaluation Report*, Canberra, 2018.

This document is online at: dfat.gov.au/ode

For further information, contact:

Office of Development Effectiveness
Department of Foreign Affairs and Trade
GPO Box 887
Canberra ACT 2601
Phone: + 61 2 6178 4000
Facsimile: + 61 2 6178 6076
Internet: dfat.gov.au/ode

Cover photo: Delivering water, sanitation and hygiene facilities to communities in Fiji. Photo credit: Fiji Community Development Program, DFAT.

The Office of Development Effectiveness

The Office of Development Effectiveness (ODE) is an office within the Australian Government Department of Foreign Affairs and Trade that is independent of aid program management. ODE monitors the Australian aid program's performance and evaluates its impact. ODE's work is overseen by the Independent Evaluation Committee, an advisory body that provides independent expert advice on ODE's evaluation strategy, work plan, analysis and reports.

www.dfat.gov.au/ode

CONTENTS

| | |
|--|------|
| FOREWORD | V |
| ACKNOWLEDGEMENTS | VI |
| EXECUTIVE SUMMARY | VIII |
| MANAGEMENT RESPONSE | 1 |
| SUMMARY OF MANAGEMENT RESPONSE | 1 |
| RESPONSES TO EACH RECOMMENDATION | 2 |
| REPORT STRUCTURE | 9 |
| PURPOSE, SCOPE AND APPROACH TO THE EVALUATION | 10 |
| 1.1 BACKGROUND | 10 |
| 1.2 PURPOSE OF THE EVALUATION | 11 |
| 1.3 SCOPE OF THE EVALUATION | 12 |
| 1.4 APPROACH AND METHOD | 13 |
| 1.5 STRENGTH OF THE EVIDENCE BASE | 16 |
| CHARACTERISTICS OF BETTER PRACTICE INVESTMENT MONITORING | 17 |
| 2.1 WHAT IS BETTER PRACTICE? | 17 |
| 2.2 OUTCOMES-FOCUSED | 19 |
| 2.3 QUALITY ASSURED | 22 |
| 2.4 EFFECTIVELY USED | 23 |
| 3.1 WHAT ENABLES BETTER PRACTICE? | 27 |
| 3.2 DFAT'S PERFORMANCE CULTURE | 28 |
| 3.3 STRATEGY AND DESIGN | 33 |
| 3.4 THE ROLE OF DEMAND | 38 |
| 3.5 MANAGING CONTRACTORS | 44 |



CONTENTS

| | |
|--|----|
| ANNEX 1: METHODOLOGY | 47 |
| ANNEX 2: CRITERION BASED ASSESSMENT FRAMEWORK | 51 |
| ANNEX 3: TERMS OF REFERENCE | 54 |
| ATTACHMENTS | |
| Attachment A: Synthesis of the literature | |
| Attachment B: Tools for DFAT investment managers | |

FOREWORD

Australia's aid performance management system is recognised as among global best practice. This is a status, however, that cannot be taken for granted. Performance management systems require ongoing attention and nurturing if they are to continue to provide the foundations that are necessary to drive an effective aid program.

I have said before that improving the quality of DFAT's investment monitoring system is a continuous challenge. Reflecting this, the goal of this evaluation is to understand the determinants of better quality investment monitoring systems to assess Australia's aid performance. This is both timely and important. The enhanced expectations of the Australian aid program as a crucial element of Australia's engagement within the Indo-Pacific region means that much depends on it.

While the evaluation focuses on investments delivered by managing contractors—which make up about 20 per cent of total Australian aid delivery—the evaluation makes a convincing case that the findings have relevance for the entire system. The core finding of the evaluation is that DFAT's performance culture—the mix of shared vision, results expectations, operational tools and workplace behaviours that define and reinforce success—is the key underlying determinant of the quality of DFAT's investment monitoring

systems. In essence, the evaluation makes a compelling case that DFAT's investment monitoring systems are only as good as DFAT's management and staff.

Ongoing reforms within DFAT to both streamline and strengthen the aid management system will go some way to address the findings and I am pleased to see that the evaluation sets out a pathway to further strengthen and build on DFAT's current reform processes. Broad recognition across DFAT's senior management that a greater focus on monitoring will benefit the department is also key and I believe this evaluation provides clear evidence for this.



Jim Adams

Chair, Independent Evaluation Committee



ACKNOWLEDGEMENTS

This evaluation was undertaken by Jim Fremming, Amy Gildea and Melissa Kamp (Coffey International Development) with guidance and support from Neal Forster (Office of Development Effectiveness). The evaluation team would like to thank the many staff from both DFAT and managing contractors who committed their time and provided insights which informed the report. Thanks are also extended to others in the Office of Development Effectiveness, especially Jennifer Noble, Leo Carroll and Tracey McMartin.

ACRONYMS & ABBREVIATIONS

| | |
|--------|--|
| AIP-R | Australia-Indonesia Partnership for Rural Economic Development |
| ANCP | Australian NGO Cooperation Program |
| AQC | Aid Quality Check |
| CAVAC | Cambodia Agricultural Value Chain Program |
| CBAF | Criterion Based Assessment Framework |
| CSO | Civil Society Organisation |
| CVB | Contracting Services Branch |
| DCED | Donor Committee for Enterprise Development |
| DFAT | Department of Foreign Affairs and Trade |
| FCDP | Fiji Community Development Program |
| FHSSP | Fiji Health Sector Support Program |
| IEC | Independent Evaluation Committee |
| KOMPAK | Indonesia Governance for Growth |
| M&E | Monitoring and Evaluation |
| NGO | Non-Government Organisation |
| ODE | Office for Development Effectiveness |
| PNG | Papua New Guinea |
| PPD | Planning and Program Division |
| SGP | Strongim Gavman Program |
| ToR | Terms of Reference |



EXECUTIVE SUMMARY

Each year the Australian Government invests around \$4 billion to promote sustainable economic growth and poverty reduction in developing countries, as part of advancing Australia's interests internationally.

Currently, more than 2,000 staff at the Department of Foreign Affairs and Trade (DFAT)—across 24 posts and 15 divisions—manage, monitor and report on the expenditure of more than 800 investments.¹ Investments are delivered through country, regional, global and thematic aid programs, with a focus on the Indo-Pacific region.

The Australian Government's development policy, *Australian aid: promoting prosperity, reducing poverty, enhancing stability*, and its performance framework, *Making Performance Count: enhancing the accountability and effectiveness of Australian aid*, place a strong focus on performance, results and value-for-money. Using performance information as part of management and learning is critical for an effective Australian aid program.

Investment monitoring systems are at the foundation of DFAT's aid management system and external accountability reporting. These systems have been developed over time to support a performance culture that generates realistic and robust information on the performance of the aid program.

However, internal reporting from DFAT investment managers over several years indicates that the quality of investment monitoring systems have been persistently lower than the quality of other aspects

of aid investments.² These systems are not providing sufficiently robust evidence to adequately manage investment performance and underpin performance reporting.

Evaluation purpose and scope

This evaluation is intended to help DFAT improve how Australian aid investments are monitored, with a focus on investments delivered by managing contractors, the most significant cohesive group of implementation partners for the aid program. Investments implemented by managing contractors comprise about 20 per cent of the aid budget, and one-third of spending by country and regional programs. This is the highest proportion of aid delivered through a single type of partner.³

Managing contractors were also chosen as the focus of the evaluation because DFAT has stronger influence over how they monitor aid investments than it does over most other partners. Lessons from working with managing contractors are likely transferable to aid delivered through grants to other partners.

Key evaluation questions

The evaluation was guided by two key evaluation questions:

1. What are the characteristics of a DFAT better-practice investment monitoring system for programs delivered by managing contractors?

¹ DFAT, Workforce Plan, International Development, September 2018, internal document.

² DFAT investment manager ratings of the quality of investment monitoring systems have been persistently lower than other criteria assessed through an internal DFAT annual aid quality check process.

³ DFAT, *Performance of Australian Aid 2016–17*, p. 19. Multilateral organisations administer more of Australia's aid budget overall, at about 40 per cent; however, this includes both core and project funding, as well as a diverse range of partners from development banks to United Nations agencies (that is, a non-cohesive group).

2. What factors contribute to, or inhibit, better-practice investment monitoring systems delivered by managing contractors? What is the relative importance of those factors? What are the management implications for DFAT?

The evaluation considers a broad range of characteristics and factors including the cultures, policies and systems of both DFAT and managing contractor organisations. It also considers their impact on monitoring practice.

Summary of findings and recommendations

The evaluation found that Australian aid investments with higher-quality monitoring systems exhibit three distinct characteristics:

- » Systems are outcome focused, from beginning to end. They both measure and guide progress towards achieving intended outcomes.
- » Systems and data are quality assured through the application of quality standards and contestability mechanisms. External resources and independent perspectives are drawn on to quality assure methods and data.
- » Systems use monitoring data well, serving different purposes and needs. Multiple stakeholders use the information that the system produces often for multiple purposes.

Four factors were found to strongly determine the extent to which aid investments demonstrate these better-practice characteristics:

- » DFAT's performance culture is the most important determinant of investment monitoring system quality.
- » DFAT's ability to set and maintain clarity about aid investment objectives is a critical pre-condition for better-practice monitoring.
- » DFAT's demand for quality monitoring data and the systems required to generate this data provides incentives for managing contractors to deliver better-practice monitoring.
- » Managing contractor responsiveness to DFAT requirements and their ability to anticipate DFAT's needs is linked to the quality of monitoring systems.

The evaluation makes seven recommendations which are summarised here and in more detail in Table 1.

Recommendations for DFAT

DFAT recognises the importance of having a strong culture of performance. With reform ongoing within the department, the evaluation identified opportunities to build on these efforts.

The evaluation recommended that DFAT:

1. Promote consistent and robust investment level monitoring, and the performance culture to support this monitoring, across the aid program.
2. Strengthen its communication of the role of monitoring in supporting performance, diplomacy and strategic objectives.
3. Check investment monitoring system quality during the investment inception phase and invest resources to ensure DFAT M&E standards are met.
4. Ensure all investment managers have technical support to establish and oversee monitoring arrangements, especially for complex, technical and/or high-value investments.
5. Standardise monitoring expectations across managing contractor contracts.

Recommendations for managing contractors







Wide variation in the readiness of managing contractors to deliver quality monitoring systems requires a clear message from DFAT on the importance of monitoring. It is recommended that managing contractors:














6. Nurture a corporate culture of performance, including by building new capability and encouraging a cohort of staff to develop and maintain M&E expertise.
7. Support simple yet adaptable monitoring approaches, strengthened by learning across investments.










Table 1. Recommendations with supporting evidence

| Recommendation | Findings supporting recommendation with primary evidence source ⁴ |
|--|--|
| <p>Recommendation 1: DFAT to promote consistent and robust investment level monitoring, and the performance culture to support this monitoring, across the aid program.</p> <p>A consistent and robust approach to monitoring investment performance across the aid program will drive overall quality and effectiveness and strengthen accountability. Stronger and more reliable measures of aid investment-level performance will have the added benefit of helping support DFAT’s broader performance culture and performance reporting systems.</p> | <p>Organisational culture is a critical determinant of how well monitoring systems function.</p> <p>DFAT’s performance culture sets the expectations against which monitoring systems deliver.</p> <p>Senior leadership prioritising aid performance management will be critical to enable change.</p> <p>Better-quality monitoring systems are evident when DFAT sets clear performance expectations.</p> <p>Managing contractors would like more engagement by DFAT on performance.</p> <p>Managing contractors want a performance culture that rewards open and genuine participation in learning and reflection and that engages with monitoring data to assess performance.</p> |
| <p>Recommendation 2: DFAT to strengthen its communication of the role of monitoring in supporting performance, diplomacy and strategic objectives.</p> <p>DFAT should better promote the purpose of investment monitoring and its role in enabling the department to realise a range of broader objectives. This is to incorporate the importance of monitoring information for managing investment risk, supporting innovation, and adaptive management. This could occur, for example, through a letter of intent from the DFAT Secretary to posts and programs.</p> <p>Revised DFAT guidance could more clearly embed the establishment and use of quality investment monitoring systems into the routine of investment managers. At each stage of the investment cycle, a to-do list for managers with responsibilities for senior staff could be identified (Attachment B outlines the proposed DFAT toolkit for investment managers).</p> | <p>Senior leaders play a vital role in championing monitoring as part of a broader performance management system.</p> <p>DFAT’s performance culture varies significantly between divisions, branches, sections and posts.</p> <p>DFAT staff experience tension between needing to pay attention to managing aid and responding to political and diplomatic priorities (with the latter characterised as both more urgent and more visible to senior management).</p> <p>The importance of monitoring systems for investment management is not consistently recognised and prioritised. Use is not optimised throughout the investment cycle.</p> |

⁴ Key for evidence sources:

| | | | | | |
|---|-----------------------------|---|----------------|---|-------------------|
|  | Survey data |  | Interviews |  | Case studies |
|  | Aid quality check reporting |  | DFAT reporting |  | Literature review |

| Recommendation | Findings supporting recommendation with primary evidence source ⁴ |
|---|--|
| <p>Recommendation 3: DFAT to institute a check of monitoring system quality during the investment inception phase and invest resources to ensure DFAT M&E standards are met.</p> <p>DFAT should check investment monitoring systems towards the end of the inception phase. The timing can be flexible, depending on the type of investment. The purpose is to ensure that DFAT's monitoring standards are satisfied early in implementation, and to identify any outstanding weaknesses in the system leading to corrective action.</p> <p>DFAT should complement this with the review of monitoring guidance and contracts discussed in Recommendations 2 and 5 so investment managers, delegates and managing contractor teams have the information and tools necessary to prepare to meet this check point.</p> | <p>There is room for improving the clarity of DFAT's strategic vision for aid investments. </p> <p>There are incentives and constraints that weaken strategic clarity and the quality of designs. </p> <p>Clear objectives, theory of change, and baselines are a pre-requisite for an effective monitoring system. </p> <p>DFAT often does not sufficiently and clearly articulate what an investment is to achieve, and how, at the design stage (or during inception). </p> <p>All better-practice monitoring systems made significant efforts during design and inception, and sometimes well into implementation, to clarify what was expected of the investment. </p> |
| <p>Recommendation 4: DFAT to ensure that all investment managers have access to technical support to establish and oversee monitoring arrangements, especially for complex, technical and/or high-value investments.</p> <p>DFAT should support the provision of ongoing technical support to all investment managers through appropriate means (for example, Canberra or post-based advisers, access to panels of external expertise, or through a community of practice). This should be supplemented through the introduction of a toolkit for investment managers, developed as part of the response to Recommendation 2. This toolkit should identify the characteristics and determinants of better-practice monitoring.</p> <p>DFAT should use externally supported quality assurance mechanisms where warranted. The need here is greatest for higher-risk, large or complex investments.</p> <p>DFAT staff responsible for managing aid investments, should prioritise routine field-monitoring, proportionate to the value and complexity of the investment.</p> | <p>DFAT demand for quality monitoring systems is not stimulating managing contractors to supply quality monitoring. </p> <p>DFAT staff often do not have the time and skills to ensure that high-quality monitoring systems are established and used effectively. </p> <p>Although aware of M&E standards, DFAT staff are not always confident in applying them. They need support. This support should focus on the tools, methods and resources DFAT staff need to carry out active, technically informed oversight of investment monitoring systems. </p> <p>Too much emphasis has been placed on DFAT staff training as a solution to the capacity gap. Ongoing technical support to support quality monitoring can be more important and valuable than training. </p> <p>Better-practice monitoring systems draw on external resources and independent perspectives to quality assure methods and data. </p> |
| <p>Recommendation 5: DFAT monitoring expectations to be better standardised across managing contractor contracts.</p> <p>Standardising language in managing contractor contracts to reinforce the importance of quality monitoring can encourage managing contractors to pay appropriate and consistent levels of attention to monitoring. In addition to the referencing of DFAT M&E standards, this could include</p> | <p>DFAT information needs are not being well met by investment monitoring systems. </p> <p>DFAT contracting and procurement systems only partially support quality monitoring systems and the right incentives for quality monitoring. </p> <p>Contract clauses vary significantly. For example, DFAT's M&E standards are explicitly referenced in some agreements as benchmarks, and not in </p> |

| Recommendation | Findings supporting recommendation with primary evidence source ⁴ |
|---|--|
| <p>better recognition in contracts of different DFAT information needs.</p> | <p>others. Contracts often do not adequately establish a framework for transparent monitoring or performance remediation.</p> <p>Investment manager expectations around progress reporting varies widely. </p> <p>More generally, there are limited references to M&E standards and other performance policies in existing templates, guidance and professional development material. </p> |
| <p>Recommendation 6: Managing contractors to nurture a corporate culture of performance, including by building new capability and encouraging a cohort of staff to develop and maintain M&E expertise.</p> <p>Managing contractors to actively support M&E prioritisation and learning among their program staff. Actions to include:</p> <ul style="list-style-type: none"> » integrating monitoring activities into the roles and responsibilities of all program staff and communicating these expectations through performance agreements » enhancing the role of informal mentoring in supporting staff in the applying of monitoring standards and tools. | <p>The capacity of managing contractor teams to supply quality monitoring is a concern for DFAT and managing contractors themselves. </p> <p>Pressure to implement can contribute to insufficient time spent monitoring. </p> <p>There is wide variation in the technical readiness of managing contractor teams to adequately deliver good monitoring systems. Training of core staff is a common weakness. </p> <p>The quality of technical and managerial backup for team leaders and M&E advisers in the field varies widely. </p> |
| <p>Recommendation 7: Managing contractors to support simple yet adaptable monitoring approaches, strengthened by learning across investments.</p> <p>Managing contractors should enhance their own management of monitoring knowledge by:</p> <ul style="list-style-type: none"> » strengthening institutional processes and capacity to undertake and learn from better quality monitoring » ensuring the ability to apply, adapt and guide delivery of quality monitoring systems in new contexts » investing in ensuring simpler, more practical, fit-for-purpose approaches to investment monitoring, and informing prospective clients of this investment. | <p>Technical oversight and knowledge of best practice M&E by managing contractors is limited. </p> <p>Managing contractors lack uniformity in the extent to which they invest in institutional knowledge management in support of consistent and forward-looking approaches to investment monitoring. </p> <p>Significant opportunities exist to transfer institutional knowledge and systems to new investments and countries. There are examples of investments where significant M&E expertise is sourced external to the managing contractor, or where the system is developed largely in isolation, leading to reinventing the wheel. </p> |



MANAGEMENT RESPONSE

SUMMARY OF MANAGEMENT RESPONSE

DFAT notes the findings of this evaluation undertaken by the Office of Development Effectiveness. DFAT recognises that investment level monitoring systems are an integral component of the aid program's broader performance management architecture, and is therefore committed to ensuring they generate credible, relevant and timely data that informs decision-making and reporting.

Within DFAT, investment level monitoring systems are complemented by other performance management and accountability measures, including annual Partner Performance Assessments of key aid delivery partners, independent evaluations of aid investments, audits by the Australian National Audit Office and periodic review by the Development Assistance Committee of the Organisation for Economic Cooperation and Development. Together, these measures give effect to the Australian Government's aid performance framework, *Making Performance Count*, and provide assurance that Australia's development assistance is achieving key objectives and providing value for money.

DFAT welcomes the evaluation's identification of examples of good practice investment level monitoring systems among DFAT aid investments delivered by commercial contractors. However, DFAT also notes with concern the findings that the quality of investment level monitoring systems and the department's performance culture is not uniform, and that a relationship exists between the two. DFAT acknowledges the need to respond to these findings promptly.

DFAT agrees with all of the actions recommended by the evaluation that are within the direct control of the department (Recommendations 1 – 5). It also agrees in principle with Recommendations 6 and 7 and, while noting DFAT is not able to direct change within external organisations, the department will encourage their adoption by the commercial contractors towards whom they are directed.

RESPONSES TO EACH RECOMMENDATION

| Recommendation | Response ⁵ | Explanation | Action plan | Responsible area(s) & timeframe |
|---|-----------------------|--|--|--|
| <p>RECOMMENDATION 1:</p> <p>DFAT to promote consistent and robust investment level monitoring, and the performance culture to support this monitoring, across the aid program.</p> <p>A consistent and robust approach to monitoring investment performance across the aid program will drive overall quality and effectiveness and strengthen accountability. Stronger and more reliable measures of aid investment-level performance will have the added benefit of helping support DFAT's broader performance culture and performance reporting systems.</p> | Agree | DFAT notes the findings that the quality of investment level monitoring systems and the department's performance culture is not uniform, and that a relationship exists between the two. DFAT acknowledges the potential benefits to be realised through promoting greater consistency in the quality of investment level monitoring systems, while recognising the challenge of achieving uniformity across a large multi-functional network with finite resources. DFAT therefore agrees to take appropriate action to promote the replication of examples of good quality investment level monitoring systems in a context appropriate manner across the aid program. | To address the technical causes of uneven quality of investment level monitoring systems, DFAT will in the first instance undertake the actions agreed under Recommendations 3 – 5, and promote the actions suggested under Recommendation 7. To address the causes stemming from DFAT's performance culture and that of commercial contractors, DFAT will undertake the action agreed under Recommendation 2 and promote the action suggested under Recommendation 6. DFAT will continue to reflect on the technical and particularly cultural causes of uneven quality of investment level monitoring systems, and take further action as appropriate. | ACD in 2019 EXB and ACD jointly on performance culture in 2019, commencing in January |

⁵ Agree/Agree in part/Disagree

| Recommendation | Response ⁵ | Explanation | Action plan | Responsible area(s) & timeframe |
|--|-----------------------|---|--|--|
| <p>RECOMMENDATION 2:</p> <p>DFAT to strengthen its communication of the role of monitoring in supporting performance, diplomacy and strategic objectives.</p> <p>DFAT should better promote the purpose of investment monitoring and its role in enabling the department to realise a range of broader objectives. This is to incorporate the importance of monitoring information for managing investment risk, supporting innovation, and adaptive management. This could occur, for example, through a letter of intent from the DFAT Secretary to posts and programs.</p> <p>Revised DFAT guidance could more clearly embed the establishment and use of quality investment monitoring systems into the routine of investment managers. At each stage of the investment cycle, a to-do list for managers with responsibilities for senior staff could be identified (Attachment B outlines the proposed DFAT toolkit for investment managers).</p> | <p>Agree</p> | <p>DFAT recognises that a performance culture at the organisational level is essential to meet its legislative obligations under the PGPA Act and associated Rules, as well as key principles of transparency and accountability. Important progress has been made in recent years, including to better define strategic objectives via the Foreign Policy White Paper and 2018-2019 Corporate Plan. Agency progress against these objectives is reported to the Departmental Executive on a quarterly basis. Better and smarter communication is necessary to embed a performance culture across a multi-function agency.</p> <p>With respect to investment level monitoring, DFAT recognises that ongoing communication of expectations is important. While this communication already takes place (for example, through letters of intent from the Secretary to Heads of Mission and First Assistant Secretaries with aid responsibilities), DFAT will consider additional measures to strengthen communication.</p> | <p>At the organisational level, DFAT will continue to refine its annual and quarterly performance processes, and make its Annual Performance Statement more meaningful for Government, the Australian public and staff. A communications campaign in 2019 will help the department advance its performance culture. DFAT will identify and utilise tools to further embed expectations on performance, including for senior managers.</p> <p>With respect to investment level monitoring, this could be formal communication from the Secretary to ensure staff at all levels understand their responsibilities. DFAT will also review existing guidance for investment managers and strengthen messaging about expectations for investment level monitoring if necessary.</p> | <p>EXB and ACD jointly on performance culture in 2019, commencing in January</p> <p>ACD on investment level actions by July 2019</p> |

| Recommendation | Response ⁵ | Explanation | Action plan | Responsible area(s) & timeframe |
|--|-----------------------|--|---|---------------------------------|
| <p>RECOMMENDATION 3:</p> <p>DFAT to institute a check of monitoring system quality during the investment inception phase and invest resources to ensure DFAT M&E standards are met.</p> <p>DFAT should check investment monitoring systems towards the end of the inception phase. The timing can be flexible, depending on the type of investment. The purpose is to ensure that DFAT's monitoring standards are satisfied early in implementation, and to identify any outstanding weaknesses in the system leading to corrective action.</p> <p>DFAT should complement this with the review of monitoring guidance and contracts discussed in Recommendations 2 and 5 so investment managers, delegates and managing contractor teams have the information and tools necessary to prepare to meet this check point.</p> | Agree | DFAT recognises the importance of establishing quality investment monitoring systems early in the life of an investment. The department supports establishing a checkpoint to promote greater consistency in implementing existing policy in this area | <p>DFAT will standardise monitoring requirements in commercial contracts (Recommendation 5) to ensure the more consistent application of the principle that investment monitoring system should be established at an appropriate time towards the end of the inception phase of an investment. DFAT will also strengthen internal communication, guidance and technical support to better equip investment managers to certify that these requirements have been met (Recommendations 2 and 4).</p> <p>In addition, DFAT will continue to regularly assess investment level monitoring systems through annual Aid Quality Checks.</p> | ACD by July 2019 |

| Recommendation | Response ⁵ | Explanation | Action plan | Responsible area(s) & timeframe |
|---|-----------------------|--|--|---------------------------------|
| <p>RECOMMENDATION 4:</p> <p>DFAT to ensure that all investment managers have access to technical support to establish and oversee monitoring arrangements, especially for complex, technical and/or high-value investments.</p> <p>DFAT should support the provision of ongoing technical support to all investment managers through appropriate means (for example, Canberra or post-based advisers, access to panels of external expertise, or through a community of practice). This should be supplemented through the introduction of a toolkit for investment managers, developed as part of the response to Recommendation 2. This toolkit should identify the characteristics and determinants of better-practice monitoring.</p> <p>DFAT should use externally supported quality assurance mechanisms where warranted. The need here is greatest for higher-risk, large or complex investments.</p> <p>DFAT staff responsible for managing aid investments, should prioritise routine field-monitoring, proportionate to the value and complexity of the investment.</p> | Agree | <p>DFAT recognises that assessing the adequacy of investment level monitoring systems, particularly of complex high value investments, requires technical knowledge and experience. While DFAT investment managers already have access to training and expert advice on monitoring and evaluation, both internal and external, DFAT will seek opportunities to further improve these support systems.</p> <p>DFAT staff undertake field monitoring of aid investments and this practice will continue to be resourced and enhanced where needed.</p> | <p>DFAT will strengthen existing systems to ensure that investment managers have improved access to advice on investment level monitoring systems.</p> <p>DFAT will continue to resource field monitoring by staff who are responsible for aid investments, and enhance where necessary. The frequency and type of monitoring will remain proportionate to risk and feasibility, noting that aid investments are implemented in a range of contexts.</p> | ACD by July 2019 |

| Recommendation | Response ⁵ | Explanation | Action plan | Responsible area(s) & timeframe |
|---|-----------------------|--|--|---------------------------------|
| <p>RECOMMENDATION 5:</p> <p>DFAT monitoring expectations to be better standardised across managing contractor contracts.</p> <p>Standardising language in managing contractor contracts to reinforce the importance of quality monitoring can encourage managing contractors to pay appropriate and consistent levels of attention to monitoring. In addition to the referencing of DFAT M&E standards, this could include better recognition in contracts of different DFAT information needs.</p> | Agree | <p>DFAT agrees that inconsistency in contractual terms regarding standards for investment level monitoring systems can lead to variation in the quality of those systems.</p> <p>The Aid Business Branch of DFAT is currently undertaking a review of the standard aid contract template, an interim result of which includes standardising the requirement for commercial contractors to program and report consistent with DFAT Monitoring and Evaluation Standards.</p> <p>In mid-2018 the Aid Business Branch also established a taskforce to develop a contract management framework and associated tools and training to support better management of contracts. Performance measures, including M&E standards, are a key area of focus.</p> | DFAT will finalise work already in train to standardise contractual requirements regarding investment level monitoring and evaluation. | ACD by July 2019 |

| Recommendation | Response ⁵ | Explanation | Action plan | Responsible area(s) & timeframe |
|--|----------------------------------|--|---|--|
| <p>RECOMMENDATION 6:</p> <p>Managing contractors to nurture a corporate culture of performance, including by building new capability and encouraging a cohort of staff to develop and maintain M&E expertise.</p> <p>Managing contractors to actively support M&E prioritisation and learning among their program staff. Actions to include:</p> <ul style="list-style-type: none"> » integrating monitoring activities into the roles and responsibilities of all program staff and communicating these expectations through performance agreements » enhancing the role of informal mentoring in supporting staff in the applying of monitoring standards and tools. | <p>Agree in principle</p> | <p>While DFAT agrees in principle with this recommendation, it cannot direct change in external organisations. However, the department will strongly encourage adoption of this recommendation, including using levers at its disposal through the contracting process. The department works closely with the International Development Contractors Community (IDCC) Board – the peak body representing international development contractors and consultants working with the Australian aid program –and member organisations on a range of initiatives to improve delivery of the Australian Aid program.</p> | <p>DFAT will provide a copy of Recommendation 6 to the IDCC for their distribution in 2019. DFAT will also endeavour to promote a corporate culture of performance within commercial contractors through the contract management framework, providing guidance to investment managers, clarifying performance reporting expectations of MCs and requirements of MCs in relation to monitoring and evaluation.</p> | <p>ACD will provide a copy of the evaluation to IDCC by end January 2019</p> |

| Recommendation | Response ⁵ | Explanation | Action plan | Responsible area(s) & timeframe |
|---|----------------------------------|--|---|--|
| <p>RECOMMENDATION 7:</p> <p>Managing contractors to support simple yet adaptable monitoring approaches, strengthened by learning across investments.</p> <p>Managing contractors should enhance their own management of monitoring knowledge by:</p> <ul style="list-style-type: none"> » strengthening institutional processes and capacity to undertake and learn from better quality monitoring » ensuring the ability to apply, adapt and guide delivery of quality monitoring systems in new contexts » investing in ensuring simpler, more practical, fit-for-purpose approaches to investment monitoring, and informing prospective clients of this investment. | <p>Agree in principle</p> | <p>While DFAT agrees in principle with this recommendation, it does not have direct control over the internal business processes and practices of external organisations. However, the department will strongly encourage adoption of this recommendation.</p> | <p>DFAT will provide a copy of Recommendation 7 to the IDCC for their distribution in 2019.</p> <p>DFAT will also assess the sufficiency of investment monitoring capability in MCs through tender responses.</p> | <p>ACD will provide a copy of the evaluation to IDCC by end January 2019</p> |

REPORT STRUCTURE

This is the report structure:

| | |
|---|---|
| Section 1 Purpose, scope and approach to the evaluation | Provides an overview of the evaluation’s purpose, scope and the approach undertaken to respond to key evaluation questions. Outlines how evidence supports the findings and recommendations. |
| Section 2 Characteristics of better-practice investment monitoring | Describes the most important characteristics that distinguish better-practice monitoring of aid investments, with a focus on investments delivered by managing contractors. |
| Section 3 Key determinants of investment monitoring system quality | Describes the most significant factors that determine the quality of investment monitoring, with a focus on investments delivered by managing contractors. Key findings are presented in this section and linked to evaluation recommendations. |
| Annexes | Describes the methodology used to undertake the evaluation in further detail, sets out the Criterion Based Assessment Framework (CBAF) in detail, and provides the original terms of reference for the evaluation. |
| Attachments | Provides the literature review and a suite of tools to assist DFAT investment managers. |

PURPOSE, SCOPE AND APPROACH TO THE EVALUATION

1.1 BACKGROUND

Each year the Australian Government invests around \$4 billion to promote sustainable economic growth and poverty reduction in developing countries, as part of advancing Australia's interests internationally. It is a core function of DFAT to deliver this \$4 billion aid program effectively.

Currently, more than 2,000 DFAT staff—across 24 posts and 15 divisions—manage, monitor and report on the expenditure of more than 800 investments.⁶ Investments are delivered through country, regional, global and thematic aid programs, with a focus on the Indo-Pacific region.

The Australian Government's development policy, *Australian aid: promoting prosperity, reducing poverty, enhancing stability*⁷, and performance framework, *Making Performance Count: enhancing the accountability and effectiveness of Australian aid*⁸, place a strong focus on performance, results and value-for-money. Using performance information as part of management and learning is critical for an effective Australian aid program.

Investment monitoring systems are the foundation of DFAT's aid management system and external accountability reporting on Australian aid. These systems have been developed over time to support a performance culture that generates realistic and robust information on the performance of the aid program.⁹

⁶ DFAT, Workforce Plan, International Development, September 2018, internal document.

⁷ DFAT, *Australian aid: promoting prosperity, reducing poverty, enhancing stability* (2014).

⁸ DFAT, *Making Performance Count: enhancing the accountability and effectiveness of Australian aid* (2014).

⁹ *ibid.*

Defining performance culture

An organisation's performance culture is the mix of shared vision, results expectations, operational tools and workplace behaviours that define and reinforce success for the organisation's performance against its expressed objectives.

The objective of a performance culture is to generate realistic and robust information on performance, drive quality and effectiveness, and strengthen accountability. This includes the adequacy of M&E systems, the culture of contestability, being informed by evaluations, and adapting to ongoing learning. It sharply focuses on results, achieving better value-for-money, and getting the best development returns on each aid dollar spent.¹⁰

However, internal reporting from DFAT investment managers over several years indicates that the quality of investment monitoring systems have been persistently lower than the quality of other aspects of aid investments.¹¹ These systems are not providing sufficiently robust evidence to adequately manage investment performance and underpin performance reporting.

Using performance information as part of management and learning is critical for an effective aid program. Substantive engagement with country partners, and the building of

¹⁰ The definition of performance culture used for this evaluation is drawn from a mix of sources including DFAT external and internal documents.

¹¹ DFAT investment manager ratings of the quality of investment monitoring systems have been persistently lower than other criteria assessed through an internal DFAT annual AQC process.

country-partner capacity relies on the availability of good investment monitoring information.

With internal reviews ongoing within DFAT, there are opportunities to look at the utility of investment monitoring and consider pathways to improvement. These internal reviews include a deep-dive examination of the investment design process and options for improvement. They also include a broad-ranging response to recent Aid Program Health Checks¹² featuring internal reform initiatives addressing strategic clarity, governance, investment designs, performance culture, capacity (aid management skills and experience), implementation and other issues.

1.2 PURPOSE OF THE EVALUATION

This evaluation is intended to help DFAT improve how Australian aid investments are monitored. It is an important addition to the Office of Development Effectiveness' (ODE) ongoing work to independently assess and evaluate the quality and effectiveness of Australian aid. It also provides comprehensive evidence to inform DFAT's ongoing implementation of reforms to strengthen aid management. This is known as the Aid Program Health Check response. A more detailed evaluation rationale is provided in Annex 3, terms of reference.

The purpose of this evaluation is to guide improvements to investment monitoring systems so DFAT can more effectively contribute to poverty reduction and sustainable economic growth. The recommendations in this report are designed to help DFAT achieve better-practice monitoring, more consistently, across aid investments. The characteristics of better-practice monitoring were identified through a comprehensive literature review and detailed analysis of 78 investments.

¹² DFAT's Contracting and Aid Management Division (ACD) coordinates regular internal "Aid Health Checks" of the quality and effectiveness of DFAT's aid programs.

¹³ According to DFAT's *Aid Programming Guide*, value-for-money engages eight principles—cost consciousness, encouraging competition, evidence-based decision making, proportionality, performance and risk management, a focus on results,

Recommendations relate to the features of investment monitoring systems themselves, as well as contextual factors influencing how effectively these systems operate. There is strong evidence in the literature to suggest that better-practice monitoring provides a basis for more efficient, accountable and effective use of funds (Attachment A).

The implications of weaknesses in monitoring

Weaknesses observed in monitoring systems pose noteworthy threats to the quality of DFAT investments, risk management practices and ability to learn from experience and adapt accordingly. Weaknesses undermine DFAT's ability to ensure value-for-money of investments because this relies on decision-making processes that have a nuanced understanding of how value can be defined and quantified.¹³

Similarly, weaknesses may represent a missed opportunity to fulfil DFAT's public diplomacy objectives through the intentional and strategic use of information about Australian aid. The Fiji Community Development Program (FCDP)¹⁴ provides an example where public diplomacy was a key component of the results framework and reflected in the monitoring system. Funded civil society organisations (CSOs) uniformly reported on activities to increase the visibility of the aid program and detailed records of FCDP's media coverage and website traffic were maintained and reported to DFAT.

The FCDP is a highly visible program of development assistance. If implemented effectively with a commitment to ongoing due diligence and the continuous improvement of interventions, the program will be 'felt' as often as 'seen' and recognised. Its activities will involve the

experimentation and innovation, and accountability and transparency. Some of these principles are clearly associable with quality of investment monitoring. It may be argued that nearly all of them are functionally affected by quality of monitoring.

¹⁴ The FCDP is a case study for this evaluation. Section 1.4 has more details the evaluation case studies.

building of strong relationships between AusAID¹⁵, CSOs and the communities they serve. In collaboration with the managing contractor, AusAID will develop an engagement strategy during the inception phase to guide the promotion of these relationships as well as fulfil the program's requirements in relation to public diplomacy.

FCDP, design document, 2011

1.3 SCOPE OF THE EVALUATION

The evaluation focused on aid investments delivered by managing contractors. This was a strategic choice of focus, for several reasons.

First, managing contractors are the most significant cohesive group of implementation partners for the aid program. Investments implemented by managing contractors comprise about 20 per cent of the aid budget, and one-third of country and regional programs. This is the highest proportion of aid delivered through any single type of delivery partner.¹⁶

Second, other implementing partners have been subject to recent independent evaluations by the ODE, whereas managing contractors have not. ODE evaluations have considered the quality of monitoring and reporting of both non-government organisations (NGOs) and multilateral organisations.¹⁷ Evidence from these evaluations was included in the analysis for this report.

Finally, managing contractors were chosen as the focus of this evaluation because they are understood to be more responsive to demand from DFAT. Therefore, DFAT has stronger influence over how managing contractors monitor aid investments than it does over most other partners. Managing

contractors serve this evaluation as a critical test case. If DFAT can better enable quality monitoring from its managing contractor partners, most of the changes to its overarching policy, systems, culture and practice are likely transferable to its other partnerships.

The scope of this evaluation encompasses both managing contractors and DFAT, including their respective cultures, policies and systems, to the extent that these impact on monitoring practice. The monitoring system for an aid investment managed by a managing contractor is created and maintained through an interface between DFAT and managing contractor staff and systems, as well as those from any other implementing partners (such as partner governments). Accordingly, the culture, policies, resources and practices of both organisations are important determinants of how well the monitoring system works.

The evaluation has gathered evidence from both managing contractors and DFAT about the quality of monitoring for aid investments, and the most critical factors determining that quality. Importantly, it also makes recommendations to both groups about how to improve monitoring quality for aid investments delivered by managing contractors. The recommendations for DFAT are presented first and followed by recommendations for managing contractors.

Key evaluation questions

The evaluation was guided by two key evaluation questions:

1. What are the characteristics of a DFAT better-practice investment monitoring system for programs delivered by managing contractors?
2. What factors contribute to, or inhibit, better-practice investment monitoring systems

¹⁵ The Australian Agency for International Development (AusAID) was integrated into DFAT in 2013.

¹⁶ DFAT, *Performance of Australian Aid 2016–17*, p. 19. Multilateral organisations administer more of Australia's aid budget overall, at about 40 per cent; however, this includes both core and project

funding, as well as a diversity of partners from development banks to United Nations agencies (that is, a non-cohesive group).

¹⁷ *Banking our aid: Australia's non-core funding to the Asian Development Bank and the World Bank* (2015); DFAT, *Evaluation of the Australian NGO Cooperation Program* (ANCP) (2015).

delivered by managing contractors? What is the relative importance of those factors? What are the management implications for DFAT?

The analytical framework (summarised in Section 1.3 and detailed in annexes 1 and 2) clarifies the scope and definition of each question and establishes the basis for evaluative assessments. Care was taken to develop, communicate and implement a transparent and objective methodology so findings and recommendations would be reliable, defensible and applicable across a broad range of investments.

Section 2 sets out the characteristics of better-practice investment monitoring. Section 3 explores the key determinants of investment monitoring system quality. The evaluation also considers the extent to which investment monitoring systems lend themselves to fostering an awareness of and focus on gender equality and disability-inclusive development.

Understanding key terms

This evaluation is concerned with aid investment monitoring systems. DFAT's policy framework does not speak directly about investment monitoring systems. Rather, DFAT describes and sets standards for elements of the overall performance management system, with a focus on points in the investment cycle where accountability for monitoring can be reinforced. These key points for accountability are at three stages—investment design, contracting and review.

The term 'investment' is used throughout this report to describe DFAT's basic unit of aid programming. DFAT's Aid Programming Guide defines an aid investment as 'an intervention designed to achieve specific outputs and outcomes and contribute to the overall objectives of a program'. In other contexts, this might be called a project or program. However, DFAT's funding arrangements vary, and some aid

investments contribute to multiple projects, programs and/or organisations.

Undertaking monitoring is an assumed responsibility for DFAT staff within much of the organisation's internal guidance, including the Aid Programming Guide, Aid Quality Check (AQC) guidance and template, and DFAT Monitoring and Evaluation Standards. Monitoring itself is not explicitly defined by DFAT. The Organisation for Economic Co-operation and Development (OECD)¹⁸ defines it as:

Monitoring is a continuous function that uses the systematic collection of data on specified indicators to provide management and the main stakeholders of a development intervention with indications of the extent of progress and achievement of objectives and the use of allocated funds.

There is limited discussion in DFAT documents of the interplay between monitoring undertaken by its implementing partners, including managing contractors, and that performed by its staff, including their role in setting the parameters for monitoring by managing contractors. This evaluation applies this broader characterisation of a monitoring system¹⁹ which is established and maintained both by DFAT staff and by managing contractor teams.

1.4 APPROACH AND METHOD

The evaluation takes a utilisation-focused approach. As such, it is informed by the priorities of its primary stakeholders. Accordingly, the evaluation recommendations also target these primary stakeholder groups:

- » **DFAT senior management** who are responsible for promoting a strong performance management culture across the department
- » **DFAT delegates** who approve investment designs and contracting arrangements

¹⁸ World Bank, 'Ten Steps to a Results Based Monitoring and Evaluation System', <https://www.oecd.org/dac/peer-reviews/World%20bank%202004%2010%20Steps%20to%20a%20Results%20Based%20ME%20System.pdf>

¹⁹ Attachment A describes monitoring systems drawn from the literature.

- » **DFAT staff** who manage aid investments and/or programs, as well as associated performance and quality staff
- » **Contracting and Aid Management Division** which is responsible for DFAT's policy and guidance on investment design, M&E and procurement services, including the quality assurance and clearance of contracting approaches
- » **Managing contractor companies and M&E consultants** who are involved in the design and delivery of the Australian aid program.

The evaluation draws on qualitative and quantitative research techniques, applied in a sequential approach. The evaluation was conducted in three phases—inception, research, analysis and reporting. A brief outline of methods for data collection and analysis is presented here. Annex 1 provides a detailed description of the phases and methods of data collection and analysis.

Data collection and analysis

Portfolio analysis: The evaluation team undertook statistical analysis of a database with details of all DFAT aid investments to identify those within the scope of the evaluation.

Desk review: The evaluation team then reviewed the literature on monitoring aid investments and the challenges associated with quality assurance in monitoring more generally. The review analysed DFAT's own policy framework for monitoring aid as well as academic and grey literature.²⁰ Results are synthesised in Attachment A.

Analytical framework: Drawing on the literature review, the evaluation team developed a Criterion Based Assessment Framework (CBAF) (Annex 2), to guide the assessment of DFAT's investment

monitoring systems. The CBAF includes 16 quality indicators in four key domains: strategy; infrastructure; capacity; and enabling environment. Its framework draws on both DFAT and external literature. The team anticipates that the CBAF can be applied to assess the monitoring systems of investments delivered by other implementation partners aside from managing contractors.

Review of quality reporting: The evaluation team identified a subset of managing contractor-delivered investments potentially suitable for case studies (N=78). These were either ongoing in 2018 or had finished in 2017 and were not exempt from DFAT quality reporting processes; yielding more recent data and opportunity to contact implementing staff. The team sourced and analysed 172 AQC reports for this sample, using the CBAF for analysis. From this analysis, the team developed a database of examples of better and sub-optimal practices. This was useful in exploring the factors contributing to better practice.

Survey of DFAT and managing contractor staff: The evaluation team designed and administered an online survey of DFAT staff²¹ (97 responses) and managing contractor staff²² (34 responses) to capture views on investment monitoring and the factors that enable and inhibit better practice.²³ The survey gathered data on factors identified in the CBAF as being the most critical to support and exemplify better practice. The resulting dataset enables comparison between the views of DFAT staff and managing contractor staff, and therefore an understanding of where perspectives diverge.

There are some limitations to the data. The response rate cannot be ascertained for either set of responses. For DFAT, 97 responses represent about 5 per cent of all staff, however not all staff are

²⁰ Grey literature is research that is either unpublished or has been published in non-commercial form.

²¹ DFAT staff are based in Canberra and overseas (posted officers and locally engaged staff).

²² Managing contractor staff included those engaged to work on specific investments (project staff) and those responsible for corporate systems and policy.

²³ Response rates are not available because the surveys were distributed by DFAT and by managing contractors internally through an online link.

investment managers or supervisors, so the response rate of the target audience is unknown. Also, no responses were received from the Pacific Division.²⁴ The managing contractor response rate is relatively low at 34, however responses were received from all organisations identified as major contracting partners for DFAT (based on financial data). Data from the survey was triangulated with interviews and case studies to confirm findings indicated by managing contractor and DFAT staff answers to survey questions.

Stakeholder interviews: The evaluation team also conducted 40 semi-structured interviews to explore stakeholder views in greater detail. Some interviews were used to develop case studies and others to investigate aspects of DFAT and managing contractor policies, systems and practices. Interviewees included DFAT and managing contractor staff working on case study investments or responsible for relevant policies on M&E performance and design. Interviewees also included individual consultants who undertake a significant amount of M&E work for DFAT.²⁵ Interview notes were coded against the CBAF, yielding further qualitative insights into factors affecting monitoring practice.

Case studies: Finally, the evaluation team developed case studies for eight investments managed by managing contractors (identified in Table 2). The team interviewed managing contractor staff, DFAT staff and consultants engaged in M&E to explore how the factors influencing the development and implementation of investment monitoring systems were related. Comprehensive documentation on implementation and results was sourced and analysed against the CBAF.

Table 2. Diversity in context, approach and experience—the eight case studies²⁶

| Investment | Value (AUD million) | M&E score ²⁷ |
|--|---------------------|-------------------------|
| Australia–Indonesia Partnership for Rural Economic Development (AIP-R) | \$112 | 6 |
| Fiji Community Development Program (FCDP) | \$20.9 | 5 |
| Tonga Skills for Inclusive Economic Growth Program (S4IEG) | \$7.6 | 5 |
| Policing and Justice Support Program in Vanuatu (PJSPV) | \$72 | 5 |
| Cambodia Agricultural Value Chain Program (CAVAC) | \$94 | 4 |
| Indonesia Governance for Growth (KOMPAK) | \$81 | 4 |
| Transport Sector Support Program (Papua New Guinea) (TSSP) | \$400 | 4 |
| Strongim Gavman Program (Papua New Guinea) (SGP) | \$302.4 | 2 |

²⁴ The reason for this is not clear.

²⁵ Identified through consultation with ODE.

²⁶ Annex 1 has details on the method of case selection.

²⁷ This reflects the last or most recent AQC M&E score.

Synthesis and reporting

The team compiled a significant and diverse set of data using NVIVO.²⁸ Each data source—interview transcript, report, policy document—was categorised or coded against all relevant evaluation questions and components of the CBAF. Analysis of coded material was then undertaken to identify emerging trends or observations and to triangulate findings between multiple sources of evidence. This iterative process is outlined here:

- » All data sources were first mapped and coded against the CBAF and then key information coded against each key evaluation question and sub-question.
- » The team triangulated data coded to each key evaluation question and component of the CBAF, identified initial findings and developed presentation material to share with DFAT.
- » The team presented preliminary findings to, and incorporated feedback from, the Evaluation Reference Group, which included representatives from DFAT's geographic divisions.
- » The team then facilitated a recommendations workshop with three key stakeholder groups—senior DFAT management, members of the Contracting and Aid Management Division and representatives from managing contractor organisations. The purpose of the workshop was to receive early feedback on evaluation findings and recommendations and identify further analytical tasks if needed.
- » Based on results of the qualitative analysis and consultations, the team developed a draft toolkit for DFAT investment managers and held a workshop to test the usefulness and refine the content of the toolkit (Attachment B).

1.5 STRENGTH OF THE EVIDENCE BASE

A large amount of data from multiple sources was collected for this evaluation. The NVIVO database used to code the materials contains 435 megabytes of data, comprising 172 AQC reports, records from 40 interviews, program documentation for 8 case study investments, 131 responses to the survey, DFAT policy documents and evaluation reports relating to monitoring, as well as academic and grey literature included in the literature review.

Triangulation revealed a substantial level of agreement between data sources. For example, interview comments were generally consistent with the quantitative results of the survey data. Table 1 in the Executive Summary sets out how data sources contributed to key findings and specific recommendations. All findings included in this evaluation are based on evidence triangulated between multiple sources. The findings are therefore robust and representative of a broad range of aid investments.

²⁸ NVIVO is a software often used in social research to assist in the analysis of qualitative (narrative) data.

CHARACTERISTICS OF BETTER PRACTICE INVESTMENT MONITORING

This section responds to key evaluation question 1. It describes the characteristics of better-practice monitoring as they apply to Australian aid investments and provides illustrative examples from investments implemented (or monitored) by managing contractors.

2.1 WHAT IS BETTER PRACTICE?

A substantial body of literature exists on what constitutes better-practice monitoring in the development sector, and more broadly. The literature argues that strong monitoring arrangements are those that are ‘planned, continuous and systematic, and documented’.

For this evaluation, quality monitoring has been defined as the extent to which investment monitoring systems enable the generation, collection and analyses of credible information on aid activities and that apply internationally recognised characteristics of good aid practice. Attachment A comprehensively summarises the characteristics of high-quality and better-practice monitoring systems, drawing on extensive literature.

This section identifies the most important features that characterise the higher-quality investment monitoring systems implemented by managing contractors on behalf of DFAT.

Aid investments with higher-quality monitoring systems exhibit three distinct characteristics:

1. **Outcomes-focused:** Systems are outcome focused, from beginning to end.
2. **Quality assured:** System and data quality is assured through the application of quality standards and contestability mechanisms.
3. **Effectively used:** Systems use monitoring data effectively, serving multiple purposes and needs.

Taken together, these three characteristics reflect a basic theory of how monitoring works. First, the monitoring system needs to generate information about what matters most, including outcomes. Second, the information needs to be an accurate and valid measure of what matters (that is, it needs to be quality assured). Third, a range of stakeholders needs to act appropriately in response to the information generated (that is, they need to use monitoring data well).

Despite these shared characteristics, higher-quality monitoring systems can look and work differently depending on their context. This section describes the characteristics of better practice and presents examples drawn from a range of investments, including the case studies.

The evaluation team used better-practice identified through the literature review to develop the CBAF (Annex 2 has further detail). The CBAF was used as the framework for analysing the quality of DFAT investment monitoring systems and practice. It categorises the characteristics of better-practice monitoring into four domains—strategic, infrastructure, capacity, and enabling environment—as well as 16 indicators. The evaluation analysed 78 investment monitoring systems to assess the extent to which they corresponded to these indicators and represented better-practice.

The **strategic domain** describes the strategic context within which a monitoring system is established and sustained. High-quality systems require an understanding of how monitoring information can

assist decision makers to set directions and guide investments. This requires strategic leadership as well as a clear understanding of the basic concepts and potential uses of M&E data and information. Evidence against this domain forms the basis of Section 2.2, which describes outcomes-focused monitoring systems.

The **infrastructure domain** describes the infrastructure needed to ensure a systematic, comprehensive and credible approach to monitoring. Evidence against this domain is significant for Section 2.3, which considers how this monitoring architecture and its products can be quality assured.

The **capacity domain** describes the capacity to supply and use monitoring system information. This requires both clarity of expectations on where and how information is intended to be used (for example, planning, policy or program development, decision making or budgeting), as well as the capacity to incorporate and use the information as part of normal business. Evidence against this domain was found to be a significant characteristic of better practice. This is considered in Section 2.4, which is concerned with the effective use of monitoring data.

The **enabling environment domain** describes a culture in which:

- » investment managers have a suitable appreciation of M&E concepts
- » there are adequate incentives for managers to use M&E information
- » investment managers report credible, unbiased and timely results.

Evidence against this domain is explored in the next chapter on key determinants of investment monitoring system quality.

The evidence base for these findings

The evidence discussed in this section draws on all sources used by the evaluation. Findings were developed by:

- » Synthesising the characteristics of better-practice monitoring through the literature review and by developing the CBAF.
- » Reviewing 172 AQC reports for the sample of 78 investments managing contractors delivered, and analysing evidence of stronger and weaker monitoring practice against CBAF indicators.
- » Using case study research to explore the contextual factors that enabled or inhibited higher-quality monitoring systems and develop lessons learned to inform recommendations.
- » Using survey results to understand managing contractor and DFAT perceptions of better practice.
- » Determining the most important characteristics of better practice based on the success case method²⁹ and triangulation with other data sources (AQC reports and survey).

To select the case studies, the evaluation team worked with ODE to identify investments which received a high or low rating for M&E, and which DFAT staff considered to exemplify better or weaker practice monitoring. The team investigated these cases in detail, applying the CBAF as a lens. Findings from the case studies were cross-referenced with other sources.³⁰

Current practice

DFAT investment monitoring is currently characterised by a wide range of technical rigour, utility for stakeholders, and alignment with the DFAT M&E standards. For example, while the M&E standards generally provide useful quality guidance for M&E frameworks, in practice these frameworks

²⁹ The success case method involves identifying the most and least successful examples and analysing them in detail.

³⁰ Other sources included AQC reports, relevant policy documents and guidelines, survey responses, and in-depth interviews with independent M&E consultants.

vary greatly in content and quality, with some offering little information on the critical point of how progress toward end-of-program outcomes will be measured and assessed.³¹

Many investments exhibit better monitoring practice in some respects and weaker practice in others, as the case studies illustrated. The standard to which monitoring systems are held might reasonably be expected to vary according to country, sector and intervention context, a reality often reflected in the narrative accompanying AQC ratings for M&E. Accordingly, the evaluation did not classify the sample of 78 investments into those that exhibit better and poorer monitoring practice. Rather, a qualitative assessment was made to determine which characteristics clearly distinguish the monitoring systems that DFAT, managing contractors and the evaluation team agree represent better practice.

2.2 OUTCOMES-FOCUSED

A better-practice monitoring system must both measure and guide progress towards achieving the intended outcomes.

Any investment and program of work requires a clear rationale, well-articulated outcomes to be achieved within the life of the investment and considered approaches to achieving those outcomes. A robust program theory is needed against which progress will be monitored. The expectations for operationalising these features into a monitoring plan are provided for by the DFAT M&E standards (Standard 2).

This evaluation found four important characteristics of outcome-focused monitoring systems. First, there is clarity about what the investment needs to achieve, and how it is expected to effect change. Second, a better-practice monitoring system distinguishes between outcomes realised for different groups of people, including, but not limited to, women and men. Third, the system generates

feedback early enough to inform changes in implementation. Finally, this feedback is used to inform any needed adjustments to the investment design, including its theory of change.

The monitoring system is built around what the investment needs to achieve

DFAT's M&E standards specify that a monitoring system must measure progress towards outcomes. This evaluation found that most investment monitoring systems aim to do this, but many are hampered by complexity, a profusion of indicators, and large logical leaps between activities and outcomes.

Some investments have monitoring systems that are guided by a clear, coherent theory of change and deliver performance data to assess progress against intended outcomes and for more in-depth learning. Other investments include weak or non-existent theories of change and little to no information on actual progress towards outcomes.

Several high-performing monitoring systems are structured around a limited number of measures that clearly articulate achievement. These come in the form of key performance measures or indicators, results, or outcomes. For example:

- » AIP-R in Indonesia³² has a clear goal to increase incomes in more than 300,000 smallholder agricultural households, and eight key performance indicators that flow from this and govern the investment.
- » KOMPAK in Indonesia structures reporting from a diverse portfolio of interventions against levels of outcome and impact and describes the type of evidence that is sought for each level.

³¹ M&E Standard 3.4, under Investment Progress Reporting, notes that 'a firm judgement of the adequacy of progress toward these outcomes is [to be] described'.

³² Investments referred to in the evaluation report are evaluation case studies unless otherwise specified.

- » CSO WASH Fund³³ (now Water for Women Fund) requires multiple implementing partners to report using a single, simplified theory of change, resulting in partners' use of the same language to describe the same parts of the change process.

These outcomes-focused indicators bring several benefits. They enable managing contractor partners to provide high-quality and synthesised reporting of outcomes, even when the intervention is complex or diffuse. They also enable comparison of performance over time, and between interventions, if the same data is collected for each. Most importantly, they build a focus on the most important activities and changes required to realise outcomes, thereby encouraging attention where it is most warranted.

The monitoring system distinguishes the outcomes realised for different beneficiaries

A high-quality monitoring system is built around the understanding that different people will interact with aid interventions differently, with different results. At the very least, this means distinguishing the differences in participation and results for women and men. Depending on the intervention, the different outcomes realised for people with disabilities, the young, the elderly, minority groups and other distinct categories may be strongly relevant to understanding results overall.

Addressing gender equality involves tracking the actual effects an investment has on women and girls, above and beyond disaggregation of monitoring data by sex. While relevance of tracking the roles of women and girls varies by investment, from the AQC evidence it appears that managing contractor-delivered investment monitoring is only partially meeting the expectations expressed in DFAT's

Gender Equality and Women's Empowerment Strategy.³⁴ A few investments track the effect of investments in ways that provide meaningful data on the different outcomes for men and women. For example, AIP-R and the Market Development Facility³⁵ collect qualitative data to help explain gender-disaggregated quantitative data on jobs, outreach, additional income and market transactions. The Market Development Facility also measures more complex indicators, especially for women's economic empowerment. This includes the degree of involvement of women in decision making around market transactions, farm practices, household expenses, and finances.

Information is collected in time to inform ongoing implementation

Monitoring systems that provide useful information in time to inform and influence ongoing implementation have several features in common. These higher-quality systems:

- » Invest in collecting adequate baseline data, in time to inform the intervention. For example, a baseline study conducted in Sri Lanka in 2016 helped the Market Development Facility to design partnerships to strengthen seafood supply chains through engaging women.³⁶
- » Collect information on early outcomes. Market system development investments³⁷ often do this well by forecasting the expected results of each intervention and comparing actual to forecasted data.
- » Begin collecting impact information as soon as possible. AIP-R, for example, will design and conduct an impact study as soon as the managing contractor team recognises early indicators of success. This means the investment has time to consider early

³³ This investment is among the broader sample of investments analysed for this evaluation.

³⁴ See, for example, Item 5 in the Strategy's Annex 1 (Gender Equality in Development), 'Build gender equality and women's empowerment explicitly into monitoring, evaluation and learning processes'.

³⁵ This investment is among the broader sample of investments analysed for this evaluation.

³⁶ DFAT, *Inclusive Economic Growth* AQC report, 2017, internal document.

³⁷ Markets systems development interventions often measure results using an approach set out in the DCED standards, which encourage a rigorous approach to performance benchmarking. <https://www.enterprise-development.org/measuring-results-the-dced-standard/>

indications of success or limited progress and pivot as required.

Baseline data forms the bedrock of an outcomes-focused monitoring system. The existence of robust baseline data provides a foundation for planning, monitoring and tracking progress against outcomes. For example, the Market Development Facility intentionally conducts tightly focused baseline studies immediately before interventions, aiming to ensure the data is current and adequate to inform the design and monitoring of the intervention.

The evaluation found nearly 70 references to the importance of baseline data and studies, mostly in AQC reports. These typically focused on the adequacy, completeness and timeliness of baseline data, and how it had helped or hindered in determining progress towards outcomes. KOMPAK, AIP-R and PJSPV invested considerable effort in identifying indicators and collecting baseline data. CAVAC took an iterative approach to baseline data collection, which was critiqued at its mid-term review for contributing to difficulties in measuring outcomes.

Feedback is actively used to adapt implementation (adaptive management)

The ability of investments to adapt to changing context—which is crucial if investments are to be effective—is dependent on the monitoring system’s ability to provide a clear picture of evolving investment status vis-à-vis program objectives. Investments that use monitoring data to learn and make decisions about implementation have certain features. They:

- » Focus on what is needed to inform decisions. Investments that use monitoring data well have often focused data collection on a few key indicators which are agreed among stakeholders to be the most important measures of success.

- » Have criteria in place to make decisions. AIP-R is a strong example, with the managing contractor team reporting a clear expectation that up to 30 per cent of investments *may* fail. The team has an internal process in place which uses monitoring data to identify which investments will succeed and fail as early as is pragmatic.³⁸
- » Have processes to facilitate use of information. Regular, informal, team-based reflection exercises were often cited as being particularly useful. For example, the Pacific Leadership Program used six-monthly reflection sessions to identify issues, such as a lack of inclusiveness, which the team later addressed through a partnership with a women’s organisation.³⁹

Australia Awards’ scholarship programs are particularly rich in monitoring data and often use it to inform and undertake adjustments in delivery. With these programs, there is typically clarity and agreement between DFAT and the managing contractors about the most important indicators of success (for example, completion rates and alumni career paths). Data from Awards’ monitoring systems has been used to inform refinements in institutional and individual targeting, marketing, application and selection processes, on-award and alumni management, and track recommendation implementation.⁴⁰

Theories of change during the design phase often need to be adjusted during implementation. A quality monitoring system will respond to this need with revisions to the theory and refinements to the monitoring framework. Failure to adapt the system has the dual impact of delaying the development of a clear M&E framework during inception and weakening the ability to report credibly on progress

³⁸ Interviews with managing contractor team, corroborated with other interview sources.

³⁹ DFAT, *Pacific Leadership Program AQC report*, 2017, internal document. This investment is among the broader sample of investments analysed for this evaluation

⁴⁰ AQC reports for a range of Australia Award investments, internal documents. These investments are among the broader sample of investments analysed for this evaluation.

towards outcomes during implementation and completion.

This evaluation found more than 80 references noting the importance of theory of change in monitoring systems, most from evaluations and AQC reports. These typically comment on the realism of intended outcomes, adequacy of the logic linking activities to outcomes, and updates or clarifications to a theory of change.

2.3 QUALITY ASSURED

A better-practice monitoring system draws on external resources and independent perspectives to quality assure its methods and data.

While monitoring systems must be tailored and context responsive, approaches, methods, tools and measures can be informed by those used for other interventions. That is, the monitoring system need not be built from scratch, and the team building it should at least refer to external resources so relevant tools and approaches are considered.

All investments must meet DFAT's M&E standards which are a point of reference for all investment monitoring systems. Other external sets of standards are available for some types of interventions, such as the Donor Committee for Enterprise Development (DCED) M&E Standard.⁴¹ A better-practice monitoring system is built with reference to these resources, even if it exceeds them in quality.

Better-practice monitoring systems also seek and use independent feedback and perspectives to quality assure the system and the data it produces. This counteracts some of the human challenges of building and using such a system, including tunnel vision, logical leaps and confirmation bias.

Monitoring standards are applied to assure high-quality data and reporting

Many higher-quality monitoring systems refer to and build upon departmental or external quality standards. In interviews, DFAT program staff were readily aware of the department's M&E standards. Standard Two, Investment Monitoring and Evaluation Systems, and Standard Three, Investment Progress Reporting, are the most relevant to investment monitoring. The standards also refer to more detailed technical guidance, such as the OECD–DAC Standards.⁴² Some managers also use these external resources.

Some investments have sector-focused resources to bolster contestability. Perhaps the most prominent example is the DCED Standard for M&E. This standard is a seven-part framework for results management and adaptive learning. In addition to technical references, DCED also provides periodic performance audits to organisations that participate with it. AIP-R has used the DCED standard and organisational support very effectively to develop and quality assure its monitoring system.

Contestability and independent review are built in to the system and culture

This evaluation found several instances of evaluations being used to strengthen monitoring systems. In better-practice systems, this is intentional and planned for. KOMPAK's monitoring system, for example, was influenced by several review processes during its inception phase. The mid-term review of KOMPAK recommended that the investment revise its program theory and invest in a more sophisticated management information system. It also proposed improvements to specific monitoring tools.

Having monitoring systems issues addressed early in implementation is important so an investment can generate and use data effectively. The Papua New

⁴¹ <https://www.enterprise-development.org/measuring-results-the-dced-standard/>

⁴² The Organisation for Economic Co-operation and Development's Development Assistance Committee (OECD–DAC) is a donor forum

to discuss issues surrounding aid, development and poverty reduction in developing countries.

Guinea TSSP case study provides an excellent example of the risk associated with weak mechanisms to verify data, especially when there are known issues with data quality. A Visual Road Condition Survey verified, for example, that 13 per cent of national roads were in good condition, where the program had reported that 46 per cent of roads were in good condition.

Collaborative accountability between DFAT and the managing contractor team was rare in the investments reviewed but identified as an important characteristic of better practice. Interviews with stakeholders from the PJSPV indicate that both DFAT and the managing contractor are interested in frank performance information and want to see it used. When DFAT has been satisfied about something or identified a gap in the program, this was raised directly with the managing contractor team, in time for it to be addressed. Any issues not addressed would then appear in the annual Partner Performance Assessment.⁴³

Technical quality assurance of monitoring systems and products

This evaluation found that building independent quality assurance into the design of an aid investment can help to incentivise the collection and use of monitoring data that would convey a more varied, textured story about performance. It can also directly strengthen contestability of reporting.

The higher-quality monitoring systems reviewed in this evaluation featured some form of independent or semi-independent group that would serve as reviewer and source of selective technical support on M&E. Such entities, including the Papua New Guinea Governance Facility Quality and Technical Assurance Group (QTAG) or the M&E House in Timor-Leste, are familiar to many DFAT program staff. Small shifts in the resourcing dedicated to quality assurance can bring out the best in already high-performing

managing contractor teams. The quality assurance can help to identify poorer performing investments sooner.

The extent of independent checking of data and reporting should reflect the value and complexity of the investment and draw upon the appropriate expertise.

2.4 EFFECTIVELY USED

When a better-practice monitoring system is in place, multiple people frequently use the information it produces, and for many purposes.

The literature is clear that a monitoring system only serves its purpose if the information it produces is used.⁴⁴ Monitoring data can serve a range of performance management, learning and accountability functions, summarised here:

- » **Performance management**—indicating progress in implementation against outcomes.
- » **Learning**—indicating what is working, what is not working, and significant changes in the context that might affect results.
- » **Accountability**—recording how funding and other resources are being used, how policies and safeguards are being complied with, and how risks are being managed.

A higher-quality monitoring system provides a continuous flow of information that is useful internally and externally. Within DFAT, the data is a crucial tool for the investment manager. Information on process, progress, problems, and performance are all key to managing for results and holding partners to account. Monitoring data also flows upwards to inform DFAT's broader performance, quality and reporting systems.

Likewise, the information from a monitoring system is important externally to those who are expecting results and wanting to see demonstrable impacts

⁴³ DFAT uses Partner Performance Assessments to assess how well implementing partners are delivering the services required.

⁴⁴ Attachment A has more detail. Information use is assumed or built into many of the literature findings on what makes a monitoring system effective (for example, demand factors).

from government action (and tax monies). DFAT, partner governments, managing contractors and other stakeholders can also use monitoring data to formulate and justify action in other arenas, including budget allocations.

This evaluation identified three characteristics that distinguish how monitoring data is used in higher-quality systems. First, a mix of data is collected, including qualitative and quantitative data. Second, reporting is utilisation-focused and tailored to meet the needs of its audiences. Third, external partners are supported to engage with and use monitoring data that is relevant to them.

A mix of qualitative and quantitative data is collected

A mix of quantitative and qualitative data is more likely to meet diverse and changing information needs. DFAT often relies on quantitative data to assess progress, and qualitative data to articulate why it matters. The development impact of investments is best articulated with a mix of both types of data. The use of mixed methods also protects systems from issues with data quality and collection. For example, if quantitative data is corrupted or unavailable, or key performance indicators change, qualitative data can help to triangulate and fill information gaps.

Independent M&E consultants familiar with DFAT have cited examples where mixed-methods approaches have been applied. One example is the FCDP, which used qualitative evaluation studies to investigate unsolved problems or questions brought to stakeholder attention by quantitative evidence. Such an approach produces higher-quality reporting and stimulates more active program learning than the more common counting approaches.

Reporting is functional, and tailored to the needs of its audiences

Reporting is functional, which means it is concise, tailored to audience, and connected to decision points. A higher-quality monitoring system reports at multiple levels. Stakeholders (DFAT staff, managing contractor representatives and M&E consultants) all conveyed that reporting occurs on at least two levels—internally among implementation partners and externally to DFAT. Implementation partner arrangements vary widely. One investment might comprise staff recruited entirely by the managing contractor (such as with most Australia Awards programs). Another investment may have a CSO or NGO reporting on the use of grant funding to a managing contractor administrator or monitoring panel (such as the CSO Wash Fund). Another investment might have managing contractor staff working directly with a partner government (as with PJSVP or TSSP). In each instance, internal reporting of monitoring information among the investment team serves an important management purpose.

Higher-quality systems invest in developing reports tailored to meeting DFAT's needs and using specifically designed reporting formats, succinct written products, and a style of mutual consultation with field partners. For some investments, the application of user-friendly performance reporting instruments serves the dual purpose of facilitating the flow of reports from partners and supporting the enhancement of monitoring capacity among these organisations.

This evaluation found that most investments have scope to improve how monitoring data is used in reporting. Only one-third of DFAT staff surveyed agreed to a 'great' or 'very great' extent that investments generally have monitoring systems that produce the information the department needs.⁴⁵ Managing contractor staff respondents were more optimistic, with nearly half believing that DFAT's information needs are met.

⁴⁵ Around 30 per cent of DFAT survey respondents agree to a 'great' or 'very great' extent. Around 65 per cent agree only to 'some

extent' or a 'small extent'. The remainder 'don't know', or 'don't agree at all'.

This discrepancy suggests that DFAT's information needs are not universally well understood. Interviews confirmed this, with several managing contractor respondents describing a process of guessing what individual DFAT staff want to see in progress reporting and noting that expectations are not standardised, despite the M&E standards. A common theme was that DFAT 'knows what it doesn't want when it sees it', but is not always able to say 'what it does want to see' in advance.

External partners and stakeholders are supported to engage with data

The most commonly cited users of monitoring system data (outside managing contractors and DFAT) are partner governments. Often, information use is limited to decision making about the investment itself. In some cases, monitoring data is used to supplement partner government's systems. An example of this is the visual survey of roads completed by the TSSP, which corrected information in the national road management system.

Better-practice monitoring also aligns with and seeks to strengthen national monitoring systems. The monitoring system for the Basic Education Sector Transformation project in the Philippines⁴⁶, has two critical areas of focus—tracking and capturing overall results of the project and strengthening the capacity of the Philippines Department of Education to clearly link priorities and strategies to program delivery outcomes. The work of the department's Planning and Program Division features quarterly meetings of the Economic Task Force, which includes the Secretary for Finance and the Counsellor for Development. These sessions offer these local leaders the opportunity to inquire about performance and support ongoing efforts.

The Fiji Health Sector Support Program (FHSSP)⁴⁷ seeks to align its M&E efforts with those of the

Ministry of Health. Most indicators in the FHSSP monitoring system are also used by the Ministry. Additional indicators or data collection mechanisms are approved by the Ministry of Health as appropriate and are intended to be adopted by it in the future for continued use post-FHSSP.

In higher-quality monitoring systems, monitoring data plays a key role in forming the case for pilot reforms to be adopted. Some investments have policy influence as an explicit aim and claim scale-up of proven interventions as part of their impact. KOMPAK is an interesting example, as it was designed to support decentralisation of basic services and implementation of Indonesia's Village Law. KOMPAK implements pilot projects, with approval from the Government of Indonesia, to test approaches to improving village governance and the allocation of village funds. It then provides advice to multiple levels of government based upon the evidence produced by pilots. KOMPAK's 2017 mid-term review found that replication of pilot interventions across districts has been substantial, and that there is promise of adoption of some reforms at national level.⁴⁸

There are some examples of higher-quality monitoring systems that enable beneficiaries to use monitoring data. KOMPAK provides an example of what this can look like. The investment has helped to establish a village information system in 307 villages. In 291 of those villages, this includes data on persons living with disabilities. The FCDP deliberately collected monitoring data in collaboration with beneficiary CSOs and communities and used data collection exercises as an opportunity to build beneficiary capacity to understand and use data.

The evidence reviewed for this evaluation indicates that few of DFAT's investment monitoring systems do well at representing the nuanced interests of

⁴⁶ This investment is among the broader sample of investments analysed for this evaluation.

⁴⁷ *ibid.*

⁴⁸ DFAT, *KOMPAK Mid-Term Review* (2017). Replication across villages and sub-districts, and across districts, has been substantial. For example, delegation of authority from districts to sub-districts

(authority to monitor basic services delivery and to support village governments) has been replicated in seven districts and three provinces.

beneficiaries or target populations. Most investments provide sex-disaggregated data where feasible, but AQC reports yielded very few examples of monitoring the effect that the investment has on women beyond their basic participation. Similarly, the outcomes achieved for other disadvantaged groups (for example, people with disabilities, children, or the poorest of the poor) are rarely monitored, beyond their participation.

Evidence gathered for this evaluation suggests there is scope to improve engaging with and meeting the needs of external stakeholders. Half of DFAT and managing contractor survey respondents reported

that investments have developed monitoring systems that supply useful, reliable and timely information to stakeholders 'to some extent' while one-fifth considered stakeholder needs to be met 'to a small extent' or 'not at all'. About one-third of each group said that needs are met to a 'great' or 'very great' extent, which might be considered better practice. Analysis of AQC reports revealed limited mention of reporting to external stakeholders, with partner government being the most frequent reporting audience outside of DFAT.

KEY DETERMINANTS OF INVESTMENT MONITORING SYSTEM QUALITY

This section responds to key evaluation question 2. It describes the most significant determinants of better-practice monitoring, drawn from an assessment of aid investments delivered by managing contractors, followed by recommendations to improve the monitoring of Australian aid investments.

3.1 WHAT ENABLES BETTER PRACTICE?

This evaluation is primarily concerned with what DFAT and its managing contractor partners can do to achieve better-practice monitoring more consistently across aid investments.

This section identifies the key determinants of better practice and presents recommendations about how DFAT and managing contractors can strengthen monitoring quality.

The evaluation identified four main determinants of the quality of investment monitoring systems:

- » DFAT's own **performance culture** and expectations are the most important determinants of how effectively managing contractor teams monitor Australian aid investments.
- » DFAT's **ability to set objectives and maintain clarity** about what aid investments are meant to achieve is a critical pre-condition for better-practice monitoring.
- » DFAT's **demand for quality monitoring data** and the systems required to generate this data incentivise managing contractors to deliver better-practice monitoring. This demand is expressed through the department's policies, procurement and contracting processes, as well as the actions of staff. All of this incentivises managing contractors to deliver.

- » **Managing contractor responsiveness** to DFAT's demand for quality monitoring and their capability to meet this demand determines the quality of monitoring systems and the information they produce.

These four determinants are next described in greater detail.

DFAT'S PERFORMANCE CULTURE

DFAT's own performance culture is the most important determinant of investment monitoring system quality.

3.2 DFAT'S PERFORMANCE CULTURE

Performance management requires that a sound monitoring system be in place to provide data to inform decision making. Accordingly, DFAT's culture of, and expectations for, performance management can be a powerful driver for managing contractors to invest in monitoring.

An organisation's performance culture is the mix of shared vision, results expectations, operational tools and workplace behaviours. These define and reinforce success for the organisation's performance against its expressed objectives. The objective of a performance culture is to generate realistic and robust information on performance, drive quality and effectiveness, and strengthen accountability. This includes the adequacy of M&E systems, culture of contestability, being informed by evaluations, and adapting to ongoing learning.⁴⁹

DFAT's performance culture is the most significant quality-enabling factor identified by this evaluation.⁵⁰ It has also been identified as an area requiring strengthening in DFAT's response to the Aid Program Health Check.⁵¹ The norms and practices connected to investment monitoring form part of DFAT's performance culture. They contribute to the broader departmental culture as well as reflect it.

The evidence base for this finding

There is evidence that DFAT's performance culture strongly influences how well its managing contractor partners monitor aid investments. The literature review at Attachment A found that organisational culture is a critical determinant of how well

monitoring systems function. Although DFAT partners conduct most monitoring on its behalf, DFAT's performance culture sets the expectations against which these monitoring systems deliver. Interviews with managing contractor staff and M&E consultants strongly supported this finding. A culture within DFAT that encourages or at least permits learning from failure, and where senior leaders champion monitoring as part of broader performance management, supports better-practice monitoring by DFAT's managing contractor partners.

Strong leadership on M&E is critical to overall success in any organisation. If the leaders, managers and decision makers within DFAT understand and prioritise using the department's own standards and policies and themselves engage in good processes, then those working under them are more likely to follow suit.

Managing contractor staff survey, free-text response

The evaluation produced several more strands of evidence that suggest DFAT's performance culture is a strong driver for quality monitoring. Investments with better-quality monitoring systems typically have a performance management system in place, which the monitoring system serves.⁵² DFAT's performance expectations sit at the apex of investment performance management systems. For example, in AIP-R and KOMPAK, DFAT staff (working with partner government and managing contractors) set the objectives that form the basis of the performance management system. The managing contractor team then builds a monitoring system to collect

⁴⁹ This definition of performance culture is drawn from a mix of sources including DFAT external and internal documents.

⁵⁰ Performance culture was a consistent theme and the singular most important determinant of investment monitoring quality as reported through interviews and surveys.

⁵¹ The fourth objective in the August 2017 Aid Program Health Check is to 'further improve our performance culture, building on what is already one of the commonwealth's strongest systems.'

⁵² Based on case study analysis.

information about performance against these expectations. Without the signal of clear performance expectations from DFAT, such a system is unlikely to be built. The case studies of S4IEG and SGP are best illustrations of this.

AIP-R demonstrates how better-practice monitoring is supported by a performance culture built between DFAT and the managing contractor team. Both parties invested significant effort at the design stage to clarify the results framework, around which the monitoring system is built. The managing contractor team felt able to request additional resources for detailed results planning, and DFAT responded with additional resourcing. This kind of interaction might not occur in the absence of a mutual performance-oriented culture.

The SGP is an example of where the performance culture across DFAT and other government implementing partners was not sufficient to enable better-practice monitoring. The parameters for a monitoring system—including clear objectives, a realistic theory of change, and clear responsibilities—were not in place. DFAT staff felt constrained to address these weaknesses.

Case study analysis

Current practice and areas for improvement

Overall, the evaluation finds that DFAT can significantly improve the incentives for managing contractor partners to invest in better-practice monitoring by continuing to build upon its own performance culture.

DFAT staff were invited through the survey to assess the extent to which sufficient leadership support

within the department creates a performance culture, encouraging staff to use monitoring data for decision making. This question tests how well DFAT's culture supports the use of monitoring data (identified as a characteristic of better practice in Section 2.4). Eighty per cent of respondents identified room for improvement against this measure⁵³, with only one in five reporting this is true to a 'great' or 'very great' extent. This indicates that support from senior leadership to prioritise higher the performance management of aid among competing priorities will be critical to enable change.

Managing contractors want a culture that rewards open and genuine participation in learning and reflection processes and engages with monitoring data to understand investment performance.⁵⁴ Where DFAT does not promote such a culture, there is a risk that managing contractor teams will default towards agreeing with DFAT perspectives or defend the investment against critical feedback that may serve to improve performance. A culture that does not allow for a collaborative and constructive approach to reviewing performance will tend to close space for dialogue. This can undermine confidence for partners to openly reflect and learn within their own teams.

Interviews and the evaluation team's review of program monitoring documents indicate that program innovation inevitably involves risk of various kinds. In this context, information from monitoring systems helps to set the scene for innovation. It can push the boundaries or challenge performance that appears to be floundering or is just good enough. A manager needs to understand this information to know what is working and what is not. Monitoring information that is reliable, empirically based, relevant to program objectives and clearly conveyed to users can be a prime resource for a manager seeking to innovate. Therefore, improvements in DFAT's monitoring practice represent an opportunity to deliver against the department's innovation

⁵³ Responses to this question were notably less positive when compared to most other survey questions.

⁵⁴ Based on interview responses from most managing contractor teams combined with results from the managing contractor survey findings.

agenda, while managing risk in a better-informed manner.

Results from the managing contractor survey demonstrate desire for more engagement by DFAT on performance. Just over half of survey respondents indicated perceptions that:

- » DFAT senior leaders should be more engaged with the performance management of aid investments
- » DFAT leadership commitment to high-quality monitoring of investments, including the use of monitoring data to make decisions, is not strongly evident.

Barriers to enhancing performance culture

The importance that DFAT places upon performance and results is not conveyed as clearly as might be desired throughout the investment management cycle. Hard check points, such as design approval processes and mandated reviews, including mid-term reviews, draw attention to an investment's results and performance management system. For most investments, however, the level of attention dedicated to these wanes significantly in-between.⁵⁵

In-depth interviews illustrated a tension between staff attention dedicated to managing aid and responding to political and diplomatic priorities (with the latter characterised as both more urgent and more visible to senior managers). Some officers, particularly locally engaged staff, were dedicated primarily to investment management and reportedly insulated from more 'responsive' work. However, most staff at post, particularly Australian-posted officers, reported that they respond first to diplomatic, political and organisational priorities, and have limited time to focus on longer-term issues of aid effectiveness.

Given the hierarchy of DFAT staff reporting and the reported tendency for staff to prioritise diplomatic concerns over development matters, some staff have found it difficult to advocate internally for more focused attention on the quality and use of investment monitoring information. This is the case even though value-for-money is an explicitly recognised foundational criterion for the department's work.⁵⁶

Interviews with managing contractor staff and M&E consultants also suggest significant variation in DFAT's performance culture between divisions, branches, sections and posts. The relative emphasis that senior staff place upon aid performance management is reported to 'set the tone' for staff working within that area of DFAT. This is consistent with literature review findings that senior leaders play a vital role in championing monitoring as part of a broader performance management system. Where this system is given lower priority, DFAT's expectations for monitoring are downplayed.

There are strong natural incentives for managing contractors and DFAT staff to emphasise the achievements of aid investments, rather than communicate a more nuanced performance story. These incentives are exacerbated when the briefing and reporting load for DFAT staff outweighs the time available to invest in actively monitoring and managing investment performance. One method to counteract this is for investment managers to carry out routine field monitoring visits as described in the Aid Programming Guide. This should reinforce program accountability and help to solidify communications with local counterparts.⁵⁷

Fostering a safe-to-fail environment that supports adaptive management counteracts perverse incentives to only report the good news. In-depth interviews revealed it is not always in the commercial interests of managing contractors to report poor

⁵⁵ Independent M&E consultants and most managing contractor teams interviewed reported this.

⁵⁶ Ensuring value-for-money is one of the 10 strategic targets for the Australian aid program. DFAT, *Performance of Australian Aid 2016–17*.

⁵⁷ For a recent perceptive review of how to ensure quality and productivity from site visits. Michael Quinn Patton (2015), 'Evaluation in the Field: The Need for Site Visit Standards', *American Journal of Evaluation*, vol. 36, issue 4, pp. 444–460.

investment performance. This can undermine objectivity in investment performance reporting from the managing contractor. If DFAT explicitly recognises that some elements of program performance may be beyond the control of the managing contractor, this can encourage a safe-to-fail environment that can better support objective reporting and collaborative management of the investment.

Building upon reforms that are underway

DFAT's response to the Aid Program Health Check includes plans for the Secretary and Senior Executive to 'reinforce the importance of strong departmental performance culture in conversations with staff and with Head of Missions and First Assistant Secretaries. Such championing by senior leaders is recognised in the literature as a highly influential factor for establishing and maintaining organisational performance culture.⁵⁸ If implemented, the evidence suggests that these conversations will be key enablers to strengthen results-based aid management.

To improve the quality of performance management processes and outcomes in any organisation, there needs to be a clear articulation of what effective performance management is. DFAT has several established organisational policies and practices on performance management in response to the *Public Governance Performance and Accountability Act 2013* (Cwlth).⁵⁹ As DFAT policies and practices continue to respond to the Act, this is likely to support a higher standard of monitoring practice and the use of information to inform programming decisions.

Recommendations

Although performance culture may at first appear to be an amorphous area of departmental management, it can be measured and strengthened, and the use of applicable progress indicators—to the extent they can be adapted to DFAT's varied needs—can help make a performance culture initiative more readily understood and carried out. The Aid Program Health Check has already identified some of these. At the investment level, progress indicators could include the percent of investment managers who:

- » conducted field monitoring visits at least twice per year
- » reported use of M&E technical support, from sources at investment level or beyond
- » reported that their immediate manager is receptive to performance information, both positive and negative
- » can cite an example of communicating bad news or learning from failure
- » can cite at least one time over the previous year when performance information underwent a contestability challenge
- » can cite at least one time when performance information led to a substantive DFAT management decision.

Within this context, it is important to keep in mind that while investment managers are at the front line in managing performance of investments, they are not alone in this responsibility. Broader accountability, including for contestable performance reporting, is shared at multiple managerial levels.

DFAT leadership publicly recognises the importance of a performance culture. The gap identified in the evaluation is the completeness of leadership follow-through in support of this culture. Program staff have

⁵⁸ Refer to Attachment A. For example, Lahey 'A Framework for Developing an Effective Monitoring and Evaluation System in the Public Sector—Key Considerations from International Experience' cites leadership as one of 12 critical factors needed to launch and sustain an effective M&E system.

⁵⁹Department of Finance: <http://www.finance.gov.au/resource-management/pgpa-legislation/>

indicated a desire for stronger demonstrated leadership commitment to high-quality monitoring, and the evaluation has observed considerable variation in performance culture across DFAT divisions, branches, sections and posts.

Recommendation 1: DFAT to promote consistent and robust investment level monitoring, and the performance culture to support this monitoring, across the aid program.

A consistent and robust approach to monitoring investment performance across the aid program will drive overall quality and effectiveness and strengthen accountability. Stronger and more reliable measures of aid investment-level performance will have the added benefit of helping support DFAT's broader performance culture and performance reporting systems.

The second recommendation concerns the consistency of DFAT's internal messaging about aid performance management, and the role of better monitoring practice in supporting this.

Consistent communication on performance management expectations to DFAT staff could improve demand for and use of information. DFAT staff at post are commonly aware of the corporate M&E standards. External stakeholders, however, reported lack of consistency in how DFAT staff at post apply the standards when assessing the quality of important documents such as investment design documents, M&E plans and progress reports.⁶⁰ Given staff knowledge of the M&E standards⁶¹, inconsistency in their application is more credibly attributed to internal messaging about their relative importance compared to other priorities, especially for time-poor staff at post.

In practice, monitoring is seen by stakeholders as something that is undertaken by DFAT partners and

overseen by DFAT investment managers.⁶² The role of DFAT staff in monitoring and engaging with investment results is underemphasised. Opportunities to advance DFAT's broader priorities of supporting performance, diplomacy and strategic objectives may be lost if DFAT staff do not see active engagement with monitoring as central to their core responsibilities.

Recommendation 2: DFAT to strengthen its communication of the role of monitoring in supporting performance, diplomacy and strategic objectives.

DFAT should better promote the purpose of investment monitoring and its role in enabling the department to realise a range of broader objectives. This is to incorporate the importance of monitoring information for managing investment risk, supporting innovation, and adaptive management. This could occur, for example, through a letter of intent from the Secretary to posts and programs.

Revised DFAT guidance could more clearly embed the establishment and used of quality investment monitoring systems into the routine of investment managers. At each stage of the investment cycle, a to-do list for managers with responsibilities for senior staff could be identified (Attachment B outlines the proposed DFAT toolkit for investment managers).

⁶⁰ As reported by independent M&E consultants and most of the managing contractor teams interviewed. This was also a consistent theme that emerged from the free text responses to the managing contractor survey.

⁶¹ DFAT staff interviewed for the case studies were aware of the M&E standards. Many referred to them without prompting. The DFAT staff survey also supported this finding.

⁶² Interviews with DFAT posts, managing contractors, M&E consultants, and responses to the managing contractor survey.

CLEAR STRATEGY AND EFFECTIVE DESIGN

DFAT's ability to be clear about what aid investments are meant to achieve is critical for better-practice monitoring.

3.3 STRATEGY AND DESIGN

This section considers the importance of strategic clarity for better-practice monitoring. It refers to clarity both with investment objectives and broader diplomatic and foreign policy objectives. Strategic clarity is therefore a product of DFAT:

- » knowing and specifying its own high-level policies and objectives
- » translating these into country-level objectives and strategies in the form of Aid Investment Plans
- » identifying aid investment objectives at concept stage and the broader objectives the investment is expected to contribute to
- » leading a design process to produce a credible, evidence-based plan about how the investment will contribute to intended objectives, and specifying what results it will be accountable for.

This is complemented by the managing contractor translating the design into an operational investment based on a realistic theory of change and performance monitoring framework.

This description of strategic clarity illustrates that managing contractor partners play only a part in the process. DFAT has ownership over more of the process, and consequently plays a vital role in maintaining clarity once the managing contractor team is contracted.

The evidence base for this finding

The evaluation literature review identified clear objectives as a pre-requisite for an effective monitoring system. All better-practice investments reviewed in this evaluation made significant efforts during design and inception, and sometimes well into implementation, to clarify what is expected of the investment.⁶³ Often, objectives are not clear from the outset.

This evaluation sought feedback from DFAT and managing contractor staff about several aspects of strategic clarity, and the effectiveness of design processes and products in articulating this. Surveys revealed that half of DFAT respondents considered there is room for improvement in the clarity of DFAT's strategic vision for aid investments. Just over half of managing contractor staff agreed.

Between 30 and 40 per cent of DFAT and managing contractor respondents reported that DFAT's management decisions are 'to a great extent' consistent with investment objectives and theories of change, and that DFAT's rationale is clear when it changes investment objectives. This leaves more than 60 per cent of DFAT and managing contractor respondents believing there is room for improvement.

Interviews with DFAT staff identified incentives and constraints that drive weaknesses in strategic clarity and the quality of designs, which in turn limit the development of monitoring systems. M&E practitioners who work regularly with DFAT corroborated what DFAT staff identified, both

⁶³ AIP-R had a lengthy design phase through which results were identified. FCDP completed the M&E framework two years into implementation. PJSVP clarified objectives inherited from a prior phase of implementation and spent significant time building an M&E framework with partners. CAVAC was still working to define its

poverty measures after its mid-term review. KOMPAK, a design and implement facility, built a results system over the first three years of implementation.

through interview and free-text responses within the managing contractor survey.

Current practice and areas for improvement

Since at least 2005, investments have been required to have an M&E framework at design. Quality requirements at design are specified in DFAT standards, including the Investment Design Quality Standards and DFAT M&E standards. In summary, DFAT requires that an investment design should:

- » clearly identify what the investment is designed to achieve (end-of-program outcomes)
- » produce credible data and evidence, including of outcomes
- » plan for the use of data and evidence to inform investment decision making and accountability
- » include plans, where relevant, to strengthen the capacity of partner monitoring systems
- » enable monitoring of gender equality related aspects of the investment
- » provide sufficient resources for the above.

In practice, sources report that monitoring details are usually minimal at design.⁶⁴ A number of reasons have been identified for this.

First, DFAT is often not clear enough at the design stage about what the investment is expected to achieve and how.⁶⁵ For some investments, including ‘design and implement’ investments and facilities⁶⁶, DFAT decides expected outcomes and activities during the inception phase. For other investments, the design process and design document fall short of specifying intended outcomes and how these will be realised (theory of change). Still for other investments, there is little knowledge at the design

phase about data availability and the practicalities of data collection. This means the monitoring system cannot be fully designed or needs to be redesigned before implementation. Without clarity at the outset, it is difficult to formulate key indicators around which to build the monitoring system.

Second, the advisers who undertake design work for DFAT rarely specialise in monitoring and may therefore underestimate the resourcing, expertise and effort required to develop a monitoring system.⁶⁷

Third, during the design process DFAT does not demand detailed information about how the monitoring system will be developed.⁶⁸

DFAT’s strategic clarity

Lack of clarity on what objectives or outcomes investments are expected to achieve is a critical factor diluting the quality of monitoring and broader performance management. This is true, most critically, for theories of change within investments. It also applies to causal links between investment upper-level outcomes and Australian strategic goals at country, regional, sectoral and thematic levels.

Capacity of DFAT investment managers to link investments to strategic outcomes is critical to articulate a clear strategy and expectations. The time and expertise needed for attention to strategy is often underestimated, and delegated staff may not be well prepared for this task. To be skilled users of monitoring information, staff do not need to be experts, but they would benefit from having sufficient knowledge to know what is needed for an investment and what to ask of a managing contractor for it to be delivered.

⁶⁴ DFAT, *Improving Quality of Investment Designs*, 2018, internal working document. Two of the sample of 10 investments had complete and adequate M&E frameworks at design. Two more developed complete and adequate M&E frameworks during implementation. The remaining six have only partially completed M&E frameworks during implementation, which do not yet meet the design standard. This was consistent with findings from interviews with managing contractors and independent M&E consultants.

⁶⁵ *ibid.* Confirmed in interviews conducted for this evaluation.

⁶⁶ DFAT defines facilities as an aid delivery mechanism that provides flexible (adaptive and responsive) services managed in an integrated way. Objectives (or end-of-facility outcomes) are specified (during design and/or inception) but the pathways to deliver them are left unspecified and developed during implementation.

⁶⁷ Based on interviews with independent M&E consultants.

⁶⁸ DFAT, *Improving Quality of Investment Designs*, 2018, internal working document. M&E was identified as the weakest area from a sample of 10 designs.

Aid investments are usually contributing to foreign policy and diplomacy objectives, but only recently have DFAT staff begun to explicitly articulate these. When these objectives are not written as part of the design and do not inform the performance system, there is a risk that the investment may appear to perform poorly by its own measures while still meeting DFAT's foreign policy and diplomacy objectives. The SGP is a good example of this.

A robust design process and product

In the absence of a firm position on how designs must address monitoring, the quality of content varies significantly. Some investments have an M&E framework at design, but most do not. Other research has suggested that the key measure at the design phase should be that end-of-program outcomes are evaluable⁶⁹ with adequate processes, resources and expertise identified to further develop the monitoring system.⁷⁰

The governance of the monitoring system is a key choice to make at design, and one that is rarely considered in detail. Typical designs require the managing contractor to staff the monitoring function with their project team. Usually, this comes in the form of one M&E adviser and some locally engaged M&E staff. However, other models are available which can increase the incentives of the project team to invest in monitoring as a priority. Examples include the presence of an independent M&E function such as Buka Hatene (Timor-Leste M&E House), the independent monitoring function of DFAT's CSO Wash Fund, and the recently established Quality and Technical Assurance Group of the PNG Governance Facility.

Investment design activities need to more systematically consider proposed approaches to monitoring, with an initial investigation of the implications of the investment theory of change for

performance measurement. Those engaged in investment design need to have ready access to examples of high-quality monitoring and additional M&E technical resources to inform specification of the monitoring system. The investment's monitoring architecture will vary from one setting to another, so flexibility is needed in transitioning from design to a fully operational monitoring system.

Establishing a monitoring system

The inception phase translates design parameters into a monitoring system. This is when the investment architecture is being put into place, including partnerships, governance structure, decision-making bodies, staff, inputs and activities. For many investments, there is a significant shift or evolution from what was conceived at design as it becomes more apparent what is and is not feasible.⁷¹ This period presents a risk of strategic drift, where the investment may focus more on what is achievable at the expense of what is difficult, but strategically important. Some adjustment of objectives is often called for. However, this must be intentional and in a partnership between DFAT and the managing contractor team so implementation ultimately supports achievement of the strategic intent and objectives for the investment.

If a significant effort is simultaneously invested into developing a monitoring system during this period, the system is much more likely to connect with, reflect and support the intervention. On the other hand, if a sound monitoring system is not established during inception, poor monitoring can persist indefinitely. It may take a 'hard gate', sometimes as significant as reaching completion of the investment and design of the next phase, to force resolution of monitoring issues if DFAT staff do not manage for this during implementation.

⁶⁹ Evaluable outcomes clearly identify how success can be measured, for example, the intended number of beneficiaries, or the quantum of change (for example, in incomes and behaviours).

⁷⁰ DFAT, *Improving Quality of Investment Designs* (2018), internal working document.

⁷¹ Interviews with DFAT posts, managing contractors and independent M&E consultants.

The SGP case study is a good example of how and why this can occur. It also illustrates how the core elements of design that underpin the monitoring system can remain in flux over an extended period, despite DFAT staff being aware of this gap.⁷² The S4IEG case study illustrates how overstretched managers at post have known about, but have not been able to adequately prioritise, resolving weaknesses in monitoring due to competing issues. Issues include staff shortages in both DFAT and managing contractor teams, environmental disasters, ministerial visits and a change of government. Such competing issues cannot be removed, however a firmer requirement for monitoring to meet DFAT standards and a checkpoint to assess this would give senior staff the information and incentive to prioritise. The KOMPAK case study illustrates how independent technical review during inception can provide DFAT with the information needed to prevent monitoring weaknesses from persisting.

KOMPAK's monitoring system has been influenced by several reviews during inception. These included two AQC rounds, an independent consultants' review, a field-based M&E assessment, and a mid-term review. This meant more independent scrutiny than is ordinarily the case for DFAT investments during inception, which led to monitoring issues being identified and addressed. The mid-term review recommended that KOMPAK consolidate its program theory into a single, simplified framework, and invest in a more sophisticated management information system. It also suggested ways to improve monitoring tools. A mid-term review may recommend similar for any investment; the difference is that KOMPAK is just beginning to implement and is flexible enough to respond.

Case study analysis

Building upon reforms that are underway

The Aid Program Health Check response indicates that senior DFAT staff will discuss the department's strategy to realise foreign and development policy objectives more regularly and openly with staff. This legitimises the conversation at investment level and enables DFAT staff to explain clearly to managing contractor partners what is expected, and to identify key performance indicators that reflect the intent of the investment. Strengthening internal communication on strategic matters increases in importance with the emergence of DFAT integrated strategies and the incoming set of new Aid Investment Plans.

A key instrument for adding quality to the design process is the recently revised template for the Design Approval Minute. The design approval checklist includes the need to validate that expected end-of-program outcomes and results are clearly described and underpinned by a sound program logic and detailed M&E framework. The evaluation team notes that it may be helpful to add that the M&E framework needs to include initial performance indicators with baselines and periodic performance targets, clearly linked to program outcomes.

In addition, among the Aid Program Health Check initiatives associated with improvements to investment design are:

- » piloting a Quality Assurance Unit
- » reinforcing peer review and independent appraisal processes
- » revising design guidance to drive more consistent design formats and greater clarity of objectives
- » preparing terms of reference for an annual DFAT investment design award.

⁷² Challenges related to strategic purpose and stakeholder demand are among common factors that the evaluation found prevent DFAT

staff from requesting higher-quality monitoring systems from managing contractor partners.

All these actions present opportunities to strengthen investment clarity which would support greater coherence and efficiency in investment monitoring.

Providing more focused guidance and training on the adequate inclusion of monitoring into investment designs can be helpful but these are unlikely to be sufficient to meet existing need. Improved awareness and follow-through on the DFAT side needs to be complemented by a more standardised focus on monitoring requirements in managing contractor contracts.

Recommendations

Instituting a check point and expectations for the monitoring system during inception could support its satisfactory establishment in a timely way. Without a check point, work to embed and improve monitoring is competing for management attention against more urgent and pressing priorities. This is true both on the DFAT side and the managing contractor side.

For the sake of efficiency, the check point should remain as straightforward as possible. An appropriately skilled external consultant or DFAT staff member⁷³ might review the monitoring system against DFAT's current monitoring standards (and monitoring systems used by similar investments) to identify critical gaps, concerns or weaknesses.

Recommendation 3: DFAT to institute a check of monitoring system quality during the investment inception phase and invest resources to ensure DFAT M&E standards are met.

DFAT should check investment monitoring systems towards the end of the inception phase. The timing can be flexible, depending on the type of investment. The purpose is to ensure that DFAT's monitoring standards are satisfied early in implementation, and to identify any outstanding weaknesses in the monitoring system leading to corrective action.

DFAT should complement this with a review of monitoring guidance and contracts discussed in recommendations 2 and 5 so investment managers, delegates and managing contractor teams have the information and tools necessary to prepare to meet this check point.

⁷³ Relying on DFAT staff to undertake such a review would require internal monitoring expertise to be built and/or shared. In the first instance, outsourcing of this function but retaining strong control

over the process and standards applied, would appear more feasible.

DEMAND FOR QUALITY MONITORING

DFAT's demand for quality monitoring data and the systems required to generate this data are expressed through the department's policies, its procurement and contracting processes, and the actions of staff, all of which incentivise managing contractors to deliver.

3.4 THE ROLE OF DEMAND

This section addresses DFAT's demand for quality monitoring information, and the systems required to deliver it. There is substantial literature on the role of demand-side factors in determining the quality of monitoring data and systems. A brief synthesis from the literature is included in this section and more details are in Attachment A.

Ownership is another way to think of the demand part of the equation. Ownership must come from those at every level who use a monitoring system, and the specific demands for performance information at each level needs to be identified. If there are levels where people do not see the need for, or have a use for, the data collected, there will be problems with quality control and ownership. The feedback loop will be disrupted. The system will degenerate, and the quality of data will decline.

A strong champion can help to ensure ownership of the system, in part through driving demand for the information it produces. A champion can stress that reliable performance data must be generated, shared and properly reported. If, on the other hand, demand is episodic or haphazard, a monitoring system will not be used and sustained. Structured requirements for reporting results, including those imposed by legislation, regulations and international development requirements can help lead to sustained, consistent demand for such systems.

In DFAT's case, demand can be stimulated when the strategic goals of Aid Investment Plans are translated into monitoring systems, when senior staff within DFAT or the managing contractor team demand meaningful performance information directly, or when other partners, such as the partner government, demand evidence of results.

This section also indicates specific reforms to procurement and contracting that the evaluation

recommends DFAT undertake to better signal to managing contractors the importance of investing in capacity for quality monitoring.

Evidence for these findings

Managing contractor and DFAT staff were asked to assess the extent to which 'there is adequate time and capacity within DFAT so quality monitoring systems are established, and to make effective use of the information provided'. These questions represent DFAT's demand for quality monitoring information and systems. The responses to these questions are one of the most telling results from both surveys:

- » Only 6 per cent of managing contractor respondents reported that DFAT has sufficient time and capacity to ensure that high-quality monitoring systems are established. Just over 80 per cent of managing contractor respondents indicated there is room for improvement against this metric.
- » DFAT respondents broadly agree with this assessment, with just 13 per cent reporting they have enough time and capacity to ensure that high-quality monitoring systems are established. Eighty-seven per cent of DFAT respondents considered there is room for improvement against this metric.
- » The responses assessing DFAT's capacity to make effective use of information are broadly similar. Six per cent of managing contractor respondents and 8 per cent of DFAT respondents report that DFAT has the time and capacity to use information effectively. A total of 79 per cent and 94 per cent respectively disagree to 'some extent', or to a 'great extent'.

Interviews corroborate these results, with many managing contractor respondents indicating uncertainty about the usefulness of their reporting to

DFAT. Several express doubts that most performance reporting is read. DFAT staff viewed reports as often being too long and onerous, and insufficiently focused on progress against annual plans. In providing feedback, DFAT was reported to focus too often on grammatical correctness, rather than engaging with substantive matters of performance. This kind of feedback indicates a breakdown in DFAT staff ownership of investment monitoring systems, or at least a failure of demand to appropriately stimulate supply.

Current practice and areas for improvement

DFAT formal policy and guidance are supportive of the effective use of monitoring data. The M&E standards and other sources provide a clear basis for this. Many DFAT staff and others working with DFAT appreciate the utility of the standards.⁷⁴ Applying the standards requires access to technical knowledge, support and guidance, confidence in providing feedback to managing contractor staff, and the use of good judgement to know when to be flexible and how to apply the standards in a range of circumstances.

The department has been weaker on the follow-through needed to ensure policies are translated into practice. Many sources noted that the standards are overly generalised to be of proper help to an investment manager, and that they could use a refresh.⁷⁵ DFAT also provides staff with M&E training; in-depth interviews with DFAT staff and DFAT survey free-text responses suggested that this training needs to focus on the specific tools, methods and resources an investment manager requires to carry out active, technically informed oversight of an investment's monitoring system.

The good-practice investments reviewed in this evaluation have typically structured learning opportunities into the investment management cycle. Such opportunities may take various forms, such as periodic and mid-term reflection sessions, or

thematic workshops using monitoring data as a starting point. The learning sessions can serve as platforms for deeper understanding of program dynamics and challenges, and secondarily reinforce capacity strengthening with local partners.

The evaluation found limited reference to the M&E standards and other performance policies in existing templates, guidance or professional development materials. Officers also reported that the value placed on performance monitoring and reporting by senior managers varies considerably.⁷⁶ Building performance management expectations more comprehensively into DFAT's existing systems, including professional development and career progression, could reinforce the priority placed on meeting existing policies.

Aid management is a core career anchor in DFAT, and managing data and results reporting needs to be more readily known as a key element within aid management. Findings from the survey of DFAT staff and interviews with managing contractor staff and M&E consultants who work with investment managers indicate they are often insufficiently equipped to carry out these tasks in an informed way. There is a need for more broadly and intensively ensuring that basic smart-user knowledge and skills in monitoring are prioritised in staff performance agreements.

There is need, then, for improvement in both the expectations for the kind of monitoring delivered through investments and the department's capacity to improve monitoring quality. In examples of good DFAT monitoring practice, investment managers are sufficiently skilled so they can be responsible users of monitoring information and are given clear (if often informal) guidance on how to support good practice. These staff demonstrate confidence in what to ask for from a managing contractor team in terms of performance monitoring, and an ability to engage

⁷⁴ Based on interviews with DFAT posts, managing contractors and independent M&E consultants.

⁷⁵ Interviews with DFAT Canberra and independent M&E consultants.

⁷⁶ Consistently reported during interviews with DFAT posts.

with the managing contractor team on the meaning and implications of the content.

Finally, incentive structures, within DFAT and beyond, play a role in potentially constraining efforts to manage investment monitoring better.

Investment managers, who often are locally engaged staff, may for institutional or cultural reasons be hesitant to challenge the status quo regarding how monitoring information is used. Australian posted staff may hold government-to-government relationships as a priority, and place relatively lower priority on monitoring, often due to competing pressures.

Contracting and procurement

The survey indicated that reform to procurement and contracting arrangements may provide some quick wins for DFAT. Only one-quarter of DFAT respondents strongly agree that contracting and procurement systems support the establishment of quality investment monitoring systems, and that contractual arrangements create the right incentives for quality monitoring and management of investments. Managing contractor respondents gave a broadly similar assessment, indicating agreement between DFAT and its managing contractor partners that improvements can be made.

Clarifying language in managing contractor contracts to reinforce the importance of quality monitoring is a step DFAT can take to encourage appropriate and consistent levels of attention to monitoring by managing contractor partners. Several DFAT staff indicated at interview that contracts did not necessarily have the levers in place to enforce expectations of quality monitoring or reporting. For example, some contracts include reference to the M&E standards, and others do not. Similarly, several managing contractor staff from case study interviews reported variation in the content of contract clauses, and even more significant variation in the extent to

which DFAT staff refer to and use the contract as a management tool.

The evaluation found that approaches followed by managing contractors to establish monitoring systems were variable. This is reflected by the extent to which managing contractors have institutionalised approaches to knowledge management on monitoring systems and the management of M&E practitioners. While recognising there are constraints to DFAT demand (as set out in previous sections), managing contractors report investing internally in improving their own institutional approaches to technical knowledge on program performance monitoring.⁷⁷ Such actions would help ensure that DFAT's intentions for improved monitoring will readily be understood, and that managing contractor efficiencies may be gained through accumulating and sharing knowledge and skills.

Barriers to effective demand

The evaluation identified challenges in supporting staff to oversee monitoring as a barrier to effective demand for quality monitoring.

Staff understanding of monitoring systems

Knowledge, skill and motivation of investment managers drive demand for and use of monitoring information. Overall, investment managers are highly motivated and display high levels of personal commitment to carrying out effective performance management.⁷⁸ The capacity of DFAT staff to absorb new information and perform new tasks within their already busy work demands is a crucial factor that influences current capacity levels.

Investment managers at Post are at times well equipped to oversee investment monitoring systems and products, but more commonly they are provided little preparation for this work and almost no

⁷⁷ Managing contractors interviewed for case studies reported internal investment in improving technical knowledge on monitoring being driven from leadership within their respective organisations.

⁷⁸ The majority of interviews with DFAT posts displayed this commitment to effective performance management. This was corroborated with findings from the DFAT and managing contractor surveys.

technical support should they seek it out.⁷⁹ On DFAT's side, use of monitoring information tends to be overly satisfied with data on status of activities and outputs and good-news stories, with insufficient attention to status of intended outcomes.⁸⁰

Several DFAT staff interviewed reported having a limited understanding of the M&E systems that may sit behind a partner's reporting. Although aware of the M&E standards, staff were not always confident in applying them when undertaking assessments of the appropriateness of end-of-program outcomes given the resources and timeframes of the investment, and in assessing the likelihood that proposed interventions will achieve expected outcomes.

Strengthen engagement between DFAT staff, contractors and development partners to help strengthen data sharing and joint analysis and learning. Strengthen the capacity—times, skills and knowledge of DFAT personnel—to be able to genuinely engage and support M&E and learning in a constructive and helpful way—that is, be part of the M&E and learning process. The current emphasis tends to be on receiving information and accountability, rather than having the capacity to engage in meaningful and helpful conversations about M&E and learning methods and processes.

Managing contractor staff survey, free-text response

Results from the survey of DFAT officers reveal that just over half of DFAT respondents rated their experience and knowledge of monitoring theory and practice as strong. However, it is evident from the survey of DFAT officers and managing contractor staff that both groups hold some concerns about demand-side capacity. DFAT Canberra has some monitoring expertise, but not a critical mass with an

organisational mandate to support adequate use of monitoring systems.

Engaging with monitoring processes and information

Interviews with managing contractor staff indicate there is often demand from DFAT officers for the good news that a monitoring system can generate, but variable appetite for other types of information. The investment teams interviewed for the case studies were unanimous in their perception that DFAT welcomes, and sometimes actively demands, performance data and stories that can be used to effectively communicate what an investment is achieving in a way that is readily understood by the public. This kind of information serves DFAT's public diplomacy purposes well, and DFAT officers tend to both ask for it and engage with the content.

Managing contractor teams reported more variable appetite from DFAT staff to engage with monitoring information that is detailed and nuanced, or that indicates a need to adjust delivery of the investment or identifies failures. Similarly, managing contractor teams reported that DFAT officers were often not available to join field monitoring, or reflection and learning exercises, both of which could deepen understanding of the investment and its performance. This sentiment is reflected in responses to the survey of DFAT staff:

- » 13 per cent strongly agree that DFAT has 'sufficient time and capacity to ensure that quality monitoring systems are established', with more than 80 per cent indicating there is room for improvement.
- » 8 per cent strongly agree that DFAT has 'the time and capacity to use monitoring information effectively', with more than 90 per cent indicating there is room for improvement.

Some managing contractor teams had invested significant effort into building a relationship with

⁷⁹ Interviews with DFAT posts, managing contractors, independent M&E consultants and responses to the DFAT and managing contractor surveys.

⁸⁰ *ibid.*

DFAT staff that permits discussion of this kind of information (for example, AIP-R). Some posts reported proactively asking for discussion with the managing contractor team around investment performance and learning (for example, Fiji Post referring to FCDP). Review of AQC reports indicated that DFAT participation in field visits and learning and reflection activities had enabled meaningful engagement with the monitoring system and data.

Recommendations

To adequately support quality monitoring and manage risk there needs to be technical support for investment monitoring. To date, emphasis has been placed on staff training as a solution for this capacity gap. The evaluation found that training can be helpful, but that ongoing technical support is even more important and valuable.⁸¹ In line with the prioritisation of learning approaches in DFAT staff performance agreements (70/20/10: experience/exposure/education), staff capability in investment monitoring should be expected to occur in multiple ways but feature an emphasis on experience such as on-the-job learning.

Recommendation 4: DFAT to ensure that all investment managers have access to technical support to establish and oversee monitoring arrangements, especially for complex, technical and/or high-value investments.

DFAT should provide ongoing technical support to all investment managers through appropriate means (for example, Canberra or post-based advisers, access to panels of external expertise, or through a community of practice). This should be supplemented through a toolkit for investment managers, developed as part of the response to Recommendation 2. This toolkit should identify

the characteristics and determinants of better-practice monitoring.

DFAT should use externally supported quality assurance mechanisms where warranted. The need here is greatest for higher risk, large and/or complex investments.

DFAT staff responsible for managing aid investments should prioritise routine field monitoring while taking into account the value and complexity of the investment.

It is also important to recognise when external expertise is needed. Some investments are too large, complex, or too technically challenging for DFAT staff to appropriately oversee the monitoring system and quality assure its products. The need for access to third-party quality assurance is greatest for higher-risk, large or complex investments. TSSP provides a good example of where independent quality assurance is warranted. In cases such as this, DFAT should build independent technical assurance of monitoring into the governance of the monitoring system, and contract appropriate experts to serve this function.

Procurement and contracting systems

Contracting is the critical moment to ensure that the managing contractor team can be held to account for monitoring. DFAT and managing contractor staff broadly agree that DFAT contracting and procurement systems support the establishment of quality monitoring systems and create the right incentives for quality monitoring.⁸²

Interviews suggest, at the same time, that there is significant variance in contract clauses. For example, DFAT's M&E standards are explicitly referenced in some agreements as benchmarks, and not in others.⁸³ This variance is a lost opportunity to promote M&E standards across the portfolio of

⁸¹ Interviews with DFAT posts, independent M&E consultants and responses to the DFAT survey.

⁸² Just under 30 per cent of DFAT respondents believe this to a 'great' or 'very great extent'. Just over 60 per cent believe it to 'some extent' or a 'small extent'. The remainder do not believe it to be 'true' or they 'don't know'. The response pattern from managing contractor staff was similar.

⁸³ This observation is based on the substantial experience of the evaluation team with DFAT managing contractor contracts. A systematic review of contract language was not included in the scope of this evaluation.

managing contractor investments. The expectations of investment managers around progress reporting were also observed to vary widely.⁸⁴ The variance can lead to inefficiencies for both DFAT and managing contractors. DFAT staff report that monitoring systems do not produce all the information that DFAT needs⁸⁵ while managing contractor teams try to determine what their individual agreement manager requires. More clarity might occur through greater contract standardisation.

The translation of monitoring requirements in designs into procurement specifications warrants attention. Currently, DFAT designs often call for:

- » a single senior international expert
- » full-time presence in-country, prohibiting flexible arrangements
- » remuneration parameters.

These factors limit the available pool of credentialed M&E practitioners and reinforce the belief that M&E practitioners are all 'equal'. Evidence from interviews suggests that different designs, sectors and country contexts call for a differentiated set of M&E skills which may not all feasibly be embodied by one professional. This approach therefore has the potential to create gaps and mismatches between the skills required and the practitioners recruited.

Contracts often do not adequately establish a framework for transparent monitoring or performance remediation. This does not provide DFAT investment managers or managing contractors with a clear understanding of expectations regarding responses to unsatisfactory performance. Some managing contractors have expressed the concern that in performance assessments a clearer distinction needs to be made between contractor performance and overall program performance (over which the managing contractor typically has limited control). A well-designed monitoring system should provide data to inform assessment of both types of performance.

Recommendation 5: DFAT monitoring expectations to be better standardised across managing contractor contracts.

Standardising language in managing contractor contracts to reinforce the importance of quality monitoring can encourage managing contractors to pay appropriate and consistent levels of attention to monitoring. In addition to the referencing of DFAT M&E standards, this can also include better recognition in contracts of different DFAT information needs.

⁸⁴ Interviews with managing contractors, independent M&E consultants and responses to the managing contractor surveys.

⁸⁵ One-third of DFAT staff surveyed agreed to a 'great' or 'very great' extent that investments generally have monitoring systems that meet DFAT information needs.

MANAGING CONTRACTORS' RESPONSIVENESS

Good practice in monitoring is a two-way street. Effective DFAT guidance and collaboration are very important, but the ability of managing contractors to deliver a good monitoring system is also significant. Managing contractors well prepared to provide quality monitoring may realise a competitive advantage. The likelihood of managing contractors investing to build this capacity will be improved by DFAT signalling that it expects, and will contract for, quality monitoring services.

3.5 MANAGING CONTRACTORS

This section addresses the capacity of managing contractors to deliver better-practice monitoring systems and respond to DFAT requirements. As DFAT's most cohesive and arguably most responsive group of implementing partners, managing contractors are uniquely positioned to establish best practice for the monitoring of Australian aid.

With the clear emphasis on meeting client needs and improving the performance of investment monitoring systems overseen by managing contractors, some clear recommendations are within the remit of managing contractors to address to lift industry standards in the monitoring of investments overall.

The evidence base for these findings

Managing contractor staff and independent contractors who design, deliver and evaluate investments for DFAT were emphatic that the relationship between DFAT, as commissioner, and managing contractors, as suppliers, is key to quality monitoring. At interview, many managing contractor staff indicated that managing contractors receive mixed messages about the importance of resourcing and prioritising monitoring. For example, one experienced M&E consultant noted that M&E expertise is rarely sought by DFAT for design work, and therefore rarely represented on design teams, even though there are appropriate M&E consultants to engage on design work.

There was substantial agreement between managing contractor staff that when a clear message is received about what DFAT wants, this will be

prioritised, otherwise business will be lost. There was a strong perception that if the managing contractor did not respond to DFAT priorities, both those expressed in contracting and the more informal emergent priorities, then the investment will ultimately be lost to a more responsive supplier.

However, it is not always sufficient for DFAT to signal what it wants. There is also the question of whether managing contractors are able to deliver. It is evident from the survey of DFAT staff that there are concerns about the capacity of managing contractor teams to supply quality monitoring. Just over one-third of DFAT respondents who were surveyed are confident that managing contractor teams have the capacity to design and oversee investment monitoring systems. The remaining two-thirds reported that managing contractors have this capacity only to 'some extent', a 'small extent', or 'not at all'. Managing contractors made a very similar self-assessment in the managing contractor survey of their own capacity to deliver high-quality monitoring.

Interviews conducted for the case studies indicated that pressure to implement can contribute to an insufficient investment of time in monitoring, often reflected in less than adequate training for core managing contractor staff. S4IEG is a good illustration of how this can occur, with the core team narrowly focusing on establishing the program and having little time or capacity to focus on monitoring. The survey of managing contractors indicated that inadequate training to equip the team to develop and operate a quality monitoring system is a common weakness. Only 20 per cent of managing contractor respondents strongly agreed that training had been adequate to equip them for their roles in monitoring investments. Managing contractor teams

rated their access to technical support, the capability of their design staff, and the availability of sufficient resources for monitoring more positively. In short, core implementation staff need the time to be upskilled in monitoring. The AIP-R case study demonstrates this well.

Current practice and areas for improvement

Knowledge management across investments

Managing contractors were found to lack uniformity in the extent to which they invest in institutional knowledge management in support of consistent and forward-looking approaches to investment monitoring. In the survey of managing contractor staff, about one-third reported they have access to adequate technical support to meet their responsibilities in investment monitoring. Another third agreed only to 'some extent', and the remaining third 'disagreed'. This is reflected in the case studies, where some investments (AIP-R and FCDP after a delay) have sourced significant expertise externally, and others have developed the monitoring system in relative isolation (CAVAC).

This is important, because managing contractors often engage repeatedly in a sequence of investments in the same sector or country, and therefore have an opportunity to realise efficiencies by sharing monitoring technical knowledge across teams and individuals. In addition, risks associated with weak monitoring are not DFAT's alone; they are shared with partners such as the managing contractors. Recognising this, it would be especially helpful for managing contractors to complement investment delivery with internally based knowledge management in the M&E technical area. The evaluation identified examples, particularly among some case studies, of lost opportunities to transfer expertise and tools from similar programs which means more reinventing the wheel (for example, S4IEG).

Managing contractors have performance cultures. While the evaluation team did not pursue this issue in depth, it is apparent that some managing contractors are quite advanced in their capacity to manage and make use of technical knowledge on investment monitoring.⁸⁶ The concern elicited by this evaluation is the apparent inconsistency in this capacity from one managing contractor to another, or from one investment to another.⁸⁷ If managing contractors were encouraged to better manage technical knowledge on monitoring it likely would lead to more consistent and higher-quality monitoring across investments.

Recruiting and managing M&E practitioners

The value-add of M&E expertise has been slow to be fully recognised by DFAT and managing contractors. This is compounded by the difficulty reported by managing contractors in recruiting and retaining practitioners with appropriate M&E expertise. All stakeholders reported there is a variable pool of qualified and high-performing M&E practitioners in the job market. An example of this dynamic is found in the FCDP and S4IEG case studies, both of which progressed past 12 months of implementation before sufficient and appropriate M&E expertise was put in place.

The limited pool of practitioners is also driven in part by limited technical oversight and knowledge of best practice in M&E by managing contractor project management teams (team leaders, contractor representatives and program managers).⁸⁸ This inadequacy contributes to human resources gaps and/or poorly performing monitoring in investments. This situation is underpinned by the performance management and quality-assurance architecture of the managing contractor organisation more broadly.

Managing contractors can contribute to DFAT's need for simple, adaptable and readily understood monitoring systems by prioritising development and

⁸⁶ Interviews with DFAT posts, independent M&E consultants and responses to DFAT survey.

⁸⁷ *ibid.*

⁸⁸ As reported by a number of managing contractors and independent M&E consultants during interviews.

application of such systems within their technical portfolios. This is especially important when the managing contractor investment team lacks the time, resources or capacity to develop such systems.

Barriers to better practice by managing contractors

The findings suggest that despite evidence of performance culture among managing contractors⁸⁹, there appears to be wide variation in the technical readiness of M&E advisers and others on investment teams to adequately deliver good monitoring systems.⁹⁰ In addition, there is a similarly wide variation in the quality of technical and managerial backup for team leaders and M&E advisers in the field.⁹¹ This effectively can leave investment teams operating in siloes and insufficiently supported by corporate performance policies and processes.

Recommendations

These recommendations target managing contractors and propose how their organisations can take a proactive approach to improving investment monitoring. These changes can be adopted without DFAT requirements changing.

Recommendation 6: Managing contractors to nurture a corporate culture of performance, including by building new capability and encouraging a cohort of staff to develop and maintain M&E expertise.

Managing contractors to actively support M&E prioritisation and learning among their program staff. Actions to include are:

- » integrating monitoring activities into the

roles and responsibilities of all program staff and communicating these expectations through performance agreements

- » enhancing the role of informal mentoring to support staff in the application of monitoring standards and tools.

DFAT's increasing reliance on external expertise to design and quality assure monitoring systems means that efforts by managing contractor partners to provide more comprehensive quality assurance for monitoring will be of particularly high value.

Recommendation 7: Managing contractors to support simple yet adaptable monitoring approaches, strengthened by learning across investments.

Managing contractors should enhance their own management of monitoring knowledge by:

- » strengthening institutional processes and capacity to undertake better quality monitoring
- » ensuring the ability to apply, adapt and guide delivery of quality monitoring systems in new contexts
- » investing in ensuring simpler, more practical, fit-for-purpose approaches to investment monitoring, and informing prospective clients of this investment.

⁸⁹ Evaluation case studies on FCDP, AIP-R, PJSVP and KOMPAK demonstrate how managing contractor teams have brought their own performance culture to support development of the investment monitoring system. The case studies of TSSP and SGP are examples of where the managing contractor role is more limited within a broader investment, and the managing contractor team has not brought a strong influence on development of the system. Finally, S4IEG illustrates how the local team engaged by the

managing contractor has not fully benefitted from the managing contractor organisation's broader performance culture, and how this could have better facilitated development of the monitoring system.

⁹⁰ As reported by DFAT posts and a number of managing contractors and independent M&E consultants during interviews.

⁹¹ *ibid.*

ANNEX 1: METHODOLOGY

The evaluation took a utilisation-focused approach and as such was informed by the specific needs and priorities of each of these four key primary stakeholders:

DFAT senior management, who are responsible for promoting a robust performance management culture across the department. They are personally responsible and accountable for their decisions and actions to ensure that a proposed aid investment represents a proper use of Australian Government resources and meets legislative and departmental requirements. Desktop research that preliminarily identified the characteristics of good-practice monitoring systems, together with evidence from the consultations, informed the findings and options discussed through participatory recommendations workshops with senior management. The use of these findings and recommendations will enable DFAT senior management to ensure that future aid investments have better-practice monitoring systems.

DFAT investment managers, performance and quality staff. Investment managers are responsible for the design, implementation and M&E of investments, with support from performance and quality staff. As the primary audience anticipated to apply evaluation findings, participatory processes were undertaken to consult with and enable this group to use the findings and subsequent recommendations and guidance material to identify practical steps for DFAT to realise better-practice monitoring systems.

Contracting and Aid Management Division, which is responsible for developing DFAT policy and guidance on investment design, M&E and procurement services, including the quality assurance and clearance of procurement and contracting approaches. Consultation and desktop research on procurement and contracting mechanisms informed:

- (a) understanding of the constraints to quality that may derive from contracting policies or procedures
- (b) prospective responses to such constraints.

Options discussed through the participatory recommendations workshops will enable this audience to use the evaluation findings to consider procurement and contracting approaches⁹² that support and complement better-practice investment monitoring systems, and to improve DFAT's approach to design and investment monitoring and the quality of support.

Managing contractor companies and M&E consultants are involved in the design, delivery and M&E of the Australian aid program. Participatory processes were undertaken to consult with and enable this audience to use evaluation findings to fine tune and adapt the way they work to better meet investment monitoring standards and expectations.

DFAT's Independent Evaluation Committee served as a technical reference group and was responsible for quality assuring the design and conduct of the evaluation. ODE convened an internal Reference Group comprising representatives from each of DFAT's geographical divisions and key policy areas

⁹² For example, the focus of M&E in request for tender documentation and submission assessments; aid procurement agreements (statement of requirements and pricing schedule).

concerning aid performance and effectiveness. The Reference Group provided guidance on the design and conduct of the evaluation to ensure the process and final products would be useful and likely to generate a positive impact.

The evaluation was conducted in these three phases:

- 1.) Inception:** This phase included understanding stakeholder needs, confirming the scope and focus of the evaluation and ensuring evaluation governance and relationships with ODE, the Evaluation Reference Group and DFAT's Independent Evaluation Committee.
- 2.) Research:** This phase included a range of tasks, including:
 - » an extensive desk review of literature on monitoring systems and the challenges associated with quality assurance in such systems generally
 - » the development of a CBAF to guide the team's review of investment monitoring systems
 - » a descriptive portfolio analysis of all investments managed by managing contractors (n=294)
 - » an extensive desk review of 172 AQCs for 78 investments managed by managing contractors, completed in 2017 or later
 - » an online survey of DFAT (97 responses) and managing contractor (34 responses) stakeholders to capture their views on investment monitoring and factors enabling and inhibiting good practice⁹³
 - » semi-structured interviews to seek the views of DFAT policy and investment managers, managing contractor project and policy staff, and M&E consultants in greater detail (40 in total)
 - » case studies of eight investments for in-depth analysis of factors associated with good practice in investment monitoring and factors explaining variation in the quality of these

⁹³ Response rates are not available because the surveys were distributed by DFAT and by managing contractors internally through an online link.

monitoring systems.

Desk-based research and consultation were undertaken throughout the research phase. A theoretical model from the literature and document review was used to develop the CBAF (that is, on the factors that influence investment monitoring systems). A hybrid approach to thematic analysis was then applied, where data collection and coding commenced with a deductive and theory-driven coding system. New codes were added iteratively through consultations. These processes informed the design of the research tools, for example by identifying the factors most likely to influence investment monitoring performance. The analysis of consultation and desk-review findings was integrated with data from the application of the research tools to develop findings and recommendations.

Desk-based research

The desk-based research enabled the evaluation team to collate and analyse a wide range of secondary data relevant to the key evaluation questions. It provided a strong foundation for qualitative and quantitative primary research that was then targeted to explore specific trends in data.

The focus of research was on:

- » global trends in monitoring systems in international development
- » policy context, data and resourcing informing the development of investment-level monitoring systems and management
- » other studies of relevance to the development of the approach, method and selection of evaluation analytical tools
- » DFAT's M&E and performance policies, including mapping DFAT's formal performance management, procurement, contracting, investment monitoring policies and objectives
- » key areas identified in the DFAT Aid Program

Health Check: improving investment designs; improving performance culture; nurturing aid management; developing skills and expertise; and sharpening focus on the implementation phase of investments.

The desk-based research led to the:

- » identifying of current practices and the key components of effective investment monitoring systems and approaches
- » identifying of guiding principles and criteria against which to assess investment-level monitoring systems and approaches
- » summarising of findings in an evidence paper covering each monitoring system component and identifying strengths, weaknesses and omissions of the literature
- » drafting of the analytical tool used by the evaluation to assess the quality of investment monitoring systems (the CBAF).

Consultation

A wide range of stakeholders from functional, corporate, policy and geographical areas within DFAT were consulted. The perspective of managing contractor partners was extremely important in triangulating evidence and assessing the extent to which the perspectives of DFAT and managing contractors align.

Consultations were conducted through semi-structured interviews with communication and engagement guided by the evaluation's stakeholder engagement and communication plan.

Consultation involved conducting:

- » preliminary targeted interviews with key internal ODE and DFAT staff to frame the evaluation and analytical framework
- » consultations with the evaluation Reference Group to seek feedback around the draft analytical framework, survey and case studies
- » an online survey of all managing contractors

identified in the initial sampling frame and across all investment managers at post

- » targeted interviews with DFAT staff and external counterparts to develop the case studies, and guide analysis against key evaluation questions
- » further broad-based consultation to collect information to answer key evaluation questions 1 and 2, which included semi-structured interviews with DFAT staff in Canberra and at post working on investment design, contracting and procurement, and performance and quality
- » targeted semi-structured interviews with a small sample of M&E consultants providing services to DFAT.

3.) Analysis and reporting involved:

- » Assessing and compiling data shortly after it was gathered. This included descriptive quantitative analysis of the investment sample as well as qualitative content analysis of relevant investment documents and narrative content from the online survey and semi-structured interviews.
- » Mapping of data to the evaluation questions, in alignment with the analytical framework.
- » Triangulating data across data sources and tools and resulting synthesis of results.
- » Facilitating three recommendations workshops with key stakeholder groups (senior DFAT management, DFAT Contracting and Aid Management Division, and managing contractors) to receive early feedback on evaluation findings and recommendations and identify further analytical tasks if needed.
- » Facilitating a workshop with a sample of investment managers to test and refine the monitoring tools for investment managers.

Methodological limitations

The major methodological issues identified during the evaluation design process and actions employed to address them are outlined here.

Sample of investments

Given the number of current or recent investments delivered by managing contractors (more than 300), a sampling approach was required to select a subset of investments to consider in more detail and provide the basis for detailed case studies. The evaluation plan proposed a sampling framework to guide selection. A key consideration of the evaluation was identifying conditions that support strong monitoring systems and how they can be more consistently replicated across the aid program. It was important to select cases that maximised learning on better-practice and less successful monitoring and, as a second priority, ensured a mix of geographic, policy and thematic areas.

Survey of DFAT and managing contractor staff

The 97 responses from DFAT staff represent about 5 per cent of all staff, and no responses were received from the Pacific Division. The 34 responses from managing contractors were from all organisations identified as major contracting partners for DFAT (based on financial data). Recognising these limitations, data from the survey were triangulated with findings from interviews and case studies.

Inclusion of facilities in the sample

Over the last 24 months, the importance of facilities as a vehicle for delivering Australian aid has increased. Several new facilities were approved and began implementation. This included the PNG Governance Facility (currently valued at \$400 million). It also included multi-sectoral investments, such as the Timor-Leste Human Development Program (covering health, water, education, nutrition, gender equality, disability and social protection) and the Fiji Program Support Facility (covering governance, education and health).

Given the nascent implementation of these facilities as a delivery mechanism, it was challenging to include such investments in the sample. The evaluation undertook some targeted desktop work and consultation to extract emerging challenges, successes and lessons where possible. KOMPAK, one of DFAT's first facility style investment, was included as a case study to help facilitate learning on monitoring for facilities.

Accounting for variation in implementation contexts

Aid investments and their monitoring systems operate within a complex environment, with investment managers and implementing partners often having limited capacity to influence factors such as partner government monitoring systems. For this reason, the evaluation took all steps possible to isolate the impacts of factors outside the direct control of investment managers and managing contractors, so the characteristics of successful and less successful investment monitoring systems could be explored.

Gauging perspectives on supply and demand factors

To build an understanding of how investment monitoring systems operate in different contexts from the perspectives of both DFAT and managing contractors, it was important to capture the views and experiences of both stakeholder groups.

Definitions of better-practice investment monitoring

The evaluation focused on collecting a realistic and efficient body of data that is relevant for decision makers in assessing the quality of monitoring systems and sufficient to provide recommendations on potential enhancements. This focus was assisted by having a very solid and shared understanding of what quality or good-practice monitoring systems are. This was defined upfront and based on credible sources from the literature, with input from ODE and DFAT. It was then reviewed, shared back and reiterated for common understanding.

ANNEX 2: CRITERION BASED ASSESSMENT FRAMEWORK

The Criterion Based Assessment Framework (CBAF, Figure A2.1) provides the evaluation team with a structure to assess the sample of DFAT

investments, based on the determinants of quality for investment monitoring systems.

Figure A2.1: Criterion Based Assessment Framework



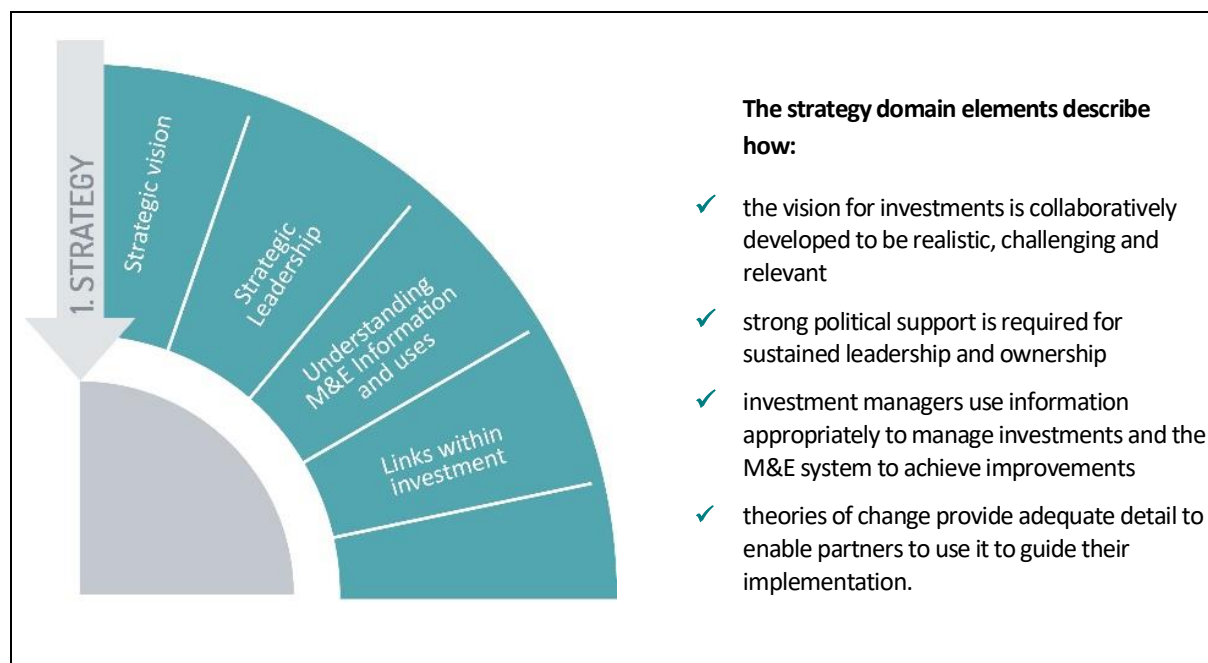
The four domains represent the key areas in which good-quality monitoring takes place—**strategy, infrastructure, capacity** and **enabling environment**. They describe the essential characteristics of good-quality monitoring systems. The CBAF forms a structure with which the evaluation team can review, question and analyse the systems and processes that form the

basis of DFAT investment-level monitoring systems. Associated with each domain is a set of four related elements that further inform the nature of the research and evaluation required. These are the core determinants of quality of each domain and are designed to provide guidance on what must be in place or addressed within investment monitoring systems to achieve sustained success within each domain.

Strategy domain

The strategy domain describes the strategic context within which the monitoring system is established and sustained. High-quality monitoring systems require an understanding

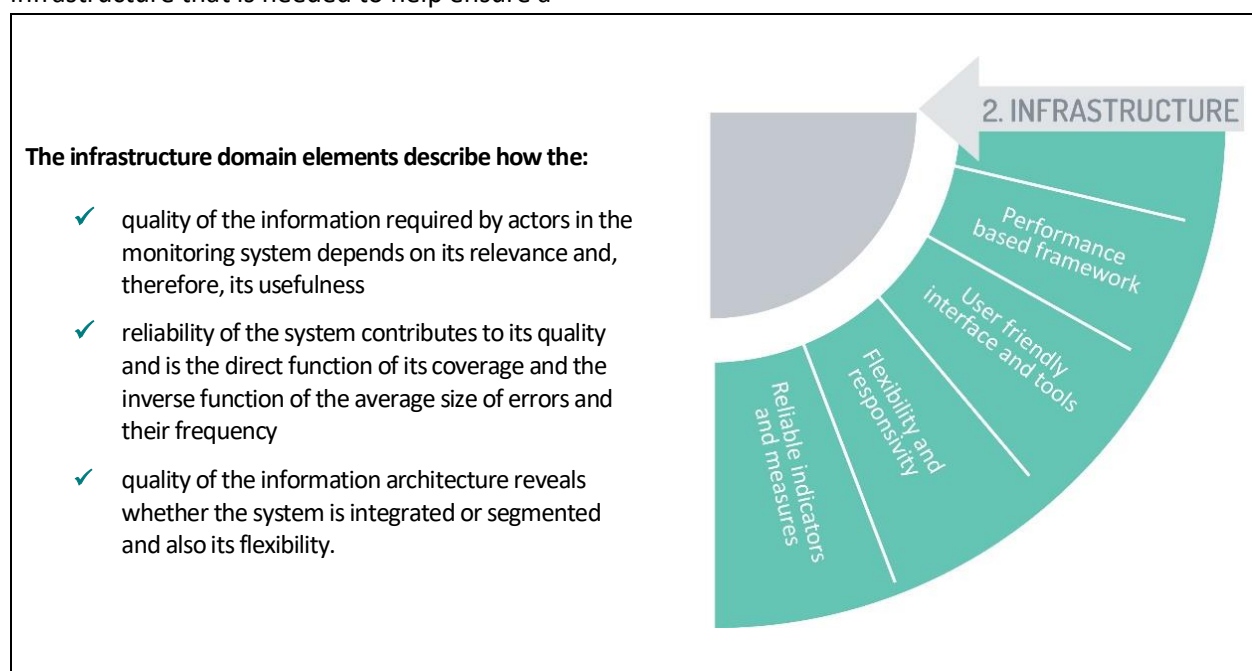
of how monitoring information can assist investment managers and decision makers to set directions and guide investments. This requires strategic leadership as well as a clear understanding of the basic concepts and potential uses of M&E.



Infrastructure domain

The infrastructure domain describes the infrastructure that is needed to help ensure a

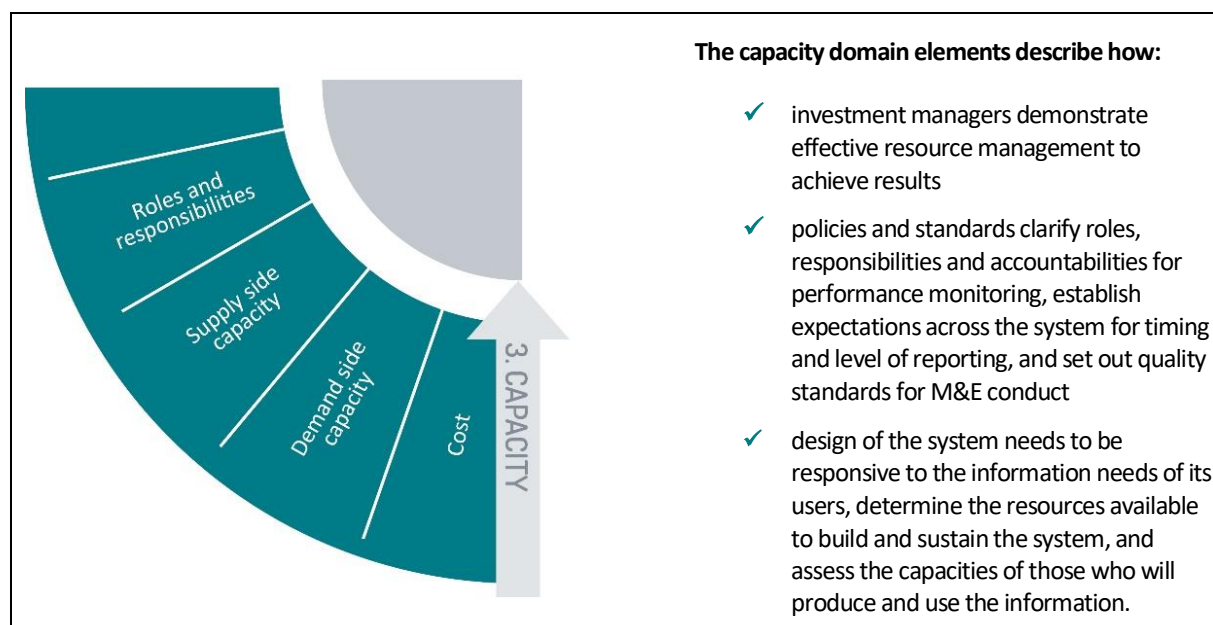
systematic, comprehensive and credible approach to M&E.



Capacity domain

The capacity domain describes the capacity to supply and use M&E information. This requires a clarity of expectations on where and how M&E information is intended to be used (for example, planning, policy or

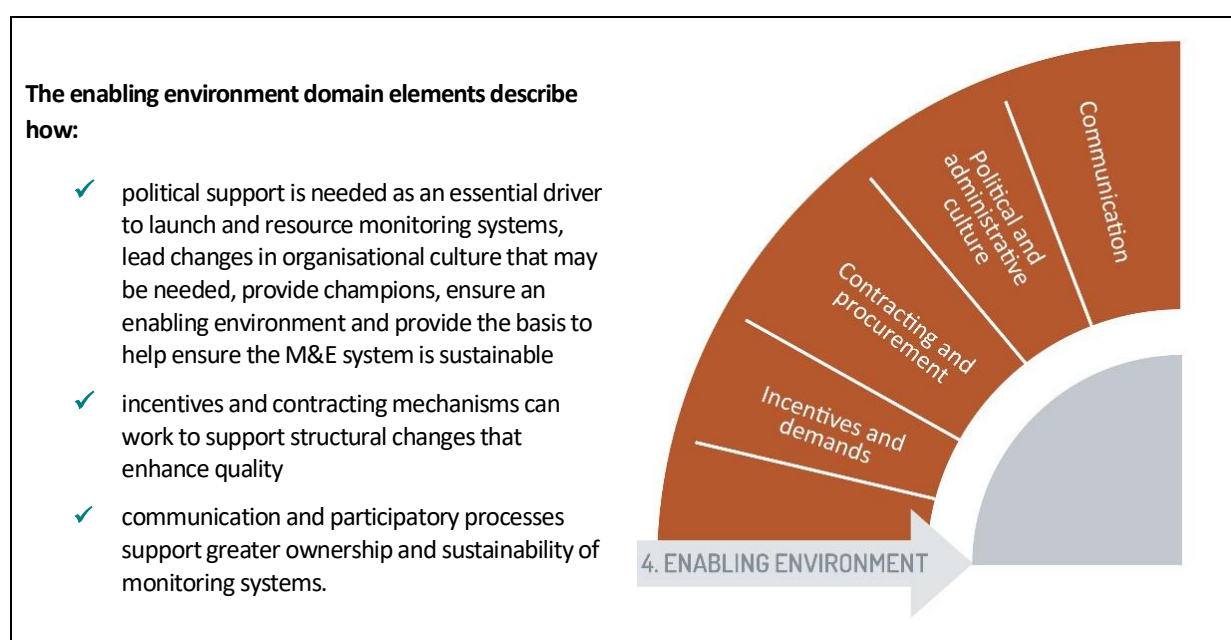
program development, decision making and budgeting). It also requires clarity of expectations on the capacity to incorporate and use the M&E information as part of the normal process of business.



Enabling environment domain

The enabling environment domain describes a culture in which investment managers have a suitable appreciation of M&E concepts, there

are adequate incentives for managers to use M&E information, and where managers report credible, unbiased and timely results.



ANNEX 3: TERMS OF REFERENCE

INVESTMENT MONITORING SYSTEMS: AN EVALUATION

SUMMARY

Investment monitoring systems are at the foundation of DFAT's aid management system and external accountability reporting. Data and evidence from investment monitoring systems are used by investment managers to complete AQC⁹⁴, Partner Performance Assessments, and report Aggregate Development Results.⁹⁵ This information informs performance assessments through Aid Program Performance Reports. It also informs decision making within country and regional programs and within the aid program as-a-whole. This is through, for example, aid governance committees and the Performance of Australian Aid report. It is therefore important that the data and evidence produced from investment monitoring systems be credible and robust.

Despite their importance, ratings by investment managers of the quality of investment monitoring systems in AQC⁹⁴ have been persistently lower than those of other criteria assessed under this system. This

also suggests that some systems may not be providing sufficiently robust evidence to underpin performance reporting on Australia's aid program or to adequately manage investment performance and risk.

The purpose of the evaluation is to help DFAT improve its investment monitoring systems and, through this, the effectiveness of the Australian aid program.

These terms of reference were drafted by ODE) in consultation with the Aid Management and Performance Branch and performance and quality staff across DFAT.

WHAT IS INVESTMENT MONITORING?

Monitoring is an intrinsic part of the framework of controls used to ensure that products or services are delivered in accordance with expectations that corrective actions required are identified and that opportunities for improvement are identified. It is axiomatic that good-quality M&E is associated with better project outcomes.

Strong monitoring arrangements are those that are planned, continuous and systematic, and documented. The ability of investments to adapt to changing context—which is crucial if investments are to be effective—is totally

⁹⁴ AQC⁹⁴ are annual self-assessments of investment quality. DFAT investment managers rate investment quality against the criteria relevance, effectiveness, efficiency, gender, sustainability, and M&E. Investment managers must give each criteria a rating between 1 and 6, with 1 to 3 representing inadequate quality and 4 to 6 representing adequate quality.

⁹⁵ Aggregate Development Results are indicators of development impact that can be aggregated across the aid program to demonstrate the contribution of Australian aid to development outcomes in partner countries.

dependent on the effectiveness of the monitoring system.

While monitoring and evaluation is often discussed together, they are distinct processes. DFAT defines evaluation as the systematic and objective assessment of an investment which takes place on a periodic basis.⁹⁶ Monitoring will often draw on evaluation as a source of evidence, and vice versa. The focus on investment monitoring systems in this evaluation complements earlier ODE reviews of the quality of aid investment evaluations.⁹⁷

WHAT IS DFAT'S APPROACH TO INVESTMENT MONITORING?

Data collected through monitoring systems is the basis for delivery partners to provide DFAT with assurance that progress is as expected, or where off track to enable corrective actions to be identified and implemented. For the majority of DFAT aid investments, investment monitoring occurs at these two levels:

- » by the delivery partner directly responsible for the implementation of the aid investment (that is, the managing contractor, multilateral organisation, NGO, or partner government)—this monitoring information is used by the delivery partner for learning, managing and reporting to DFAT
- » by DFAT which has oversight of investment implementation and quality as well as responsibility for meeting corporate reporting requirements. While DFAT principally draws on monitoring undertaken by delivery partners for information, it may supplement this with, for example, field visits to investment sites. In many cases,

monitoring is outsourced to third parties, with responsibility for fulfilling DFAT's interests in monitoring the performance of the delivery partner.

DFAT's requirements and guidance for investment monitoring have been fairly consistent over time. Since at least 2005, investments have been required to have an M&E framework as part of their design. Quality requirements for M&E at design have been specified in various DFAT quality standards.

In summary, DFAT's investment monitoring should:

- » produce credible data and evidence, including outcomes
- » generate data and evidence to inform investment decision making and accountability
- » where relevant, build the capacity of partner monitoring systems
- » enable monitoring of gender equality related aspects of projects
- » be provided with sufficient resources needed to carry out the above.

Although DFAT has requirements and quality standards in place for investment monitoring systems, there is a gap between guidance and practice. This may be because of weaknesses in the designs themselves (related to, for example, program logic). It may also be because there are no hard gates to ensure the quality of monitoring systems. Delegates are able to approve investment designs that do not have well-considered M&E frameworks.⁹⁸ There is no standard as to the level of resources required for monitoring. This depends on factors such as the level and nature of demand from decision makers and stakeholders, risk, historic

⁹⁶ In DFAT, evaluations are independent. This means they are led by a person who is not involved in investment management.

⁹⁷ ODE, Review of Operational Evaluations completed in 2014.

⁹⁸ Where DFAT contracts out the delivery of projects or programs, it will commission a design team or delivery partner to identify the broad parameters for M&E as part of the

investment designs, which is then further refined in the early implementation of an investment. More detailed M&E plans are usually separately approved by DFAT investment managers in the inception phase.

performance, complexity, size, political interest, and the form of aid being used.

DFAT investment managers use the information from monitoring systems to inform decision making and, as noted earlier, to report on performance internally, especially through AQC. These checks also provide an opportunity for investment managers to reflect on the quality of the investment's M&E system, as they are required to assess whether these systems are of adequate or inadequate quality.

RATIONALE FOR THE EVALUATION

The majority of investments are assessed as having 'adequate' or 'better quality' M&E systems. However, despite the importance of investment monitoring systems, the self-assessed quality of those systems within DFAT suggests monitoring is one of the weaker aspects of the department's management of aid performance. Furthermore, ODE AQC spot checks have consistently found M&E ratings to be one of the least robust of all quality criteria. This, in turn, affects the confidence in the robustness of assessed performance against other criteria. Moreover, the persistence of these lower ratings suggests there may be systemic constraints to improving the quality of investment monitoring. The quality of investment monitoring is also highlighted as a concern in a number of recent ODE evaluations. For example:

- » *A window of opportunity: Australian aid and child undernutrition* (April 2015) found that while many investments identified indicators in their design phase, these indicators were often inappropriate, focusing on outputs instead of outcomes or impact. Key measures such as stunting rates were regularly omitted.
- » *Banking our aid: Australia's non-core funding to the Asian Development Bank and the World Bank* (September 2015) found

that greater attention needs to be paid to the quality of monitoring arrangements for Bank-executed projects. It also found that DFAT should improve the quality of its engagement with such projects to help ensure better-quality M&E.

- » *Investing in Teachers* (December 2015) found that only one-third of investments included learning outcome-oriented indicators. Even fewer investments collected data on these.

A number of recent ODE evaluations have examined programs with strong monitoring systems—for example, evaluations of the Eastern Indonesia Roads Improvement Program (2017), the Civil Society Water, Sanitation and Hygiene Fund (2016) and the Australia–NGO Cooperation Program (2015). A key question for this evaluation is whether the conditions that supported the establishment of strong systems in these cases can be replicated more consistently across the aid program.

EVALUATION PURPOSE

The purpose of the evaluation is to identify options for DFAT to increase the quality and use of investment-level monitoring systems and promote learning.

EVALUATION SCOPE

The core focus of the evaluation is on the monitoring systems of aid investments delivered by managing contractors and associated partner government entities. There are a number of reasons for this:

- » Aid delivered through managing contractors accounted for 20 per cent of the total aid budget in 2015–16. This is the highest proportion by a single type of

delivery partner/approach.⁹⁹ This is equivalent to around one-third of country and regional program aid budgets.

- » DFAT has considered the monitoring systems of other major delivery partners, such as multilateral banks and NGOs.¹⁰⁰ Furthermore, evidence considered by the multilateral bank evaluation suggests that managing contractor M&E may be relatively weaker in quality.¹⁰¹
- » DFAT has a greater ability to directly influence the monitoring systems of managing contractor-delivered aid investments through the design, contracting and implementation processes compared to most other aid delivery arrangements. This increases the potential relevance and management utility of the evaluation.
- » Focusing on a single type of delivery partner ensures that the scope of the evaluation is manageable and can be completed to a high standard.

The evaluation will not verify whether the M&E ratings from AQC's reflect the quality of investment M&E. This is the role of the annual ODE AQC Spot Check.

Where an evaluation focuses its empirical analysis on investment monitoring systems for projects delivered by managing contractors, consideration is given in the reporting phase to evidence from previous ODE evaluations on how DFAT can most effectively monitor multilateral development banks and NGO projects that DFAT funds bilaterally.

⁹⁹ DFAT, *Performance of Australian Aid 2015–16*, p. 17. While funding through multilateral organisations makes it higher at 41 per cent of delivered administered aid in 2015–16, this represents a diverse range of delivery partners and covers both core and non-core funding approaches. The percentages of aid delivery by other partner types in 2015–16 are NGOs at 11%, Australian public sector organisations at 11%, university and academic institutions at 8%, developing country governments at 4% and other parties at 4%.

¹⁰⁰ DFAT, *Banking Our Aid: Australia's non-core funding to the Asian Development Bank and the World Bank* (2015) and DFAT,

Evaluation of the Australian NGO Cooperation Program (ANCP) (2015).

¹⁰¹ The AQC M&E ratings for projects implemented by managing contractors was found to deteriorate over the life of the investment as indicated by the proportion of projects given unsatisfactory ratings for M&E and those given highly satisfactory ratings. DFAT, *Banking Our Aid: Australia's non-core funding to the Asian Development Bank and the World Bank*, ODE, p. 71).

EVALUATION AUDIENCES

The evaluation has five main audiences:

- » **DFAT delegates** that approve designs as well as financial and procurement arrangements for aid investments. These delegates are personally responsible and accountable for their decisions and actions in ensuring that a proposed aid investment represents a proper use of Australian Government resources and meets legislative and departmental requirements. This audience can use evaluation findings to ensure that future aid investments have better-practice monitoring systems.
- » **DFAT investment managers, performance and quality staff.** Investment managers are responsible for the design, implementation and M&E of investments, with support from performance and quality staff. This audience can use evaluations findings to identify practical steps for enabling better-practice investment monitoring systems.
- » **Aid Management and Performance Branch** is responsible for DFAT policy and guidance on investment design and M&E. This audience can use the evaluation findings to improve DFAT's approach to design and investment monitoring and the quality of support.
- » **Contracting Services Branch** is responsible for procurement services, including the quality assurance and clearance of procurement and contracting approaches. This audience can use evaluation findings to consider procurement and contracting approaches¹⁰² that support and complement better-practice investment monitoring systems.
- » **Managing contractor companies and M&E consultants** involved in the design and delivery of the Australian aid program can use the evaluation findings to fine-tune and adapt the way they work to better meet the needs of investment monitoring.

¹⁰² For example, the focus of M&E in request for tender documentation and submission assessments; aid procurement

agreements (statement of requirements and/or pricing schedule).

EVALUATION QUESTIONS

The evaluation examined the questions outlined in Table A3.1

Table A3.1: Evaluation questions

| Primary question | Sub-questions (indicative) |
|--|--|
| 1) What are the characteristics of a DFAT better-practice investment monitoring system for programs delivered by managing contractors? | a) Does the investment monitoring system generate information that meets the different needs of stakeholders? |
| | b) What is the overall utility of the information produced by the monitoring system? |
| | c) What is the quality of data generated: <ul style="list-style-type: none"> » Is the monitoring system producing credible and robust data and/or evidence? Is there robust baseline data? » Can progress towards outcomes (immediate/intermediate/final) be accurately reported? Does the monitoring system reflect the investment program logic/theory of change? Does the monitoring system capture the contribution of Australian aid to outcomes? » Do monitoring systems adequately cover gender equality? » Do monitoring systems use and strengthen national systems (where relevant)? |
| | c) To what extent is the monitoring systems used for investment management, learning and accountability? <ul style="list-style-type: none"> » Do monitoring systems support adaptive and politically informed management? » Is the monitoring system flexible enough to allow program managers to recognise and adapt to changes in context or poor results? |
| | d) Is the data generated in a timely and efficient manner? |

| Primary question | Sub-questions (indicative) |
|--|---|
| 2) What factors contribute to, or inhibit, better-practice investment monitoring systems delivered by managing contractors? What is the relative importance of those factors? What are the management implications for DFAT? | <p>a) What is the relative importance of the design, procurement, inception and ongoing management period of the investment in producing a 'better-practice' monitoring system?</p> <ul style="list-style-type: none"> » Design: Rationale and approach to concept and design. Design document meets DFAT design standards related to M&E (DFAT M&E Standard 1). » Procurement: Procurement approach and outcome reflects and supports design intent with respect to monitoring; whether monitoring system (and quality assistance) is contracted to the investment delivery partner or a third party. » Inception: Inception period occurs in which finalisation of M&E system products is explicit and quality assured. » Ongoing contract management: Is the contract set up to allow a program to adjust to monitoring information? <p>b) What is the relative importance of monitoring demand factors?</p> <ul style="list-style-type: none"> » Capability: Experience and quality of support available to DFAT investment managers; quality of DFAT monitoring guidance (DFAT M&E standards 2, 3 and 7) and technical support expertise. » Resourcing: DFAT staffing levels and time dedicated to engage in monitoring. » Incentives: formal and informal influences (including personal motivation) of DFAT investment managers, managing contractor/team leader (for example, extent to which monitoring systems are integrated with investment management and decision making). » Relationship between DFAT investment managers and M&E system contractors, including understanding of respective roles and responsibilities. <p>c) What is the relative importance of monitoring supply factors?</p> <ul style="list-style-type: none"> » Capability: Experience of M&E contractor; quality of DFAT monitoring guidance (DFAT M&E standards 2 and 3). » Resourcing: Role and time dedicated to M&E activities. » Incentives: To include formal and informal influences. » Relationship between DFAT investment managers and M&E system contractors, including understanding of respective roles and responsibilities. <p>d) What is the relative importance of the broader enabling environment?</p> <ul style="list-style-type: none"> » To what extent does DFAT corporately value and encourage the use of investment monitoring system data for investment management, learning, accountability and public diplomacy? » Incentives for quality M&E, including DFAT executive commitment to and messaging on investment M&E. |

EVALUATION METHODOLOGY

The evaluation method is expected to be outlined in detail in the evaluation plan. This should include consideration of the:

- » development of an approach to sampling a small group of investments for detailed considerations
- » development of an approach for capturing stakeholder views
- » consideration of AQC data on the quality of investment-level M&E and the determinant of this
- » analysis of the consistency and completeness of contractor reporting against M&E frameworks during implementation of programs.

The sampling strategy to be developed is likely to take into account AQC data about the self-assessed strengths and weaknesses of DFAT's investment-level monitoring, stakeholder views and the quality of investment monitoring and reporting documentation.

Efforts will be made to ensure that the sample has a broad geographic mix (Pacific, South-East Asia, and South West Asia). The evaluation plan may consider the identification of areas of enquiry linked to existing challenges associated with investment monitoring related to areas such as:

- » policy influence and advocacy
- » institutional strengthening for service delivery
- » multi-country and/or regional investments
- » flexible and adaptive investments.

Sample investments should also be:

- » designed no earlier than 2012, to help ensure that sufficient information can be gathered on the design process and for the

bulk of the implementation period to be post-integration of AusAID, the Australian Agency for International Development, and DFAT in 2013

- » implemented for at least two to three years, to ensure reasonable time for an M&E system to have been developed and implemented.

EVALUATION PROCESS

Evaluation team

The team should comprise up to three team members that collectively address these characteristics:

- » internationally recognised expertise and demonstrated capacity for thought leadership over issues associated with the monitoring and management of outsourced aid investments
- » proven experience in producing high-quality evaluation reports, reviews and/or research reports for publication, preferably on topics relevant to the evaluation
- » established track record of producing succinct and engaging analytical material on topics relevant to the evaluation for publication
- » demonstrated technical expertise in conducting methodologically rigorous reviews or evaluations
- » demonstrated commitment to rigour in evaluation methods, including in the rigorous application of methods that can capture stakeholder views (surveys, interviews and focus groups)
- » demonstrated technical expertise in analysis of quantitative data and ability to present arguments visually through tables, figures and other infographics
- » demonstrated understanding and experience of the aid contracting industry, and of working with bilateral and multilateral donors in the implementation

of large investment projects.

The distribution of roles, responsibilities and resources between team members is negotiable, but should account for the nature of the proposed contributions, availability of members, the seniority of members, and their proposed contributions to the evaluation.

There is a potential conflict of interest if any team member (and/or the source consulting company) has been previously involved in the development of investment monitoring systems and the broader performance management system for DFAT. This will be principally mitigated through an appropriate contracting conflict of interest clause for team members and/or the source consulting company. It will also be mitigated through ongoing management oversight of the evaluation team by ODE.

Evaluation roles and responsibilities

Contracted evaluation team. This team will work cooperatively and closely with ODE throughout the evaluation. It will be responsible for delivering evaluation products in accordance with the contract and the agreed evaluation plan to an acceptable quality standard (as outlined in DFAT's M&E standards).

ODE. This office will be responsible for managing the evaluation from concept to publication, including through the contractual relationship with the evaluation team. ODE will work in an integrated way with the contracted team. ODE staff will contribute to the evaluation plan and participate in data collection, analysis and report writing, under the direction of the evaluation team leader. ODE has primary responsibility for managing stakeholder engagement, including the relationship with the Independent Evaluation Committee and the evaluation Reference Group. ODE is responsible for the publication

of the final report and implementation of its dissemination strategy.

Independent Evaluation Committee. This committee is responsible for ensuring that ODE evaluations and reviews are high quality. It provides expert technical assessment and advice in relation to evaluation methods and the use of evidence to support findings and recommendations. The Independent Evaluation Committee will comment on the quality of the draft evaluation plan and the draft report. It will also endorse the final report.

Evaluation Reference Group. This group comprises representatives from the Aid Management and Performance Branch and performance and quality units from DFAT's major geographic aid divisions.¹⁰³ It will provide guidance and advice to ODE and the contracted evaluation team to increase the relevance and use of the evaluation. Evaluation Reference Group members will provide comments on the terms of reference, the draft evaluation plan and the draft report. They will also participate in briefings with the evaluation team. This includes in the initial briefing at the inception phase, during the discussion on early findings following the data collection phase, and in the recommendations workshop. In these roles, evaluation Reference Group members are expected to represent the position of their division.

Evaluation deliverables

The evaluation team will provide DFAT with these reports:

- » Evaluation plan, maximum of 20 pages, that meets DFAT M&E quality standards (Standard 5: Independent Evaluation Plans).
- » Draft report, maximum of 24 pages plus

¹⁰³ South-East Asia Maritime Division, South-East Asia Mainland and Regional Division, South West Asia Division, and Pacific Division.

annexes, and including an executive summary, maximum of four pages (Standard 6: Independent Evaluation Reports). A separate guidance attachment for investment managers, maximum 10 pages, that describes DFAT better practice and a checklist of enabling factors.

- » Final report (maximum 24 pages plus annexes and guidance attachment), that incorporates any agreed changes within seven days of receipt of feedback. The report should provide a succinct and clear presentation of key findings, lessons learned and recommendations. It should meet DFAT's accessibility guidelines and otherwise be fit for publication.

Ethical issues

The evaluation will be undertaken in accordance with the Australian Evaluation Society Guidelines for Ethical Conduct of Evaluations. The evaluation team will recognise the sensitivity of the project and will maintain strict confidentiality of all data, information and documentation provided or obtained during the course of the project.

Contracted members of the evaluation team will be asked to declare any DFAT investment M&E systems they had designed, provided advice on, or implemented. These systems should not be included in the evaluation's sample to avoid any real or perceived conflicts of interest.

