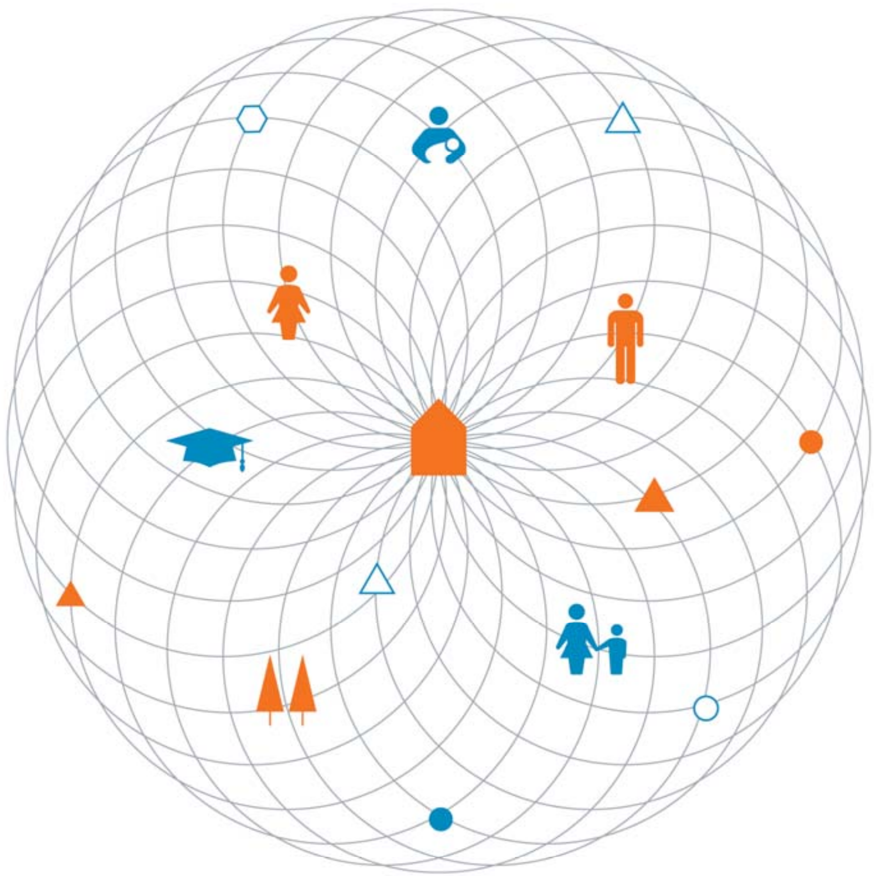
Coffey's submission on Performance Benchmarks for the Australian aid program

Consultation Paper



Global change
starts at the
heart of our
communities

Summary

Coffey is focussed on delivering value for money, quality outcomes and effective results for its clients.

We welcome this consultation process and the opportunity to share our experiences and learning. It recognises many areas where performance management and benchmarking could be enhanced to improve the effectiveness and quality of Australian aid.

We have identified 10 recommendations from a managing contractor's perspective, that we believe will improve accountability of the aid program; increase overall performance and value for money; better integrate investments with foreign and trade policies; and help justify investments to partner governments and taxpayers.

On defining and assessing the aid program, Coffey recommends:

1. Clearly articulating a performance management framework based on a hierarchy of outcomes and achievements that provides the frame of reference for implementing partners, and clearly justifies investments to the Australian public and international community. It would include;

    a. Australia's national interests

    b. Global commitments such as the millennium development goals and post-2015 targets

    c. Regional (Asia pacific) economic, social and environmental obligations

    d. Country program goals including policy and trade dialogue

    e. Program and sector wide objectives

    f. Project specific objectives

2. Developing performance benchmarks and indicators at the country, program and project level that are contextually framed, including degree of difficulty (risks and constraints), to accurately reflect achievement, contribution and value. They should be designed to encourage continuous improvement, innovation and adaption to locally changing circumstances and new research. They should take into account positive and negative externalities and unintended consequences to reflect costs and level of effort.

3. Within the performance management framework, the definition of what constitutes success, at all levels, needs to be clearly articulated, agreed by implementing partners, and frequently reviewed and reaffirmed. A common issue during implementation of many interventions is that each partner has a different view of what success means, and therefore different perspectives on what performance measures and benchmarks are relevant.

4. Producing a public document, "The Impact of Australian Aid" every 3-5 years to summarise the impact the aid program is having across the region and the progress being made through effective aid delivery. This reference would summarise progress against performance benchmarks or indicators, contextually framed, with contributions from all implementing partners. Completion reports and independent evaluations would be a source of material through capturing lessons and to frame stories. It would encourage longitudinal studies focussing on the contribution of Australian aid to improving regional economic, social and environmental conditions over a longer timeframe.

On linking performance to the aid budget, Coffey recommends:

5. Performance incentives be developed on a case by case basis. While it is generally acknowledged that incentives are good to direct funding towards good performance, it can also lead to perverse behaviours. If performance indicators are contextually framed according to degree of difficulty, then it follows that performance incentives should also be contextually framed and carefully defined to drive successful implementation.

6. Implementing “Design and Proof of Concept” projects as a precursor to scaling up larger interventions. Interventions are often based on a program theory that incorporates many assumptions about how the world works. They are models that are only proven when implemented in the real world. Linking funding to the results of a “Design and Proof of Concept” stage, allowing for refinement of the design and performance benchmarks during this phase, may reduce uncertainty and increase the specificity of programs to allow contracting with lower risk and greater confidence in partnership arrangements.

7. Adapting to changing circumstances to increase effectiveness and performance through the use of flexible contracting arrangements should be rewarded. Despite good intentions, contractors can be penalised by trying to adapt to changing circumstances due to fixed contractual arrangements.

On improving the assessment of the performance of implementing partners, Coffey recommends:

8. Objectively and independently assessing the performance of the partnership (including DFAT, partner governments and managing contractors) within the context that each operates and appropriately resourced to meet the objectives. The current focus is the contractor performance assessment (CPA) which is directed towards the managing contractor yet many aspects of performance are beyond their control, with other partners playing a significant role in decision making during implementation, affecting resources and risks. A fairer process should involve a performance assessment of the partnership.

9. Independent assessment of the performance of implementing partners (including managing contractors) within the context of the performance of the partnership, in particular the extent to which the partner government is meeting its commitments and obligations.

10. For large programs and sector-wide initiatives, appoint independent evaluation managers that work in parallel with implementing partners from the design phase to a period beyond implementation to assess the efficacy and effectiveness of design and implementation, the performance of the partnership arrangements, and the potential for sustained benefits.

Discussion and Rationale

How should performance of the aid program be defined and assessed?

A Performance Management Framework

Performance management systems are seen by implementers as necessary to determine the level of success of a program, and the contribution each actor makes to higher level goals, such as the millennium development goals, regional security, improved economic conditions and poverty alleviation. Measuring our contribution should focus on Australia's prestige as a donor of merit; its strength in building regional relationships and its position within the Asia-Pacific as an influencer for growth, development and regional stability.

From an implementer's point of view, the performance of the aid program should be defined and assessed within a clearly articulated framework that connects the logic of a project or program to agreed higher order outcomes at the regional and global levels. In this context there needs to be clear evidence of the contribution of each project or program is making to the overall goals of the aid program and Australia's national interests. This evidence needs to be in the form of key performance indicators that are measurable, defensible and expressed in terms of progress and quality and not just quantity.

It is difficult to clearly attribute specific long term development gains to a particular intervention unless the overall framework is clearly mapped out in relation to other programs and activities within specified timeframes. Such a framework helps explain the purpose of a particular intervention to the Australian public, which it may not understand if treated in isolation. It helps explain Australia's contribution to development progress which is necessary to also appreciate levels of accountability for results. Similarly it is important to link aid programs and activities to partner government programs and the extent to which this support is successful in promoting policy dialogue for reform.

A performance management framework provides a documented structure for reporting on the level of success and performance of actors, including demonstrating value for money to Australian tax payers.

Recommendation 1 - Performance Management Framework

A performance management framework for Australian aid should be based around a hierarchy of results and achievements that supports justifying investment to the Australian public against:

- *Australia's national interests*
- *Global commitments – honouring global commitments to the millennium development goals (MDGs) and international agreements on aid effectiveness. Evaluating the contribution Australia making in comparison to other countries.*
- *Regional obligations – regional security, poverty alleviation, trade, economic development, social welfare and environmental protection*
- *Country program goals – establishing bilateral relations and opening policy and trade dialogue*
- *Program (sector-wide) objectives – achieving local goals across sectors such as health and education. There are many interrelated projects being conducted by a range of stakeholders, each one contributing to similar overarching goals. Evaluating the impact Australian aid is having by comparison to other actors and being recognised for its contribution.*
- *Project specific objectives – the success or otherwise of particular interventions in terms of value, policy reform, sustainable benefits, innovation, learning and replication.*

Key Performance Indicators and Benchmarks

Generally performance indicators, benchmarks and baselines are necessary to drive continuous improvement and provide information to decision makers (and constituencies) about priorities for funding; what is working; and what is not. However, performance is dictated by many factors including many vested interests that can distort actual measurements. Simplification may lead to a disregard of these factors resulting in recommendations for improvement that have little impact.

Robust and rigorous M&E systems are needed to understand the performance of an intervention taking into account internal and external influencing factors. It is also important for tracking and reporting on changes to particular problems to address the root causes and not just the symptoms.

Development programs operate within open systems. So while the efficacy of the design and performance of implementers may be high, the results may be poor due to factors beyond the control of the implementers or contractors. These need to be taken into account, as well as the externalities, or costs and benefits, to external parties that did not choose to be involved and are not factored into resourcing. It is often dealing with these externalities (such as corrupt action by third parties or local political imperatives) that can unexpectedly affect resourcing and results. Managing contractors are generally used to manage these types of risks but they are not necessarily accounted for when assessing performance.

Performance indicators should be established on a project by project basis and related to changes in behaviour and achievement (as a function of the operating environment) and what can be realistically achieved rather than in absolute terms. We should be looking at the rate of (or ability to) change and whether these changes are persistent or not.

Key performance indicators need to be carefully chosen at the project and program level to drive goal attainment and contribution to overarching impacts. They should be determined:

- Within the context of achieving sustained benefits, altered behaviours and institutional change.
- Through consultative processes and "proof of concept" approaches to ensure they are realistic and to avoid introducing perverse behaviours (performance indicators being achieved at the expense of longer term outcomes and impacts, or easy pathways chosen (the path of least resistance) at the expense of hard to reach target beneficiaries).
- By being contextualised within the political and cultural environment, and measured objectively relative to the risks, constraints and barriers that need to be overcome during implementation.
- Specific to each operational environment and measured against a baseline to determine rate of change rather than absolute values.
- Through evaluation of each phase of a program to understand change over time against inputs and processes.

Performance indicators can be perceived by some as limiting innovation and the opportunity to build on success. Key performance indicators should include measures for innovation, lessons learned and dissemination of knowledge to the broader community.

There should be significant value assigned to understanding failures and why things did not work, as much as there is to celebrating successes and documenting what worked.

Factors affecting the achievement of performance targets also need to be quantified or qualified to place the level of success within context, and with the focus firmly on measuring potential positive impact on target beneficiaries.

As noted in the Aid Effectiveness Review (2011), performance indicators need to judged against degree of difficulty. Training female teachers in remote areas of Afghanistan is significantly different in terms of level of effort and approach compared to training female teachers in the Philippines. Yet the current headline indicator does not differentiate this, nor does it recognise that teacher training alone does not guarantee achieving any impact on target beneficiaries (ie disadvantaged children in remote areas). To appreciate the true value of achievement an indicator system should be used to reflect the challenges (risks and barriers) and potential impact (change to people's lives) that meeting targets have.

Factors that affect achieving targets, relating to degree of difficulty, need to be factored into performance indicators:

- Readiness for change of the target organisation or beneficiaries
- Actual demonstration of committed leadership
- Aligning interventions politically and culturally
- Conflict and fragility
- Realistic ability of counterpart governments to commit resources and staff to the program and replicate or scale up the intervention
- Realistic ability of counterparts to sustain the capacity to perform key functions resulting from interventions.

There is a propensity to measure physical components of a program or more tangible intermediate results and outputs and use them as key performance measures or headline indicators. These are usually easy to measure and provide tangible results for public reporting, such as number of people awarded scholarship or kilometres of road built. However, these can detract from the actual attainment of program goals such as improved governance, improved delivery of basic services by partner institutions, or building the capacity of local governments.

For instance contractors or implementers can train teachers and build classrooms, health centres or roads to achieve tangible targets without regard to improved governance or capacity of partner institutions, yet the program may be deemed a success by reporting on these achievements in isolation. Once the intervention stops, the benefits may deteriorate rapidly through lack of local capacity and maintenance. Measuring sustained benefits is more relevant to the success of aid.

Key performance indicators and benchmarks should be weighted towards institutional, social, economic and environmental reforms and policy changes (softer targets). The Australian Government needs to find ways to “sell” these to its constituency as critical to Australia’s and the region’s best interests as well as the typical headline indicators. Although key performance indicators such as policy and institutional reforms, transparency, accountability and improved capacity may not be as newsworthy, they can be more important for long term change, regional stability and economic growth than physical infrastructure achievements.

Performance benchmarks need to be expressed as a function of risk and constraints. Risks need to be factored into performance management and key performance indicators in order for valid comparisons to be made between locations and environments. They need to be location specific and contextualised.

Performance benchmarks which do not recognize project context can be misleading. In many assessment systems generic or abstract criteria are universally applied to assess performance of projects. This includes the modifications of standard DAC criteria (Relevance, Effectiveness, Efficiency, etc.), which are now applied to evaluations, reviews, QAE and QAI. While this may be a useful approach for broad comparison, it makes objective and contextually relevant assessment extremely difficult in many cases. To ensure that contextual issues can be appropriately considered, ‘achievement of outcomes and objectives’ must be reasserted as the primary focus of all performance assessments.

Current standardized approaches to assessments make little or no distinction between different forms of aid. This makes objective assessment difficult because different forms of aid legitimately have different priorities. For example ‘success’ under emergency responses or in conflict settings needs to be defined quite differently from ‘success’ in a ‘standard’ development setting. Such distinctions are currently not made (note the standard DAC criteria mentioned above were not designed to be used in emergency or conflict settings). Not only does the current ‘one-size-fits-all” approach discount many significant achievements in emergency and conflict settings, it also sets up incongruities between the existing assessment framework and other, more context sensitive, official policies (including humanitarian and fragile states relevant policies). This incongruity renders objective assessment unfeasible, because it is simply impossible for projects to satisfy two sets of divergent policy.

Performance indicators should be established and agreed at the beginning of a program with widespread consultation and sensitivity testing. They need to be realistic and constraints understood and appreciated. Agreements should not be reached with partners until the M&E plan has been established and targets investigated within the context of the program and local political economy. This may take 6-12 months for a complex or long term program. Ideally this should commence during the design phase of a program and extend through a "proof of concept" or pilot stage.

Recommendation 2 – Key Performance Indicators and Benchmarks

Develop performance benchmarks and indicators at the country, program and project level that are contextually framed, including degree of difficulty (risks and constraints), to accurately reflect achievement, contribution and value. They should be designed to encourage continuous improvement, innovation and adaption to locally changing circumstances and new research. Indicators should also account for positive and negative externalities and unintended consequences to reflect cost and effort.

Defining Success

Development programs are based on a program theory and design that make many assumptions and contain a low degree of specificity. They are based on conceptual models of how the world works and how it will change with specific interventions. There is a high degree of subjectivity with stakeholders often having differing views of what success, and therefore good performance, looks like. These views can be personality or relationship driven, as well as defined by time frames.

Most designs and programs are aspirational with targeted goals beyond the life of the program. They are constantly changing due to dynamic circumstances which, in terms of implementing contractors, reset commercial relationships at various points in the process and therefore changing what constitutes success and good performance.

There is a need to increase the dialogue between partners at the early stage of the program to clearly articulate what success looks like. The intended results of the intervention need to be clearly articulated and explained in terms of levels of performance, timeframes and quality of results. There is a need to be realistic about the development objectives by doing frequent reality checks and revising these objectives under changing circumstances, which may require challenging previous political commitments.

Recommendation 3 – Defining Success

Within the performance management framework, the definition of what constitutes success, at all levels, needs to be clearly articulated, agreed by implementing partners, and frequently reviewed and reaffirmed. A common issue during implementation of many interventions is that each partner has a different view of what success means, and therefore different perspectives on what performance measures and benchmarks are relevant.

Justifying Aid Investments to the Public

The Aid Effectiveness Review (2011) found "the absence of a single, easily comprehensible scorecard on the effectiveness of the Australian aid program as a whole". It is debatable as to the wisdom of such an approach. A single scorecard that the public understands will contain little useful information and require "dumbing down" very complex issues into a set of broad ratings which may be used to direct aid investment through public opinion.

Consideration should instead be given to producing an *"Impact of Australian Aid"* report that is comprehensive enough to provide scorecards by country, sector and program linked to an integrated performance management framework to demonstrate how Australia's best interests are being served and the difference our aid makes to the region through various partners. This could be similar to the "State of Environment Reports" produced by government using baselines, benchmarks and the degree of difficulty operating in each environment (through an analysis of the local political economy and socio-economic indicators). It should focus on the rate of change of key performance indicators (not absolute values) for each country or location and sector. It should be an indicator of development progress in relation to partner government performance and policy commitments.

If produced every 3-5 years it will encourage longitudinal studies into the sustainable benefits of aid and follow up studies (post implementation), making partners more accountable for longer term impacts and outcomes. Material would be sourced from independent evaluations and completion reports, as well as development partners funded by DFAT. This could be a point of convergence for the international development industry to celebrate the success of Australian aid in our region, highlighting impact, progress, lessons and innovations. Such an initiative may help drive improvements in overall development communications by managing contractors, knowing that results may be published.

Coffey has produced a publication called "Changing Lives – Stories from the Australia Papua New Guinea Incentive Fund (2000-2008)". This type of publication turns aid interventions into real stories about how people's lives are changed through aid. It is supported by performance statistics. It makes understanding Australian aid more accessible to the general public. While annual "Aid Effectiveness Reviews" are useful in terms of accountability, they are generally less appealing to the general public and do not involve contributions from the broader international development industry.

Recommendation 4 – Justifying Aid

Produce a public document, "The Impact of Australian Aid" every 3-5 years to summarise the impact the aid program is having across the region and the progress being made through effective aid delivery. This reference would summarise progress against performance benchmarks or indicators, contextually framed, with contributions from all implementing partners. Completion reports and independent evaluations would capture lessons and frame stories. It would encourage longitudinal studies focussing on the contribution of Australian aid to improving regional economic, social and environmental conditions over a longer timeframe.

How could performance be linked to the aid budget?

Linking investment to performance

Linking investment to performance is necessary for any government or business to demonstrate it is delivering value for money and achieving a good return on investment. Aid should be no different.

Within the international development sector, however, and from the view of an implementer, it is how performance is measured and assessed that is of the biggest concern given many intangibles and long time frames to achieve outcomes and impacts.

In some instances linking the level of investment to performance, particularly outputs, can drive perverse behaviours and undermine the achievement of program outcomes and impacts. It can also stifle innovation and ability to adapt to changing circumstances.

It is generally considered good to incentivise aid to a certain extent but this needs to be done on a case by case basis (project by project) and contextualised to the local political economy and operating environment. Blanket targets attached to incentive payments in certain environments can lead to perverse behaviours and resources being redirected to meet aid targets instead of essential services.

Performance targets need to be carefully structured and interlinked to achieve the long term results or outcomes. Otherwise short term gains may be sought at the expense of longer term outcomes. For instance the number of classrooms built may be linked to partner government capacity to maintain them.

Incentivising partner government performance through performance targets also needs careful consideration. Past experience shows that poor performance is often not penalised even though targets are not met by partner governments in order to preserve good working relationships. Similarly targets that are not met by a partner organisation can result in victimisation or blame being attributed to specific individuals in the organisation if significant funding is involved. This acts as a disincentive for others to take up the challenge in the future.

For managing contractors it is possible to have part of the management fee performance based. However, accountability for meeting performance targets and results needs to be clearly defined in terms of the ability to control and influence decisions that affect achieving targets. Many of these decisions are outside the domain of a managing contractor and therefore they should not be held accountable under these circumstances. Understanding program governance – who is making critical decisions on delivery and inputs – is required by all actors in order to appreciate how performance is affected. In linking investment to performance this means appreciating how the partnership as a whole and they contribute each makes (DFAT, Partner Government and implementing contractor) contributes to overall performance. Managing contractors cannot be held accountable for long term results but they can be held accountable for maintaining the integrity of the design, scope of services and implementation plan to achieve those results.

In terms of partnership performance it is important to understand what financial contributions the counterparts are making and their invested stake in the program. Similarly DFAT's contribution can be significant outside of aid allocations. There are many hidden costs absorbed by DFAT and other government agencies to manage and implement a program that are not necessarily counted towards the overall aid investment.

Recommendation 5 Linking performance to incentives

Performance incentives be developed on a case by case basis, and contextually framed. While it is generally acknowledged that incentives can be useful to direct funding towards good performance, it can also lead to perverse behaviours. If performance indicators are contextually framed according to degree of difficulty and the local environment, then it follows that performance incentives should also be contextually framed and carefully defined to drive successful implementation.

Design and Proof of Concept Approach

Many environments in which we work are highly dynamic and influenced by the local political economy and particular features of the operating environment. Intervention designs can take 2-3 years and are based on political scenarios that may have been valid at the time. Designs are based on a program theory, or theory of change, that depends on a range of assumptions. Performance indicators are derived from the theory of change to reflect desired outputs and outcomes. From experience rarely do these assumptions all prove true, and often the program design requires change and reinterpretation. This can affect resourcing during implementation; the need to sometimes radically adapt; and a revision of performance indicators.

Often programs are rapidly mobilised with monitoring and evaluation frameworks designed within 3 months of commencement. This is generally insufficient time to address aspects such as the quality of partner government systems; test all the assumptions; evaluate the quality of available data sources; establish accurate baselines; confirm the efficacy of the design; and validate the robustness of program theory in the field. Contracts should therefore allow for performance indicators (and M&E Plans) to be revised after 6-12 months through a reality check of development objectives and projected results. Larger projects should have a "proof of concept" phase to iron out any design issues prior to scale up. Although there is pressure for a program to commence as soon as possible once the design is finished, there is good justification for a "Design and Proof of Concept" contract in some instances. From experience some more complex designs take 2-3 years to reach the implementation stage due to uncertainty, risk and local dynamics. A "Design and Proof of Concept" approach may actually speed up the process and reduce uncertainty and risk.

From an implementing contractor perspective one of the biggest issues affecting performance is adherence to a design that won't meet the set outputs or achieve desired results due to changed circumstances or false assumptions. Faulty designs can come from designers who may not have had the full engagement from partner governments. There may have also been several political changes in the country since the design was created. This needs to be addressed early on, between the managing contractor and client.

One option would be to establish "Design and Proof of Concept Projects" prior to scale up or full implementation. This could be based on an initial Concept or Pre-feasibility Study. The Proof of Concept (PoC) phase would test the design and implementation strategy, including the capacity and commitment of counterparts. Following PoC the program would be redesigned and scaled up for implementation. KPIs would be based on the local context and operating environment. Local benchmarks and baselines on performance would be well researched and applied to measure rate of progress or improvement.

Recommendation 6 – Design and Proof of Concept Projects

Implementing "Design and Proof of Concept" projects as a precursor to scaling up larger interventions. Interventions are often based on a program theory that incorporates many assumptions about how the world works. They are models that are only proven when implemented in the real world. Linking funding to the results of a "Design and Proof of Concept" stage, allowing for refinement of the design and performance benchmarks during this phase, may reduce uncertainty and increase the specificity of programs to allow contracting with lower risk and greater confidence in partnership arrangements.

Adaptive Management

One critical aspect of performance improvement is the ability to adapt quickly to changing environments in response to performance indicators. All development projects operate in a highly volatile operational and political environment. There needs to be the ability to adapt to external factors in order to optimise performance. Contractually this is currently quite difficult. While necessary improvements can often be identified quickly, implementing the contractual changes through variation orders (or through formal partnership agreements) can take time due to due diligence processes. Often the window of opportunity for significant improvement can be short and delays have further detrimental impacts on progress. When the necessary contractual changes are made the impact of the changes on improved performance is sometimes diminished. Often contractors may respond and then be financially disadvantaged creating a disincentive for future adaption or innovation.

An adaptive management framework needs to be introduced to all programs to enable adaption to changing circumstances and to be more responsive. This is necessary to ensure investment is linked to performance relative to changing environments, innovation and improved knowledge.

Recommendation 7

Adapting to changing circumstances to increase effectiveness and performance through the use of flexible contracting arrangements should be rewarded. Despite good intentions, contractors can be penalised by trying to adapt and improve outcomes due to fixed contractual arrangements.

How can the assessment of the performance of our implementing partners be improved?

Assessing Partnership Performance

Understanding the stakes of the stakeholders in the program is necessary to appreciate how performance assessment may be perceived and controlled. There needs to be open and transparent discussion about partnership performance (DFAT, partner governments and the managing contractor) otherwise the ability to implement positive improvements that could enhance program delivery, will be limited.

There needs to be broader consultation with stakeholders using an objective approach that acknowledges the role of the partnership in overall performance. A 360 degree approach could be used so that the contribution of partner governments and DFAT make to performance is evaluated as part of the process.

For example DFID welcomes feedback on their performance from managing contractors during the assessment process. If we are serious about overall performance improvement then ‘the implementing partnership’ needs to be assessed in an open and transparent process. DFID also commissions independent assessment teams who go out to each project annually to evaluate the performance of programs. A recent example in Pakistan involved 12 evaluators interviewing stakeholder’s and government partners, examining policies and procedures, and outputs against targets. AusAID have done a similar thing in the past where there was no buy in or agenda from the assessment team as an independent unit. It can be seen as a very good process – transparent and objective. It is important that all stakeholders have a voice and points of difference are adequately explored within the proper context.

The Independent Review of Aid Effectiveness (2011) found that DFATs system of independent evaluations is not working well and requires reform. From a managing contractor’s perspective it appears that independent evaluations as they are currently performed are under resourced and lack systematic review and structured analytical frameworks that are fit for purpose. There should not be a simple standardised approach but a carefully designed research methodology appropriate to the intervention, issues and partnership arrangements. Generally they rely heavily on value judgements from experts with past experience that may not apply in current settings. Due to resource and time constraints they usually do not use statistically reliable data or a comprehensive evidence base that is fully cross-referenced.

The performance of the partner government in meeting program commitments needs to also be reinforced. It appears quite common for commitments to soften during the course of a program, particularly cash contributions and capital equipment that may have been committed to gain support/approval for the intervention initially. This can leave the managing contractor and DFAT exposed to additional costs and risks. Judging partner government performance is difficult for diplomatic reasons, however, consideration needs to be given to performance benchmarks that reinforce:

- Demonstrated commitment to agreed policies and principles as a basis for mutual cooperation
- Agreeing to participate in strategic assessments of program progress to ensure they are meeting their goals or demonstrating progress in the right direction
- Agreeing to participate in performance feedback on supported programs
- Honouring commitments to policies, procedures and resourcing for bilateral programs
- Withdrawing financial support or terminating programs if commitments are not honoured or progress is not made

Recommendation 8

Objectively and independently assessing the performance of the partnership (including DFAT, partner governments and managing contractors) within the context that each operates and appropriately resourced to meet the objectives. The current focus is the contractor performance assessment (CPA) which is directed towards the managing contractor yet many aspects of performance are beyond their control, with other partners playing a significant role in decision making during implementation, affecting resources and risks. A fairer process should involve a performance assessment of the partnership.

Assessing the Performance of Managing Contractors

Currently there are two assessments being performed by DFAT as the organisation transitions. The original CPA included 4 major criteria (Organisation, Approach, Personnel, Project Management) and 18 sub-criteria. The transitional assessment criteria now include: delivering lasting results and impacts; maximising value for money; working collaboratively and facilitating communication; be accountable and transparent; effectively managing risk and fraud; and support and implement government policies and priorities.

Irrespective of the criteria, there needs to be a consistent approach to assessing performance so that it is objective and not subject to personal bias or quality of relationships. Ideally it should be based on an independent review of program performance and the partnership with a distinction made between the two. Often DFAT (previously AusAID) will use the results from independent evaluations (mid term reviewsif available) to input to contractor performance assessments. This is a positive practice to moderate assessments but they are not always available. Coffey has noticed distinct differences in practice and scoring between locations (the same level and quality of service scores differently across the region) hence the need for consistency in approach.

The rationale for using independent evaluators to assess performance of contractors is based on the need to preserve professional working relationships of the partnership and achieve an objective and constructive result. Currently a lot of responsibility resides with local staff in DFAT to rate performance of the contractor without necessarily having a detailed operational knowledge. Often they do not have the evaluation experience, technical expertise or subject matter knowledge to contextualise the response. The responses are often subjective and weighted towards a specific issues. These factors can cause tension within the partnership. This is critical for managing contractors when performance is linked to milestone payments or performance is used to influence subsequent contracts. It is considered more objective when it can be linked to independent evaluation reports.

Coffey conducts internal project “health” checks and continuously monitors project performance. In some instances the perspectives of the client and contractor are substantially different when it comes to performance assessment. The current CPA process invites managing contractors to provide comments but it is unsure how this information is taken into account. It is not a completely transparent process. There is uncertainty about how disputed assessments are reconciled. Contractors also need to know what they are being assessed on. Not all areas are treated equally. The areas of focus should be clearly explained and the performance assessment process clear. Some projects are difficult to assess (e.g Governance) so agreement is required on how such a program will be assessed.

Currently CPAs are carried out annually. In some instances it may be more appropriate to carry out 6 monthly independent assessments so that the program can adapt in response to changing circumstances. These should be fit for purpose depending on level of investment and reputational risk, the changing environment, and level of assessment required.

From experience the areas of assessment should emphasise:

- Achieving quality results and delivering value for money within the context of the operating environment. This requires a clear definition of what constitutes both quality of results and value for money on a case by case basis. It does not just mean being “cost aware” and demonstrating savings, but understanding what drives productivity and performance to get the best return on investment. This includes efficiency, effectiveness, economy and equity.
- Potential for sustained benefits – establishing capacity, policies, systems and processes for sustaining the benefits should be emphasised over short term gains.
- Risk assessment management – Government usually manages risk by passing it onto a private contractor. How risk is managed is critical to overall performance and diplomatic relations. The contractor has a role in protecting the reputation of government and representing the government in-country. We should be judged on how well we manage this and advise the government on risk exposure. However, if the Government subsequently gets involved in day to day operations of the contractor, then it needs to also share some of the accountability and risk. This needs to be taken into account (hence the need for a partnership evaluation).

- Collaboration and value adding – creating greater value through collaboration with other programs, donors and implementers to deliver more with less. Coordination and collaboration with donors is usually not factored into programs as a costed activity yet it is necessary and can pay high dividends for partners.
- Innovation and learning – introducing radical or incremental innovation to change the way programs are delivered and sharing this information to the broader development community
- Adaptive management – the ability of the partnership to adapt to changing circumstances or new knowledge in order to continuously improve and take advantage of emerging opportunities
- Collective efficacy – the strength of the partnership to work together to overcome obstacles and challenges to achieve goals

While the idea of linking past performance to the awarding of future contracts is not new, and generally a standard evaluation criteria for awarding contracts across a range of disciplines, it cannot be too specific or limited to only DFAT's CPA performance assessment. This would limit the pool of applicants for a bid and make it difficult for new entrants to enter the market. If the CPA is used then it needs to look at aggregated performance taking into account an analysis of variance across contracts and over a period of time (for instance a contractor that has only done one contract for DFAT and scored highly, may rate well compared to a contractor who is a consistent performer that has done many contracts under many difficult environments). Using a variety of sources, as is currently done, is probably the fairest process.

Recommendation 9

Independent and objective assessment of the performance of managing contractors within the context of the performance of the partnership should be undertaken. This will reduce subjectivity and remove bias influenced by the nature of personal relationships between the contractor, partner government and DFAT.

Independent Evaluation Managers

To understand what works and why requires a robust and rigorous M&E which can run either in parallel as an independently funded service, independent evaluation panel, or embedded within the program as part of the managing contractor's M&E arrangements.

The concept of an independent Evaluation Manager or Evaluation Panel, overseeing the M&E of the program is one option that could be explored if the evaluation extends to partnership performance and requires continuity throughout the life of the program. Coffey has operated as an Evaluation Manager on many large and complex sector-wide programs for DFID and there appears to be significant merit in this approach if resourced appropriately. It allows the ability to invest in longer term, statistically valid approaches, without impacting on the delivery of the program. However, there needs to be close cooperation with implementing partners and continuous engagement. It also provides for an independent adviser on changes to implementation strategies and resourcing to improve results based on performance data. The current arrangement leads to debate between the funder and the implementer (both with vested interests) about resourcing needs or changes to scope, often resulting in delays.

There are advantages and disadvantages of this approach. The advantage is that the independence should reduce bias, or perception of bias, associated with embedding as part of the managing contractor, and enable an outside view of partnership function, governance, performance and targeted beneficiaries. As an independently contracted function it could also run to an extended time frame. Ideally the Evaluation Manager should get involved from the design phase (to establish the performance indicators and baseline) and extend beyond the life of the program (to measure potential for sustained benefits). A monitoring function (and data analysis) function is still required by the managing contractor for project management, but an independent evaluation would look at the performance of the partnership, and quality of implementation and results. The disadvantage is the additional cost and contracting arrangements, although the Evaluation Manager could operate across projects and programs for a given sector to provide benchmarking that is contextualised within the same sector and country. This provides greater value for money and addresses many of the issues that current limit performance improvement.

Recommendation 10

For large programs and sector-wide initiatives, appoint independent evaluation managers that work in parallel with implementing partners from the design phase to a period beyond implementation to assess the efficacy and effectiveness of design and implementation, the performance of the partnership arrangements, and the potential for sustained benefits.