

## **More than the sum of the parts: assessing country level effectiveness in the Australian international development program**

John Winter, Jo Hall and Susan Ivatts<sup>1</sup>

### Abstract

Within the directions of the 2006 White Paper the Australian aid program is required to strengthen its performance orientation at the country program level as well as at activity level. In doing so it is facing the same challenges as other aid organisations, namely how to determine the effectiveness of aid especially in circumstances where aid is not the major factor influencing change. To compile a more convincing picture of results (and one that is ultimately more useful for management purposes), AusAID country programs must raise the levels at which success is measured from the output to the outcome level and describe modest progress over longer time periods. At the same time good practice and Australia's international commitments require greater alignment of monitoring and evaluation with the performance systems of partner countries, where the effect of a single donor is impossible to separate out. In telling a more valid story about results, however, AusAID faces an institutional and political culture that has preferred the certainty of outputs. In 2005 and 2006 AusAID implemented a process of annual review in Fiji using the principles of contribution analysis which attempted to draw together internal performance information with an exploration of the context as a way of determining effectiveness and seeing the Australian aid presence as more than the sum of its program outputs. The approach promises some benefits for program management, for consideration of the shape of future activities, and for articulating the returns from Australian aid.

---

<sup>1</sup> John Winter is the Senior Performance Assessment Adviser, AusAID. Jo Hall is a Director in the Office of Development Effectiveness, AusAID. Susan Ivatts was First Secretary, Australian High Commission, Suva 2003-2006. The views contained in this article are those of the authors and do not necessarily represent the views of AusAID.

## **More than the sum of the parts: assessing country level effectiveness in the Australian international development program**

*“The complex world of human beings cannot be fully captured and understood by simply adding up carefully measured and fully analysed parts. At the systems level (the whole program, the whole farm, the whole family, the whole, organisation, the whole community), there is a qualitative difference in the kind of thinking that is required to make sense of what is happening.”*

*Michael Quinn Patton (quoted in Iverson, 2003)*

### Background

The 2006 White Paper (AusAID 2006a) commits AusAID to increasing the effectiveness of its development programs, and therefore to being able to better demonstrate its impact. This is nothing new. All development agencies, including AusAID, have sought since the late 1990s to demonstrate not only that individual aid activities work but also that aid is a significant determinant of progress in developing countries (Binnendijk 1999 p 9). Possibly the flagship product of the concern for aid effectiveness is the World Bank’s Annual Review of Development Effectiveness. AusAID is committed to a similar report of which the first is due to be published later this year. However, as attested by swathes of literature on reporting on results of development programs, this is not simply a matter of collecting activity level performance information, assuming that it exists, and aggregating the results. It requires the collection of quite different sorts of information and a different process of enquiry that asks not only “Is what we do working?” but also “Are we getting the best possible results<sup>2</sup> from the money and staff time we are investing in the country?”.

Senior management in AusAID have made it clear that they expect significant change both in the organisational and the cultural approach to performance assessment within the organisation; and the development of the ability to account for the effectiveness of country programs, not just projects. AusAID has changed its formal structures of strategy setting and performance reporting over the last year to give effect to these expectations. These changes will be covered in the forthcoming *Annual Review of Development Effectiveness* and are outside the scope of this paper, except insofar as they signal to program managers that consistent and defensible approaches to performance reporting are required at the level of country programs. This paper illustrates through the experience of one country program how AusAID staff are adapting culturally to changed expectations.

### Issues for the Australian aid program

AusAID has had systems for setting country strategies and reporting on the results for some years<sup>3</sup>. In attempting to revitalise its vision and practical implementation measures for country program performance assessment, AusAID is following a well trodden path among international development agencies. However, while there is a good deal of practice to take into account, it is an area that all donors are struggling with and where measurement systems are less well established than they are at activity level (Ireland et al, p 427). Some of the methodological, epistemological and institutional issues which AusAID has had to come to grips with are set out in this section.

---

<sup>2</sup> For the purposes of this paper, we use the definition of results used by the Canadian Treasury Board: “The end or purpose for which a program or activity is performed and referring exclusively to outcomes”

<sup>3</sup> The last pre-White Paper iteration of these was in 2005.

### *Multiple ownership of results*

As a signatory to the Paris Declaration (OECD 2005), Australia is committed to a planning and monitoring approach aligned with the plans of its developing country partners and to collective endeavours towards agreed outcomes, increasingly relying on and strengthening partner countries' own performance systems, and harmonising with other donors. While this strengthens developing country ownership, it means that aid agencies are increasingly attempting to integrate into their results chain the data and practices from partner institutions whose capacity to collect data is weak and who are a long way from practising what would be recognised as results based management. Moreover, the normal functions of management review and audit in respect of determining results do not operate in as clear a way as they would if they were being carried out entirely for domestic accountability purposes (Ireland et al p 420). Much of the evidence needed for reporting results is owned by partner Governments and the terms of use need to be negotiated. Processes for interim and end-of-cycle evaluations have to take account of the need for joint analysis and agreement on conclusions. There is some evidence from Africa (Lawson et al, 2005) that this diffusion of ownership goes part way to eroding the well worn observation that the relationship in development assistance between the evaluating and the evaluated institutions is fundamentally unequal; but in the short term it is undoubtedly an additional complication. At the heart of this developing relationship is the paradox that the better behaved a donor agency is over alignment and harmonisation the more difficult it is to determine just whose results are being measured (Binnendijk 1999, p 11).

### *Information requirements*

Understanding of the changes that are taking place at a country level and the ways in which aid activities may be contributing to them may require new thinking about what information to collect and when. Rather than collecting only information arising from the agency's own reporting mechanisms, program managers need to map a more consistent flow of information about their context, decide what is useful, and structure a process for routine collection and analysis of data from disparate sources. These sources may vary in profundity from evaluation reports to press reporting.

### *Measuring contribution to outcomes is harder than measuring outputs*

Moving from a project based focus to accounting for results at the country program level both requires addressing the attribution challenge of outcomes as opposed to outputs and puts a spotlight on the quality of monitoring and evaluation arrangements at the activity level. Where such arrangements exist, typical findings from AusAID programs have been that

- Information beyond the boundaries of the activity (i.e. at a development outcome level) is less extensive and precise than information at output level and below
- Links to country strategy objectives are difficult to sustain because the stated ambitions of the country strategy fall well beyond the credible influence of the activities<sup>4</sup>

Both activity level monitoring and evaluation and setting strategy objectives and their corresponding success factors need attention. However, we know from experience of applying program logic to increasingly complex situations that simply improving and extending the tools used to measure progress at the activity level (as with many other development agencies, the AusAID standard is the

---

<sup>4</sup> A phenomenon known as the "missing middle" (White and Booth 2003)

logframe), for example by setting indicators for success at the country program level, is difficult. As complexity increases, progress becomes less linear and it is harder to be simplistic about the way in which one factor affects another.

#### *Complex and long term results*

Synergies between activities need to be captured (for example, donors may have more impact through sector programs on human resource management within a public service as a whole than through a project labelled Public Service reform). So does the effect of what is called *policy dialogue* – the deliberate effort of donor officials to influence partner government decision making – on top of the effect of the agency’s packaged aid activities. In some cases, the main point of the aid activity itself may be to get donor officials a legitimate seat at the table in order to be able to influence decision making.

Since the outcomes to which we aim to contribute are frequently long term and changes in higher level indicators (for example infant mortality rates) lag behind the measures aimed at changing them, we need information about intermediate processes and relationships as well as about outcomes, particularly if the performance frameworks we are working to include outcomes which are subject to higher levels of doubt about attribution (Bourguignon and Sundberg, 2007). As the level at which an agency seeks to measure higher level indicators increases, indicators are less easy to rely on and need to be supplemented by a set of research questions<sup>5</sup> that start with “How?” or “To what extent?”

#### *Changing culture and incentives*

Managing the change from collecting activity information to understanding and identifying program level outcomes is a significant challenge for country program managers. Organisational change that attempts to stimulate the use of performance information at higher levels of complexity has profound implications for the resources devoted to capturing and analysing it and for the nature of accountability in an organisation (Binnendijk p 23). It requires sustained management commitment and communication. Accounting for results at country level is intrinsically more difficult than listing outputs. Getting people to believe that the more difficult job is the one that is wanted, and that honesty is rewarded, requires a cultural shift. As Ireland et al observe:

*“The challenge of facilitating changes in the agency’s culture should not be underestimated. The entrenched set of values, attitudes and behaviours may not be oriented toward a commitment to honest and open performance reporting, a reorientation away from inputs and processes towards results achievement and in encouraging a learning culture grounded in evaluation.”*

In describing the difficulties of introducing results based management into a number of multilateral institutions, Flint (2003) notes that they arise not only from the inherent challenge of introducing such a measurement system into an unreformed environment (ie that of the governments with which donors are working) and of changing the emphasis from the project to the program level, but also from the dead weight of incentive structures within donor agencies themselves.

For AusAID the starting point is one of a set of external accountabilities grounded (before the White Paper) in activity and output information. The quality target for the outcomes for which the agency receives public funds is that 75% of activities should be rated satisfactory or better. Examination of AusAID annual reports to 2005/6 (AusAID 2006b and previous) shows that effectiveness has in the

---

<sup>5</sup> We are grateful to Paul Nichols for this distinction.

past been overwhelmingly presented through the citation of outputs. While this situation has its roots in the way that output based accounting operates in the Australian public service, it also reflects past pressures for a form of accountability well known to other development agencies which “encourages a focus on the project level and on being able to attribute the impacts at the same level.” (Iverson p 40). This underlines the profoundly political nature of attribution within development programs<sup>6</sup>. Aside from the accountability side of the equation, a desire to improve results is also at the core of wanting to understand more about what difference aid is or is not making.

### The Fiji experience

Fiji is not aid-dependent. The Australian aid program runs at about \$30 million a year. Although this makes Australia a significant donor, the program represents only about 1% of Fiji’s public expenditure. A large proportion of the Australian program is taken up with three sector support activities in education, law and justice and health.

In 2005, the program managers, anticipating the move to performance reporting required by the White Paper, set out to account for the impact of Australian aid on Fiji’s development. The approach taken was to build up a plausible association between significant outcomes recorded in the sectors where AusAID was active and the support provided through the Australian programs using the principles of contribution analysis (Mayne 1997). The way that this approach played out in one of the sector programs has been documented in a paper presented to the 2006 conference (Kotvojs 2006). The purpose of this paper is to trace how the approach provided a conceptual framework that eased the corporate challenges surrounding the introduction of performance measurement at the country level.

Contribution analysis is a way of “reducing uncertainty in our knowledge about the contribution of a program” (to quote Mayne). Its key features are the use of multiple sources of evidence, and taking account of alternative causes of observed changes (see annex). It appeared to offer a way forward for country program performance measurement in the Fiji context because

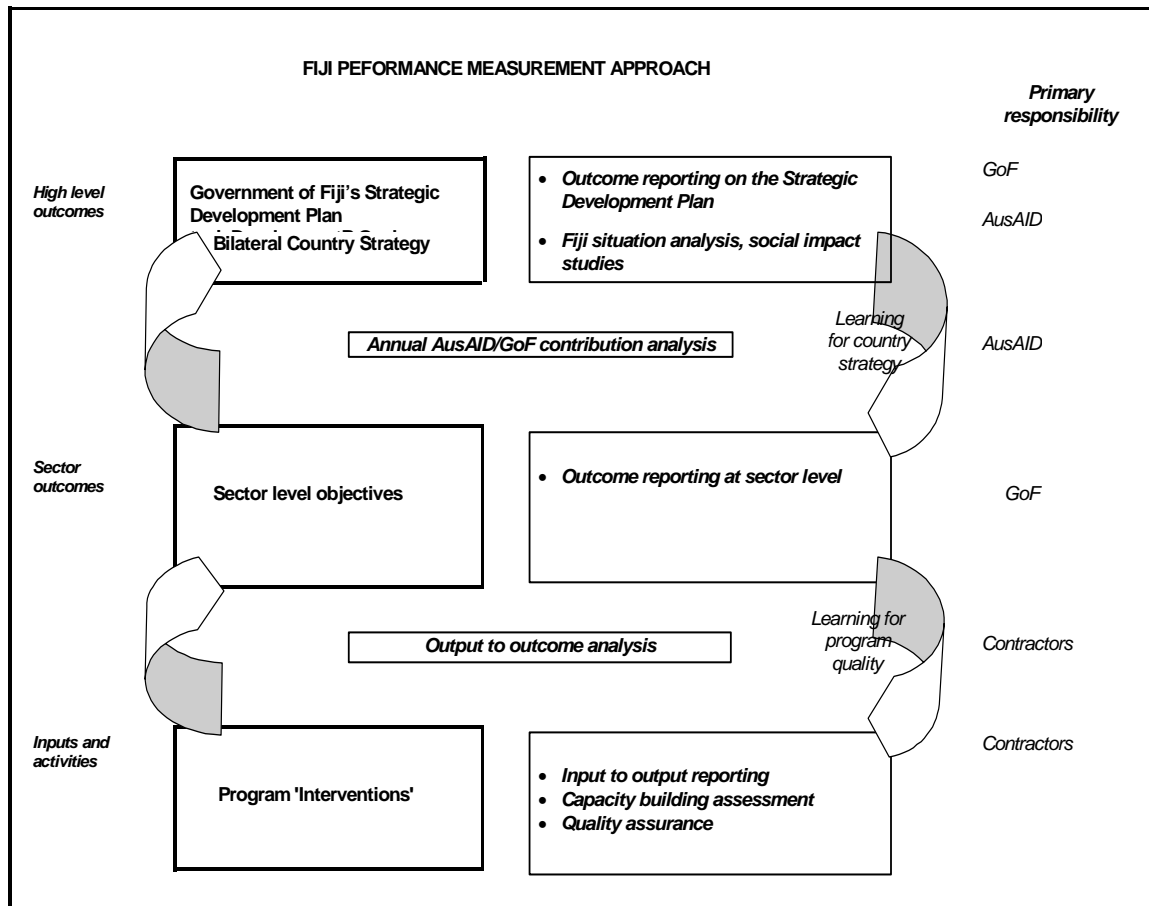
- Aid, and especially Australian aid, is only a small contributor to change
- It involved the identification of significant changes, which in principle could be done in collaboration with Fiji government agencies and using their data, and on which AusAID wanted, as participants in the development process, to base its policy dialogue with the Fiji authorities
- By requiring the exploration of alternative causes, it encouraged program managers to increase their understanding of the context and provided a framework for doing so

---

<sup>6</sup> Whether development agencies choose to tell a significant story of change with less precision about attribution or discuss less significant change but with a higher degree of certainty that aid, and particularly the aid provided by a single country’s taxpayers, has brought about the change, is highly dependent on the choices made by politicians about the level at which they choose to exercise their accountability. In the UK, for example, the period from 1997 saw the Department for International Development concentrating in its annual reporting on progress against the Millennium Development Goals. While changes at this level are highly significant and an attempt was made to impute the contribution of British aid, this approach reflected not only the system of high level public service agreements instituted by the new Government but also a political decision that the principal role of reporting was to show that progress was possible on a broad front in tackling poverty, rather than to prove the link between aid activities and outcomes. Other countries have taken a more conservative attitude to measuring attribution.

- It appeared to provide a sufficient degree of attribution to Australian support<sup>7</sup> as well as deepen understanding of what does and doesn't work effectively
- It combined the benefits of the "historical" or "top down" approach to country program performance assessment (identifying change and then identifying the causes (aid and non-aid) which explain that change) and the "evaluative" or "bottom up" (seeking to trace the impacts of aid activities by extending and aggregating their internal logic)
- It encouraged program managers to assign value to non-financial inputs such as policy dialogue or donor co-ordination.

Two cycles of annual review and planning were undertaken in mid 2005 and mid 2006. By mid 2007 the coup in Fiji had changed the nature of the aid program so substantially that it was decided to suspend the annual review process while a revised strategy was developed. The process envisaged is illustrated in Figure 1 below.



**Figure 1. Approach to performance measurement for the Australian bilateral aid program with Fiji**

<sup>7</sup> It is no accident that the approach was suggested by a public servant from Canada, where political demands for attribution are similar to those in Australia.

In an ideal world, contribution analysis of development programs rests on two uncontested sets of data: observed changes (improved quality of education, improved primary health care), and the outputs delivered by the programs. The a priori logic that links the two should be articulated and not be over-ambitious (for example, it is highly unlikely that Australian support will lead other than in a marginal way to wholesale improvement in Fiji's health indicators; intermediate observed changes such as the availability of drugs or medical staff are much more likely to be explainable with reference to aid inputs, so a judgement needs to be made as to where to pitch the results of any given program).

The second category of data (outputs delivered by the programs) was naturally easier to collect than the first. Evidence for the performance assessment was drawn from (i) annual performance reports completed by private sector contractor teams engaged to implement Australia's aid programs in the three sectors; (ii) independent annual technical reviews of the three main programs completed by external technical experts and (usually) Fijian Government staff. As a part of this approach, the private sector contractors were expected to apply a more rigorous approach to collating their annual performance reports so that claims of achievements were backed with multiple strands of evidence where feasible (e.g. using official reporting data, anecdotal feedback from community members, findings from reviews completed by other organisations etc), explaining and describing through good qualitative analysis and reporting how the outputs they were delivering were making a difference. However, during the pilot period even this proved (predictably) difficult as implementers strove to use performance frameworks that were mainly aimed at providing detailed information on inputs, activities, and immediate benefits, to demonstrate progress against higher level outcomes. Nevertheless by the time of the second annual review process in 2006 the usefulness of reporting from this point of view had improved.

The first set of data, on observed changes, was much more difficult to pin down. In the absence of adequate data, for the first year implementing teams were asked to formulate up to three propositions about positive changes in the sector and to demonstrate the logic chain linking their own interventions with those changes. For example, the education sector team put forward the proposition that "the quality of leadership and management in the education system has improved". In the second year these propositions were formulated by AusAID after interrogating the very limited performance data available from the Fiji Planning Department. For 2007, following a major overhaul of Fiji's Strategic Development Plan and associated monitoring, it was hoped that better data would be available from official sources against which Australia could judge its contribution; but as noted above, the coup has disrupted this process.

External review teams and the external facilitator contracted to conduct separate focus groups were briefed to take a semi-structured approach to interviews/sessions whereby the initial question was "what, if any, are the most significant recent changes in the sector?" Follow-up questions were based on responses. After this initial discussion, the propositions were put to the respective interviewee(s) who were asked if they could support the proposition (and based on the response, then asked why they could or couldn't support the proposition) and if they agreed with it what might have caused the change. The results were recorded and possible conclusions about contribution triangulated between the contractor reports, external reviews of the programs, and responses from focus groups (the latter two were conducted during the same period).

Finally, the evidence of results was assessed at a one day workshop with the aim of capturing headline results information for the annual country report, and agreeing the implications for AusAID's next annual work program. The timing of this process around April/May each year was chosen to meet

both the GoF's annual corporate planning and budget cycle (with the aim of maximizing the use of findings and reinforcing the practical value of assessment processes to GoF) as well as providing timely contributions to AusAID's annual report to Parliament. Out of this process came a strong sense of the synergies between the programs (for example, in improving financial and human resource management across the agencies supported by Australia) and agreement that AusAID had the basis to move on to more strategic engagement with Government over public service reform.

AusAID's intention was to complete this final assessment with the Government of Fiji, providing a higher level analysis of the Australia-Fiji program overall, utilising the sector level information available, as well as additional data and analysis. However, perhaps reflecting the modest role aid plays in Fiji it was difficult to secure high level Government engagement in this process, which meant this second stage was only completed to a certain extent. As such, the 2005 final assessment was predominantly an AusAID affair. In 2006 an extended final assessment session was held with representatives from Government of Fiji, civil society and other donors where AusAID presented what it believed was its significant contribution in the respective sectors, and sought constructive challenge from the broader audience. This provided another opportunity to explore alternative explanations for changes presented and some broader reflection of progress against Fiji's Strategic Development Plan.

#### *Lessons learned*

In the absence of good data at the outcome level it is recognised that the results obtained from these first two iterations were not especially robust, although by subjecting each activity to the test of whether it had contributed to any noticeable results they did enable the program team to draw up a reasonable consensus of the strengths and weaknesses of the whole program. When further iterations become possible, the program team is confident that the concept and the results can be refined (for example, it will be necessary to be much more rigorous about the treatment of alternative explanations). Of more interest than the robustness of the results, in the context of the organisational and cultural changes referred to in the background section, are the following lessons:

- Contribution analysis is a way of breaking down and managing the continuum between high level outcomes that everybody is interested in and nobody can claim full credit for to outputs which are readily attributable. It is possible to focus on a series of high level outcomes, while also affirming that it is the responsibility of donors collectively, working with government, to determine the contribution of aid, and beyond that for individual donors to assess the contribution of their own program to the extent that is domestically necessary.
- For an agency geared to producing primarily output data, contribution analysis offers a relatively safe way of moving to say something defensible about higher level outcomes. In order for it not to be so safe as to be meaningless, it should take place in the context of a country strategy in which objectives are set at the highest level of outcome at which it is reasonable to link to aid activities.
- Providing a framework that moves management attention from activities up the results chain also encourages thinking back down it, in the sense of asking, for any given result, what else Australia might have done to contribute to it – hence the emerging value of non-financial inputs.

Other factors commonly encountered when moving to measure results at a country level, but not necessarily connected to the contribution analysis model, were also present in the experience. As



expected, the monitoring frameworks at activity level were severely tested and as part of the learning experience began to be better aligned to higher level outcomes (Kotvojs describes, for example, the process of weaving the contribution story into individual advisers' work plans). The process was resource intensive, but integrated into staff work programs in a way that enriched the perception particularly of junior staff of what it was possible for an aid program to aim at, and the strengths and weaknesses of the Australian approach. The experience demonstrated that a collective process of country program assessment involving AusAID staff, contractors, partner country counterpart staff and external informants can be useful in reinforcing higher level objectives and informing the positioning of programs without imposing unacceptable transaction costs on partner Government officials and key informants.

### *Conclusion*

Contribution analysis provides a practical means of applying inductive methods to a stretched logic chain. It is less obviously useful in cases where development aid provides a much larger proportion of the resources aimed at bringing about change. It is not a substitute for proper end of cycle evaluation. But it has proved useful to a set of program managers coming at routine country level performance measurement for the first time. For an agency moving away from the safety of output measures it offers a bridge to outcome reporting, which although it can be resource intensive, requires no special methodological expertise.

### *References*

AusAID: Australian Aid: Promoting Growth and Stability – A White Paper on the Australian Government's Overseas Aid Program, 2006a

AusAID: Annual Report 2005-2006, 2006b and previous reports in the series

Binnendijk, A: Results based management in the Development Co-operation Agencies: A Review of Experience, DAC Working party on Aid and Evaluation, 1999

Bourguignon, F and Sundberg, Aid Effectiveness: Opening the Black Box, paper presented to the American Economics Association, 2007

Flint, M: Easier said than done: A Review of Results Based Management in the Multilateral Development Institutions, for the UK Department for International Development, 2003

Ireland, M, McGregor, J and Saltmarshe, D: Challenges for Donor Agency Country Level Performance Assessment: A Review, in *Public Administration and Development*, 23, 2003

Iverson, A: Attribution and Aid Evaluation in International Development: a Literature Review, for IRDC, 2003

Kotvojs, F: Contribution Analysis: A new approach to evaluation in international development, paper presented to the Australasian Evaluation Society, 2006

Lawson, A, Gerster, R and Hoole, D: Learning from experience with performance assessment frameworks for General Budget Support, for Swiss State Secretariat for Economic Affairs, 2005

Mayne, J: Addressing Attribution through Contribution Analysis: Using Performance Measures Sensibly, discussion paper, 1999

OECD: The Paris Declaration on Aid Effectiveness, 2005

White, H and Booth, D: "Using development goals to design country strategies" in "Targeting Development", ed Black, D and White, H, 2003

## **Annex 1**

### **Principles of contribution analysis <sup>8</sup>**

#### **• Acknowledge the problem**

Too often, the measuring and particularly the reporting of performance through performance measurement systems completely ignores the attribution problem. A first step is simply acknowledging that there are other factors at play in addition to the program and that it is therefore usually not immediately clear what effect the program has had or is having in producing the outcome in question.

#### **• Present the logic of the program**

While logical frameworks are often used, there are limitations in terms of attribution when using these to describe a program of support rather than a project. The logical reasoning behind the program that explains what it is supposed to be accomplishing and how needs to be articulated.

#### **• Identify and document behavioural changes**

By trying to identify and then document the changes in attitudes, knowledge, perceptions and decisions taken by program target groups, which logically link to the outcomes being observed, a good understanding of the actual impact the program is having can often be acquired.

#### **• Use discriminating indicators**

Considerable care is needed in selecting indicators of performance, often maximizing the specificity of indicators gives a better indication of what the program is trying to achieve.

#### **• Track performance over time**

In cases where the program activities have varied over time, showing that outcomes have varied in a consistent manner with the variation in activities can strengthen the argument that the activities have indeed made a difference.

#### **• Discuss, and test alternative explanations**

Dealing with alternative explanations explicitly is often the best way of buttressing an argument in favour of the program's impact. This entails:

- identifying the most likely alternative explanations;

---

<sup>8</sup> Taken from Mayne, op cit

- presenting whatever evidence or argument you have to discuss and, where appropriate, discounting these alternative explanations; and
- presenting whatever evidence there is that the program is a more likely explanation.

- **Gather additional relevant evidence**

The stronger the case that can be made, the stronger is the conclusion about the program's contribution. Additional data collection might include a review of the relevant literature, surveys, tracking of relevant external factors, field visits, or focus groups. In addition, evidence about the contribution of the program can be gathered directly, most often through the use of expert opinion, through a structured survey or a focus group of experts, and/ or through review of program files, secondary analysis (studies that others have done in the program area) or case study evidence.

- **Gather multiple lines of evidence**

While no one piece of evidence may be very convincing, a larger set of different and complementary evidence can become quite convincing.

- **When required, defer to the need for an evaluation**

In some cases, if the various lines of evidence point in different directions there may be little one can say with enough credibility about the contribution of the program. In this case, the best strategy may be to simply acknowledge that one does not know and suggest that an evaluation be carried out.