Australian Government
Department of Foreign Affairs and Trade

Australian Aid

EVALUATION OF DFAT INVESTMENT LEVEL MONITORING SYSTEMS

ATTACHMENT A: SYNTHESIS OF THE LITERATURE

December 2018

OFFICE OF DEVELOPMENT EFFECTIVENESS

dfat.gov.au/ode

CONTENTS

INTRODUCTION

This paper draws on a comprehensive literature review undertaken as part of the Office of Development Effectiveness (ODE) evaluation of investment level monitoring. The literature review involved research of publicly available policy papers and statements, published and internal Department of Foreign Affairs and Trade (DFAT) evaluation reports and studies on monitoring systems, and searches of online databases for peer-reviewed papers.

The literature review has been used in two ways to inform the evaluation. First, to summarise evidence from the literature about what a quality monitoring system is and what the determinants of a quality monitoring system are. Second, to use the findings from the literature to develop a Criterion Based Assessment Framework (CBAF), to guide the assessment of DFAT's investment monitoring systems.

This paper is structured as follows.

- Section 1—Introduction.
- Section 2—Defining quality monitoring systems. This chapter sets out the basic concepts and definitions of investment monitoring.
- Section 3—Determinants of quality monitoring systems. This chapter investigates the critical elements of, and influences on, effective monitoring systems based on the lessons learned from research experiences and the literature.
- Section 4—Monitoring systems analytical models. This chapter presents a number of analytical models used to describe monitoring systems.
- Section 5—Criterion Based Assessment Framework. This chapter presents the CBAF developed for this evaluation.

DEFINING QUALITY MONITORING SYSTEMS

This chapter sets out the basic concepts and definitions of investment monitoring. It defines what monitoring systems are and what quality monitoring is according to the literature.

2.1 MONITORING WITHIN DFAT

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. Undertaking monitoring is an assumed responsibility for DFAT staff within much of the organisation's internal guidance, including the *Aid Programming Guide*, the *Aid Quality Check*[1] guidance and template, and the *Monitoring and Evaluation Standards.*[2]

Monitoring itself is not explicitly defined by DFAT. The Organisation for Economic Co-operation and Development (OECD)[3] defines monitoring this way:

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

2.2 WHAT IS A MONITORING SYSTEM?

Monitoring systems endeavour to effectively measure progress towards development objectives. Fundamentally, the investment monitoring system supports the clarification of goals and objectives. DFAT, partner governments, managing contractors and other stakeholders can also use monitoring systems to formulate and justify budgetary requests. Information on progress, problems and performance are all key to an investment manager striving to achieve results.

Systematic monitoring and reporting on the *quality* of aid activities directly supports program management, lesson learning and improvement. It also adds to the accountability of funds committed for specific aid objectives.

A functioning investment monitoring system provides a continuous flow of information that is useful both internally and externally. The monitoring system gives ongoing information (by way of select indicators) on the direction of change, pace of change and magnitude of change. It can also identify unanticipated changes. All are critical to knowing whether policies, programs and projects are moving in the intended direction.

It is useful to highlight that a focus on monitoring systems requires an understanding of the interdependencies within such systems between key stakeholders. A singular focus on DFAT culture, systems and processes is insufficient to realise change and embed good investment

[1] Aid Quality Checks are annual self-assessments of investment quality. DFAT investment managers rate investment quality against these criteria: relevance; effectiveness; efficiency; gender; sustainability; and monitoring and evaluation (M&E). Investment managers must give each criteria a rating of between 1 and 6, with 1 to 3 representing inadequate quality and 4 to 6 representing adequate quality.

[2] DFAT, 'DFAT Monitoring and Evaluation Standards.' (2014), https://dfat.gov.au/about-us/publications/Documents/monitoring-evaluation-standards.pdf

[3] World Bank, 'Ten Steps to a Results Based Monitoring and Evaluation System.' https://www.oecd.org/dac/peer-reviews/World%20bank%202004%2010_Steps_to_a_Results_Based_ME_System.pdf

monitoring systems. To achieve meaningful change there needs to be a very strong focus on how the department works in *partnership* and *interacts* with its supply-side partners through a wide range of mechanisms.

2.3 WHAT IS A QUALITY MONITORING SYSTEM?

There is a substantial body of literature on what constitutes better-practice monitoring in the development sector and more broadly. The literature outlines ample good-practice guidelines for the technical components of monitoring systems (for example, the '10 Steps to a Results Based Monitoring System'[4], by the Word Bank, and '12 Components Monitoring and Evaluation System'[5], by the Joint United Nations Programme on HIV/AIDS (UNAIDS). Many of these technical, output-driven components are captured in DFAT's Monitoring and Evaluation (M&E) Standards.

The literature argues that strong monitoring arrangements, or quality monitoring arrangements, are those that are 'planned, continuous and systematic, and documented'.[6] For the purposes of this evaluation, quality monitoring has been defined as 'the extent to which investment monitoring systems enable the generation, and collection and analyses of credible information on aid activities and that apply internationally recognised characteristics of good aid practice.'

[4] ibid.

[5] UNAIDS, '12 Components Monitoring and Evaluation System Assessment'. http://www.unaids.org/sites/default/files/sub_landing/files/1_MERG_Assessment_12_Components_ME_System.pdf

[6] World Bank, 'Ten Steps to a Results Based Monitoring and Evaluation System.' https://www.oecd.org/dac/peer-reviews/World%20bank%202004%2010_Steps_to_a_Results_Based_ME_System.pdf

DETERMINANTS OF QUALITY MONITORING SYSTEMS

The literature recognises a set of consistent characteristics or system requirements to building and maintaining investment monitoring systems. These characteristics are subject to the influence of broader determinants in the political and organisational environment that affect the quality of the system. These determinants involve activities such as creating the right incentives and providing sufficient financial, human, and technical resources for organisations, managers and staff to carry out monitoring tasks. The determinants most relevant to this evaluation are summarised next.

3.1 STRATEGY

Alignment of investments with departmental strategies and policies is pivotal to quality monitoring. An understanding of how monitoring information can assist investment managers and decision makers is critical. This requires strategic leadership. It also requires an understanding of what investments set out to achieve as well as the basic concepts and potential uses of monitoring information.[7]

If staff are unclear about what their objectives are, or if departmental priorities frequently shift, this can have an adverse effect on learning and the sustainability of the quality of the investment monitoring system.

Strategic vision

Alignment of investments with the strategic aims of the Australian aid program and specific country programs involves ensuring:

- clarity of goals and rationale for Australian engagement
- clarity of strategic outcomes
- extent to which the articulation of strategic outcomes lends itself to performance M&E
- robustness of the investment and strategic program theory.

Any investment and program of work requires a clear rationale, well-articulated outcomes which need to be achieved within the life of the investment or parts thereof, and approaches or interventions to achieving these outcomes. There needs to be a robust theory of change so adequate resources are available and proposed interventions are likely to achieve desired outcomes.

Different interpretations of what investments are expected to achieve across stakeholders can result in over-emphasising some outcomes, while not giving enough attention to others. This, in turn, can result in insufficient achievements across the scope of the investment, and a compromise of value-for-money.

The development of DFAT Aid Investment Plans and investment designs is an important starting point. Work is required to design the Performance Assessment Framework and exploratory evaluations that underpin Aid Investment Plans, so they generate credible information as expected. This involves the extent to which indicators and evaluations have been fully operationalised within investment M&E frameworks, which define all methods for data collection and analysis, and the quality of data expected from other sources such as from partner government systems. This requires a high level of expertise and is the foundation of DFAT's ability to report confidently.

[7] ibid.

Strategic leadership

A successful investment monitoring system must have sustained leadership. While it is important to have competent managers overseeing the implementation of investments, programs and projects, there must also be strong political support at the very highest levels of DFAT (particularly at post) and within managing contractor organisations.

It takes strong and consistent leadership to institute a quality monitoring system. Bringing credible or results-based information into the public arena can change the dynamics of diplomatic and institutional relations, budgeting and resource allocations, personal political agendas, and public perceptions of governmental or aid program effectiveness.

DFAT must be in the driver's seat in leading demand for monitoring information to ensure quality control and ownership of results.

Understanding of M&E information and use in decision making

Understanding what information is required by decision makers, what information is available and whether there are formal policies or requirements on how monitoring information gets used is critical to good-quality monitoring.

Monitoring systems should be built in such a way that there is a demand for results information at every level where data is collected and analysed. If people are not involved or if there is no ownership and people begin to lose interest, the result will be poor data collection and reporting.

The literature argues that often decision makers seldom have all the information they need when they need it. Even without perfect data, however, a monitoring system that can provide some analytic feedback will help policymakers make more well-informed decisions.[8]

If demand for information is episodic or haphazard, investment monitoring systems are not going to be used and sustained. Structured requirements for reporting results, including legislation, regulations and international development requirements, can help lead to sustained, consistent demand for such systems. Demand can be stimulated when the strategic goals of the Aid Investment Plan are translated into monitoring systems.[9]

Thus, there is an absolute necessity to collect no more information than is required. Time and again, monitoring systems are designed and are immediately overtaxed by too much data collected too often—without sufficient thought and foresight into how and whether such data will actually be used. Complexity and overdesign are constant concerns.

Links within investments

In an ideal situation, project-level performance data would be fed into and linked to program assessments that, in turn, would be linked to investment, sectoral, regional, and country goals and targets. In other words, staff (both DFAT and managing contractors) at each level would have a clear line-of-sight into, or understanding about, the other levels and how they relate. Monitoring at the project level that is not clearly aligned with higher-level goals is not useful beyond the given project.[10]

Information must flow freely between levels to be truly useful. Each level must help inform the next level to achieve the desired results. It is also important to ensure that within a level there is a commitment to horizontally use and share information from the collection and analysis of data. The goal is to create a monitoring system that is transparent and aligned from one level to the next.[11]

[8] ibid.

[9] DFAT, 'DFAT Investment Level Monitoring and Evaluation Systems Report' (final version completed), unpublished internal report.

[10] ODE, 'Research for better aid: an evaluation of DFAT's investments.' https://minerva-access.unimelb.edu.au/bitstream/handle/11343/192345/research-for-better-aid-an-evaluation-of-dfats-investments.pdf?sequence=1&isAllowed=y

[11] The Information Systems Group, *Vision and Value,* 'Evaluation of information systems and monitoring arrangements for the programmes supported by the European Social Fund,' (2017), http://www.ec.europa.eu/social/BlobServlet?docId=2202&langId=en

3.2 INFRASTRUCTURE

The quality of the information required by stakeholders depends on its relevance and, therefore, usefulness. This is often determined by the technical or infrastructure components of how the system is built. The higher the relevance, the higher the quality of the information.

The reliability of the system also contributes to its quality and is the direct function of its coverage and the size and frequency of any errors in the data.

Basic foundations—performance-based framework

Investments must first have, or establish, a basic foundation—a traditional implementation-focused M&E system. Establishing a foundation requires basic statistical systems and data, as well as key budgetary systems. Data and information must be of appropriate quality and quantity. Investment monitoring systems need to provide information on their baseline conditions, that is, where they currently stand in relation to a given program or policy.[12] The completed matrix of outcomes, indicators, baselines and targets becomes the performance framework. It defines outcomes and plans for the design of a monitoring system that will, in turn, begin to provide information on whether interim targets are being achieved on the way to longer-term outcomes.

The performance framework becomes the basis for planning, with attendant implications for budgeting, resource allocation, staffing and so forth. The framework can and should be a relevant guide to managers. It should be frequently consulted and considered during the process of managing toward desired outcomes.

User friendly interface and tools

The quality of information produced by investment monitoring systems is a function of factors such as relevance of the indicators system, possibility to perform comparisons with other data and between investments, projects or regions, and frequency of updates and release of information. Finally, the ease with which useful data can be accessed and extracted from any system is telling about the overall usefulness of the system itself.[13]

Access

The literature stresses that access to data is rooted more in behaviours and demand for information than in enabling or limiting technical or information technology-driven conditions.

Comparisons

The usefulness of information is often reduced by the difficulty to perform comparisons between regions, investments or projects. Good examples of monitoring systems in the literature highlight cases where data facilitates and identifies best practice examples, as the system can communicate related figures and data.[14]

Frequency of data collection, analysis and use

It is important to build a continuous system of data collection and analysis. Data collection, analysis and reporting should be aligned throughout the various levels of investments. Data analysis and reporting at program levels then feed into the larger database in determining progress toward the desired outcomes.

The timing of data collection may be problematic and could mean that at times the information available is only partially pertinent. This has implications for the way strategic and operational decisions can be made on investment direction.

[12] UNAIDS,. '12 Components Monitoring and Evaluation System Assessment.' http://www.unaids.org/sites/default/files/sub_landing/files/1_MERG_Assessment_12_Components_ME_System.pdf

[13] World Bank, 'Ten Steps to a Results Based Monitoring and Evaluation System,' https://www.oecd.org/dac/peer-reviews/World%20bank%202004%2010_Steps_to_a_Results_Based_ME_System.pdf

[14] The Information Systems Group, *Vision and Value*, 'Evaluation of information systems and monitoring arrangements for the programmes supported by the European Social Fund,' (2017), http://www.ec.europa.eu/social/BlobServlet?docId=2202&langId=en

A continuous stream of data can provide significant information on trends and directions over time. The more often measurements are taken, the less guesswork there will be on what has happened. More data points enable managers to track trends and understand project, program and policy dynamics. The more time that passes between measurements, the greater the chance that events and changes in the system might happen that may be missed.

Performance findings should be used to help improve investments, projects, programs and policies. Analysing and reporting data yields important, continuous information about the status of projects, programs and policies. It can also provide clues to problems that arise during implementation and create opportunities to consider improvements in implementation strategies.[15]

Flexibility and responsivity

Interoperability between systems may be a cause of inflexibility. Many such discontinuities have been identified in different cases in the literature. An example is a lack of compatible interfaces that makes it difficult to centralise data coming from different sources. This can be particularly challenging for investments where multiple grantees use different systems for data collection and reporting which managing contractors and/or DFAT must aggregate and make sense of.

A case in point is whether investment monitoring systems are developed on the basis of a pre-existing monitoring system and can draw on a larger pool of credible information made available and/or can integrate with a new system or as part of a wider monitoring platform.

In some cases, data manipulation—inputting or extracting data—is difficult and time consuming. Also, cross tabulations can be difficult or not possible. It is important to ascertain if the overall architecture of the system makes the manipulation of data cumbersome and time consuming.

Reliable indicators and measures

Performance indicators can and should be used to monitor outcomes and provide continuous feedback and streams of data throughout the investment, project, program or policy cycle. In addition to using indicators to monitor inputs, activities, outputs and outcomes, indicators can yield a wealth of performance information about the process of and progress towards achieving these outcomes. Information from indicators can help to alert managers to performance discrepancies, shortfalls in reaching targets, and other variabilities or deviations from the desired outcome. Thus, indicators provide investments with the opportunity to make mid-course corrections, as appropriate, to manage toward the desired outcomes.[16]

Constructing indicators takes work. They should be constructed to meet specific needs. They also need to be a direct reflection of the outcome itself. Over time, new indicators may be adopted, and others dropped. This is to be expected. Every indicator has cost and work implications. Therefore, they should be chosen carefully and judiciously. There should be clarity and agreement in the monitoring system on the logic and rationale for each indicator from top-level decision makers to those responsible for collecting data in the field.

Indicators ought to be adequate. They should not be too indirect, too much of a proxy, or so abstract that assessing performance becomes complicated and problematic. Indicators should be monitorable, meaning they can be independently validated or verified. This is another argument in favour of starting with quantitative indicators as opposed to qualitative ones.

Indicators should be reliable and valid to ensure that what is being measured at one time is what is also measured at a later time and that what is measured is actually what is intended. Caution should also be exercised in setting indicators according to the ease with which data can be collected. Too often, the selection of indicators is based on how readily available the data is, not on

[15] ODE, 'Research for better aid: an evaluation of DFAT's investments,' https://minerva-access.unimelb.edu.au/bitstream/handle/11343/192345/research-for-better-aid-an-evaluation-of-dfats-investments.pdf?sequence=1&isAllowed=y

[16] United Nations. 'United Nations Development Assistance Framework guidance,' https://undg.org/document/2017-undaf-guidance/

how important the outcome indicator is in measuring the extent to which the outcomes sought are being achieved.[17]

Factors that jeopardise data accuracy are more frequently related to ill-definition of indicators. If the monitoring system is to be a useful management tool, it needs to be manageable, and not overloaded with too many indicators. Otherwise, too much time will be spent managing the system that produces the data, and not enough time will be spent using the data to manage.

As with agreeing on outcomes, the interests of multiple stakeholders should be taken into account when selecting indicators. Outcomes need to be translated into a set of measurable performance indicators. The selection process should be guided by the knowledge that the concerns of interested stakeholders must be considered and included. It is up to M&E advisors, together with relevant team leaders or facility directors, contractor representatives and DFAT investment and program managers, to distil stakeholder interests into good, usable performance indicators.

Thus, outcomes should be disaggregated to make sure indicators are relevant across the concerns of multiple stakeholder groups, not just a single stakeholder group. Just as important, the indicators have to be relevant to investment managers, because the focus of such a system is on performance and its improvement.

Baselines

Quality monitoring systems will enable reporting against baseline and intermediate measurements to determine whether progress has been sustained, whether there was only a short spurt of improvement, or whether early improvements have all disappeared. Comparing actual outcomes to targets is central to reporting results.

Performance into the future (setting of targets) cannot be projected without first establishing a baseline. The baseline is the first measurement of an indicator. It sets the current condition against which future change can be tracked. For instance, it helps to inform decision makers about current circumstances before embarking on projecting targets for a given investment, program, policy or project. In this way, the baseline is used to learn about current or recent levels and patterns of performance.

Importantly, baselines provide the evidence by which decision makers are able to measure subsequent policy, investment, program or project performance and/or engage in policy dialogue with partner governments.[18]

The challenge is to obtain adequate baseline information on each performance indicator for each outcome. This can quickly become a complex process. It is important to be judicious in the number of indicators chosen, because—as discussed earlier—each indicator will need data collection, analysis and reporting systems behind it.

Coverage of data

The coverage of data permitted by the monitoring arrangements is one factor contributing to the overall reliability of data produced. The selected performance indicators, and the data collection strategies used to track those indicators, need to be grounded in the realities of what data systems are in place, what data can presently be produced, and what capacity exists to expand the breadth and depth of data collection and analysis.

Every indicator constitutes its own miniature monitoring system, so the first consideration in starting to build the information system for that indicator is what sources of information potentially can supply the relevant data. It is important to collect only the data that is intended to be used.[19]

[17] Sustainable Measures. 'Characteristics of effective indicators.' http://www.sustainablemeasures.com/node/92

[18] Sevone, '6 steps to an effective performance monitoring strategy.' https://www.sevone.com/white-paper/6-steps-effective-performance-monitoring-strategy

[19] United Nations, 'United Nations Development Assistance Framework guidance,' https://undg.org/document/2017-undaf-guidance/

Verification and validation

Verification and validation of data take place at different levels, and according to different modalities.

The United States Agency for International Development (USAID), for example, values the quality of data provided by partners, and as such conducts Data Quality Assessments (DQAs) in accordance with Automated Directives System (ADS) in an effort to understand and increase the quality of the data it reports on regularly.[20] According to the ADS, the purpose of a DQA is to ensure that the USAID Mission and technical offices overseeing an activity are aware of the strengths, weaknesses and limitations of their performance data as well as the extent to which the data can be trusted to influence management decisions. A DQA of each selected performance indicator helps validate the usefulness and integrity of the data.

The ADS mandates that 'data reported to USAID/Washington for Government Performance and Results Act reporting purposes or for reporting externally on Agency performance must have a data quality assessment within the three years before submission.' Through a DQA, Missions should ensure that the data being reported are measured against five data quality standards—validity, integrity, precision, reliability and timeliness (abbreviated V-I-P-R-T).

The ADS requires Missions to:

» review data collection, maintenance and processing procedures so procedures are consistently applied and continue to be adequate
» identify areas for improvement, if possible
» retain DQA documentation in performance management files and update the information within three years.

The DQA is also an opportunity for building capacities and improving reporting quality. Thus, the DQA helps end-users of USAID data to know the strengths and limitations of the data on which their programs report.

3.3 CAPACITY

To be effective, monitoring needs to be positioned as far more than a technical instrument for change. It is not enough to simply create highly trained monitoring capacity and expect that organisations and systems will eventually become more effective. There is a need to also address the institutional capacity that is generating the demand for and supply of performance information.

Designing and building a monitoring system that can produce trustworthy, timely and relevant information on the performance of investments, projects, programs and policies requires experience, skill and real institutional capacity both within DFAT and managing contractors.

This capacity for a quality monitoring and reporting system has to include, at a minimum, the:

» ability to successfully construct indicators
» means to collect, aggregate, analyse, and report on the performance data in relation to the indicators and their baselines
» need to have managers in place with the skill and understanding to know what to do with the information once it arrives.

Building such capacity within DFAT and managing contractors for these systems requires long-term effort.

Capacity within DFAT

Often large assumptions are made about what program managers or generalist public servants can reasonably do with performance management. Critical areas of performance management are commonly delegated to staff that do not have the skills and/or experience to perform them well. Two common but high-risk areas are the development of:

[20] USAID, 'How to conduct a data quality assessment (DQA): an aid memoir for a COR/AOR,' http://pdf.usaid.gov/pdf_docs/pnaec151.pdf

» program theory in investment designs or country program strategies (Aid Investment Plans)
» performance assessment frameworks for investments or, more commonly, country program strategies.

The Evaluation Capacity Building Program[21] found that being able to assess how robust an investment's program theory may be is particularly demanding.

Although it is highly desirable for DFAT teams to be very closely involved in strategic or investment-level designs, the generation of program theory requires specialist skills which not all staff may have. Some important required expertise includes:

» knowledge of how to apply the principles of program theory to design
» familiarity with what works and does not work in development programming in complex contexts
» knowledge of important theories or of the general literature on topics such as individual behaviour change, organisational change, effective partnerships, getting knowledge into policy and diffusion of innovation
» knowledge of how to balance political and technical imperatives
» understanding of common issues in the political economy of development
» ability to deal with the preferences of powerful voices
» practical experience in program delivery, especially at the interface between programs and target beneficiaries.

Anyone who has worked in aid for a long time appreciates the inefficiencies and disappointing results associated with weak designs. It is cost effective to identify and access the highest quality technical assistance available.

This skill is also important well beyond the design stage. During annual planning, for example, seemingly small decisions can quickly add up to strategic drift or the selection of ineffective interventions. If a program manager responsible for the oversight of an investment is not able to connect work plans and outcomes, then important risks quickly present themselves.

Developing a Performance Assessment Framework also requires particular expertise. Skills include:

» appropriately applying basic measurement theory
» operationalising indicators (identifying suitable methods of data collection, analysis and tool development)
» assessing the reliability and validity of data sources
» making sensible compromises in the challenging and resource-constrained environments in which the aid program works.

Program staff are capable of taking on quite challenging technical roles if well supported. However, some aspects of these tasks will always require technical expertise. For example, a program manager can make an overall assessment of the quality of an M&E plan, but if there are questions about non-payment for a product, it is useful to have technical assistance to provide a high-level assessment to support this decision. Small investments in competent technical expertise at the right time can build strong foundations and limit the burden from problems encountered when peer reviewing designs, implementing programs and preparing Aid Program Performance Reports.

Management of quality investment monitoring systems requires the right behaviours and cooperation from the executive, senior management, program staff, procurement officers, implementation partners, investment M&E practitioners, national partners, designers, and independent reviewers. Focusing on the skills and incentives of any group in isolation of the others cannot result in the desired organisational change and likely will not represent value-for-money.

[21] DFAT, 'Evaluation Capacity Building Program: Submission to ODE Review on M&E Systems,' unpublished internal report.

Capacity within managing contractors

Some of the most critical issues in implementing and sustaining quality monitoring systems for managing contractors are the challenges in recruiting and holding talented staff who can design monitoring systems and build and manage them.

Findings from previous DFAT-commissioned studies[22] into investment monitoring systems suggest that the qualifications and base competency of M&E practitioners is a partial constraint on effective M&E. Anecdotal evidence affirms the difficulty that DFAT managers and managing contractors face in recruiting suitably qualified and experienced M&E practitioners even though, globally, there is no shortage of qualified and experienced people.

Providing oversight and direction on strategic performance information during the design phase of investments requires a particularly high level of expertise. This type of advice is not always found in program-level M&E advisors. It requires:

» a high level of expertise in program theory and design (investment, portfolio, program, and activity levels) in dynamic and complex settings

» a particularly high level of credibility to allow advisors to act as effective change agents

» the ability to provide constructive and collegiate advice to program-level M&E advisors working across the investment itself and the investment portfolio.

These types of advisors are hard to find, and at times are often difficult to get access to as required by services contracts.[23]

Furthermore, frequently, a careful assessment of the competency of M&E advisors is not made—in particular, one that would reflect the real capacity to create, use and sustain the system. A carefully completed assessment helps build a solid understanding of how to design the system in order to:

» be responsive to the information needs of its users

» determine the resources available to build and sustain it

» assess the capacities of those who will produce and use the information.

Roles and responsibilities

Roles and responsibilities and existing structures available to monitor development goals are critical to good investment monitoring systems. Clear roles and responsibilities and formal organisational and political lines of authority must be established. The organisation and people in charge of collecting, analysing and reporting performance information must be clearly defined.

Cost of monitoring systems

The cost profile of monitoring systems is a key influencing factor in facilitating the quality of information available and its accessibility to stakeholders.

Important here is to consider:

» the opportunity cost identified with the use of any part of the technology platform used for monitoring investments—both human resources and time necessary—to feed in and extract data

» the wealth of information provided and whether it is accompanied by a highly flexible and uncomplicated use of the database (if one is in use), both to input and extract data

» the overall benefit derived from its use, that is, the usefulness of the information provided and if the information is sent upwards and used to inform decision making

» the time necessary to input information and for data validation and verification

» the benefits that accrue to other stakeholders who may contribute or use the information

[22] 'DFAT Investment Level Monitoring and Evaluation Systems Report (final version completed),' unpublished internal report.

[23] KPMG, 'Report for DFAT PNG Health and HIV M&E Service Provider,' internal document.

- how user-friendly any information technology program is and how complex the architecture of any modules or user interface may be
- how easy the monitoring system is to learn.[24,25]

3.4 ENABLING ENVIRONMENT

A focus on monitoring systems requires an understanding of the interdependencies within such systems. This includes creating the right incentives, contracting arrangements and communication between different departments and external partners, and the broader political and administrative culture.

Incentives and demands

It is important to determine whether incentives exist—political, institutional or personal—and how they may influence an investment monitoring system. A multitude of pressures may need to be responded to. These will drive incentives for building, managing and using an investment monitoring system.

Investment managers have the option of offering monetary incentives to managing contractors for good performance and sanctions or non-monetary implications for performance that fails to meet expectations or falls short of intended outcomes. This includes poor performance reviews.

Incentives are also needed to encourage the use of performance information. Success needs to be acknowledged and rewarded and problems addressed. Messengers must not be punished, organisational learning must be valued, and budget savings must be shared. It is imperative that results information be used. Simply providing information to investment managers is not enough. Investments by a range of donors have used different approaches to providing such incentives. These generally fall into active or passive measures.

Active measures include:

- formal reviews through regularly scheduled meetings at which performance is assessed
- senior management attention, either as the chair of the formal review or direct engagement in monitoring and following up on performance expectations
- non-monetary rewards, such as public recognition with an award or honour.

Many of these active measures are blended for greater impact. Former United States Vice President Al Gore's High-Impact Agency Initiative, and the United Kingdom's Prime Minister's Office's six-month performance reviews, are examples of this.

Passive measures include:

- performance contracts, such as formal agreements between managers and staff on targets which imply a formal review at the end of the contract period
- peer pressure, such as a scorecard of performance for each investment that is made widely available, so investments can be easily compared
- public embarrassment
- approval or monetary incentives if performance improves or targets are achieved, either on an individual or overall basis.[26]

Contracting and procurement

Contracting and procurement plays a number of roles in ensuring quality monitoring systems.

[24] World Bank, 'Ten Steps to a Results Based Monitoring and Evaluation System,' https://www.oecd.org/dac/peer-reviews/World%20bank%202004%2010_Steps_to_a_Results_Based_ME_System.pdf

[25] The Information Systems Group, *Vision and Value*, 'Evaluation of information systems and monitoring arrangements for the programmes supported by the European Social Fund.' (2007), http://www.ec.europa.eu/social/BlobServlet?docId=2202&langId=en

[26] An example of individual basis incentive includes by tying senior management pay or bonuses to a Partner Performance Assessment. An example of overall basis incentive includes by tying overall managing contractor pay or bonuses to performance, or by trying to link the managing contractor's budget to its performance, that is payment by results.

First, it is a mechanism for engaging the right expertise by way of individual M&E advisors. Second, it is the mechanism through which to engage managing contractors and articulate performance expectations and incentives.

The fastest way to ensure that contractors identify competent M&E advisors is to build into contracts the requirements to meet clearly defined standards and ways by which contractors will be held to account for those standards.

Outcome-based performance targets can be included in grant or funding agreements with outcomes compared against targets. Rewards and penalties based on performance can be delineated in such contracts. If there are no data on which to base decisions, those decisions can be arbitrary.

DFAT's basis of payment for performance falls into these four categories:

1. management fee tied to milestones
2. management fee tied to Partner Performance Assessment scores
3. management fee tied to results
4. incentive payments (that is, bonus payments).[27]

The United Kingdom's Department for International Development (DfID) called its 2014 payment by results strategy *Sharpening Incentives to Perform.*[28] It was seen to promise an approach to financing that makes payment contingent on the independent verification of results, aligns incentives with partner governments for development outcomes, and encourages innovation. At the strategic level, pure payment-by-results is not suited to DFAT's business model and is, indeed, largely unproven in aid as being an effective mechanism.

Political and administrative culture

Monitoring systems will inevitably, even if infrequently, produce data that may be sensitive. If it is clear that only politically popular or politically correct information will be allowed to emanate from the monitoring system, the system will be vulnerable and compromised from the beginning. It will not be seen as credible. Some organisations have a culture where accountability tends to be associated with blame. This has the effect of discouraging openness and learning. In other organisations, it is more acceptable to own mistakes and see them as opportunities for learning, recognising there is often as much to learn from poorly performing projects as there is from success stories.

Champions in managing contractors and DFAT are critical to the sustainability and success of a monitoring system. A highly placed M&E champion can be a strong advocate for more well-informed decision making. Viability is dependent upon the information being viewed as relevant, trustworthy, useable and timely.

Monitoring systems with marginally placed champions who are peripheral to the decision-making process will have a more difficult time meeting these viability requirements, hence the criticality of roles and responsibilities and where the monitoring champion is placed in the investment organisational structure.

Communication

The new realities of governance and donor-performance expectations require an approach that is consultative, cooperative and committed to consensus building. The voices and views of stakeholders should be actively solicited. Engaging key stakeholders in a participatory manner helps to build consensus and gain commitment to reaching the desired outcomes.

If the investment monitoring system is to provide continuous performance feedback as a management tool, continuous communication is important. Monitoring results should be continuously disseminated to provide feedback to decision makers. Typically, the more senior the recipient, the less need there is for extensive detail and explanation. Aggregated, succinct data

[27] Advice received from DFAT Contracting and Aid Management Division, 2017 (internal email).

[28] DfID, 'Sharpening Incentives to Perform: DFID's Strategy for Payment by Results,' https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/323868/Sharpening_incentives_to_perform_DFIDs_Strategy_on_Payment_by_Results.pdf

relevant to the specific issue will be more appropriate. For this reason, personal briefings—especially to high-level officials—can be another effective means of communicating performance findings. Further down the managerial chain, it is more likely that more operational data will be required. This may require tailoring information into the preferred format for each decision maker and end user.

MONITORING SYSTEM ANALYTICAL MODELS

Analytical models for assessing monitoring systems generally highlight that monitoring arrangements cannot be viewed in isolation from the institutional, social and technical contexts in which they develop. Management and maintenance of quality monitoring systems requires creating the right incentives and providing sufficient financial, human and technical resources for organisations, investment managers, and staff to carry out monitoring tasks.

To account for these different dimensions, a framework for analysing information systems and monitoring arrangements in the form of a matrix is required. One the one side, such a matrix comprises structure (extant organisational arrangements, resources and capacity), process (what is done, by whom and how) and outcomes. On the other side, the matrix comprises organisational and/or institutional factors, social and/or stakeholder perspectives (that is, monitoring arrangement participant perspectives), and technical systems functions.

It is argued that this approach enables evaluators to account for the fact that the success of monitoring arrangements does not depend only on technologies employed or the design of data stores and information flows. It also, crucially, depends on social and institutional factors.[29] Social factors means the willingness of different participants in the overall program to generate data and exchange information with one another. Institutional factors mean the formal layouts that allocate decision making power to different actors and design a certain implementation process. Key here for managing contractors is the investment design and human resourcing and subsequent roles and responsibilities dedicated to the investment.

Some studies in the literature found it useful to go down to a fairly disaggregated-level of criteria to account for the different environment of investment monitoring cases. A second way to obtain performance criteria that can be measured in a more objective and comparable manner is to identify certain specific tests, like feedback and horizontal flows, that are normally associated with higher or lower performance.

Precedent analytical models used within ODE evaluations

It was useful to consider precedent analytical models in determining useful frameworks for assessing investment monitoring systems within the DFAT context for this evaluation.

A previous ODE-commissioned analysis (internal) on investment monitoring systems[30] developed an analytical framework that enabled the evaluation team to systematically disassemble M&E systems from the standpoint of three main classes of stakeholder. The systems were analysed from these perspectives:

- **Demand-side.** DFAT's investment managers who have oversight of implementation and quality and, as such, define the scope and standard expected of M&E products.
- **Supply-side.** M&E practitioners responsible for designing M&E systems and/or ensuring the delivery of M&E products that meet the requirements of the demand side.
- **Enabling-environment.** Senior DFAT management who must balance M&E against other priorities and ensure adequate

[29] The Information Systems Group, *Vision and Value*, 'Evaluation of information systems and monitoring arrangements for the programmes supported by the European Social Fund,' (2007). http://www.ec.europa.eu/social/BlobServlet?docId=2202&langId=en

[30] DFAT, 'DFAT Investment Level Monitoring and Evaluation Systems Report (final version completed), unpublished internal report.

support and resources are available to demand-side actors. The enabling environment also extends to broader structures and institutions that senior management works within, industry norms and standards, and the extent of agreement about reasonable standards of M&E between demand and supply actors.

The underlying assumption reflected in the analytical framework was that successful M&E requires the demand side, supply side and enabling environment to work in concert. For each of the three M&E stakeholder classes, three core elements for investigation were defined and a subset of 25 criteria elaborated against which ratings were awarded. This gave structure to primary and secondary data collection.

Alternative analytical models

While the literature[31] acknowledges that one-size-does-not-fit-all, a framework can serve as a useful guide or diagnostic tool to monitoring system development. It can do so in planning, assessing progress and identifying gaps, and as a communication vehicle and springboard to inform and educate technical and non-technical personnel on the ways knowledge building and innovation are being or can be introduced into aid investments.

In addition to precedent models used in ODE evaluations, alternative frameworks were explored through the literature review for this evaluation. Three key models were considered in further detail. These are the:

» **Lahey's model**[32] for developing an effective M&E system in the public sector. This framework identifies four broad building blocks needed for an effective monitoring system—vision, enabling environment, infrastructure to supply monitoring information, and infrastructure to demand and use monitoring information. A set of 12 critical success factors recognise the importance of positioning monitoring within a broader context rather than simply in a technical one. This recognises the political support factors needed to launch and sustain an effective monitoring system. The 12 factors are drivers, uses, leadership, commitment, resourcing, accountability, technical capacity, infrastructure to supply monitoring information, infrastructure to use monitoring information, oversight, values and ethics, and sustainability.

» **The 7s framework**. This framework[33] was developed in the 1980s. It has been widely used by human resource managers as a key to higher organisational performance. The model's key point is that all seven areas—strategy, structures, systems, style, shared value, staff and skills—are interconnected and recognise that a change in one area of an organisation requires a change in other areas of an organisation for it to function effectively.

» **Gartner's Business Analytics framework.**[34] This framework defines the people, processes and platforms that need to be integrated and aligned to take a more strategic approach to business intelligence, analytics and performance management initiatives.

---

[31] Lahey, R. 'A Framework for Developing an Effective Monitoring and Evaluation System in the Public Sector— Key Considerations from International Experience,' www.ecdg.net/wp-content/uploads/2011/12/Framework-for-developing-an-effective-ME-system-in-the-public-sector-2009_Lahey_good.doc A Framework for Developing an Effective Monitoring and Evaluation System in the Public Sector—Key Considerations from International Experience.

[32] ibid.

[33] Bradach, J. (2012). 'Organizational Alignment: The 7-S Model.' Harvard Business School.

[34] Chandler, N et al. (2011). 'Gartner's Business Analytics Framework,' http://www.gartner.com/imagesrv/summits/docs/na/business-intelligence/gartners_business_analytics__219420.pdf

CRITERION BASED ASSESSMENT FRAMEWORK

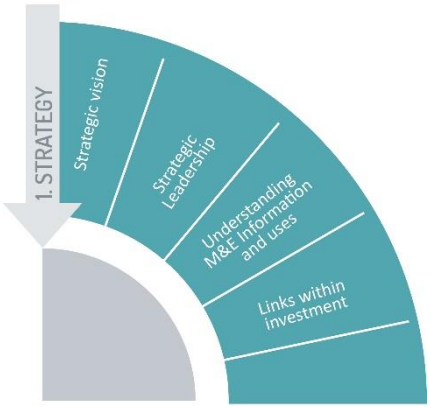
The CBAF (Figure 2) provided the evaluation team with a structure based on the determinants of quality for investment monitoring systems from which to assess DFAT investment-level monitoring systems implemented by managing contractors. The analytical models considered in the previous section were analysed for their fit-for-purpose and adapted to create a simplified hybrid framework for application to the DFAT context and for this evaluation. Each domain and component are summarised in this section.

Figure 2: 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. 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.



The strategy domain elements describe how:

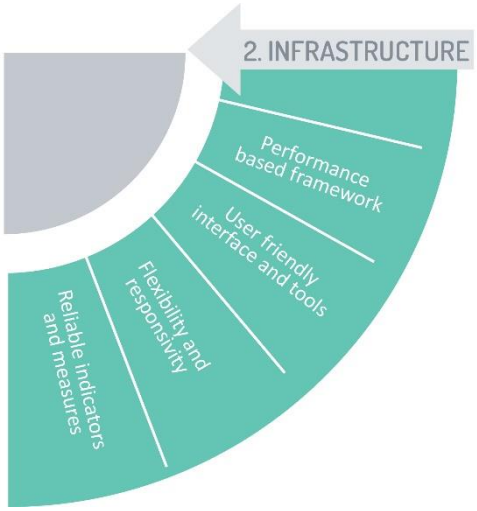
- ✓ the vision for investments is collaboratively developed to be realistic, challenging and relevant
- ✓ strong political support is required for sustained leadership and ownership
- ✓ investment managers use information appropriately to manage investments and the M&E system to achieve improvements
- ✓ theories of change provide adequate detail to enable partners to use it to guide their implementation.

Infrastructure domain

The infrastructure domain describes the infrastructure needed to help ensure a systematic, comprehensive and credible approach to M&E.

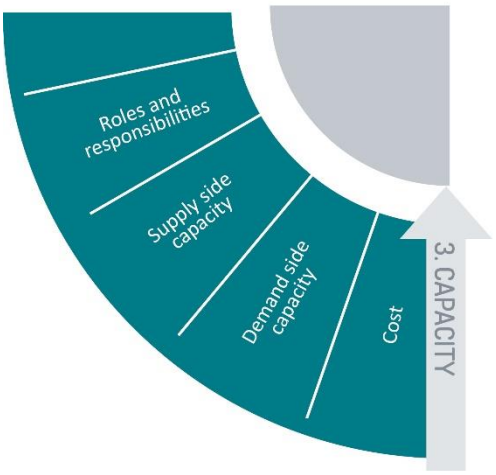
The infrastructure domain elements describe how the:

- ✓ quality of information required by actors in the monitoring system depends on its relevance and, therefore, its usefulness
- ✓ reliability of the system contributes to its quality and is the direct function of its coverage and the inverse function of the average size of errors and their frequency
- ✓ quality of the information architecture reveals whether the system is integrated or segmented and also its flexibility.



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.



The capacity domain elements describe how:

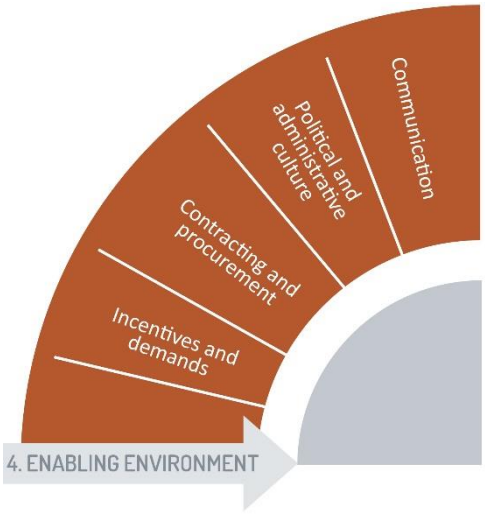
- ✓ investment managers demonstrate effective resource management to achieve results
- ✓ policies and standards clarify roles, responsibilities and accountabilities for performance monitoring, establish expectations across the system for timing and level of reporting and set out quality standards for M&E conduct
- ✓ design of the system needs to be responsive to the information needs of its users, determine the resources available to build and sustain the system, and assess the capacities of those who will produce and use the information.

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.

The enabling environment domain element describe how:

- ✓ political support is needed as an essential driver to launch and resource monitoring systems, lead changes in organisational culture that may be needed, provide champions, ensure an enabling environment, and provide the basis to help ensure the M&E system is sustainable
- ✓ incentives and contracting mechanisms can work to support structural changes that enhance quality
- ✓ communication and participatory processes support greater ownership and sustainability of monitoring systems.



ACRONYMS & ABBREVIATIONS

ADS	Automated Directives System
CBAF	Criterion Based Assessment Framework
DFAT	Department of Foreign Affairs and Trade
DfID	Department for International Development
DQA	Data Quality Assessments
M&E	Monitoring and Evaluation
ODE	Office for Development Effectiveness
UNAIDS	Joint United Nations Programme on HIV/AID
USAID	United States Agency for International Development