Measuring performance or performing measurements? Measuring the impact and sustainability of the PARBICA Recordkeeping For Good Governance Toolkit

Danielle Wickman

Danielle Wickman works in performance audit at the Australian National Audit Office. Prior to this she spent 15 years at the National Archives of Australia working across many areas including governance, publishing, arrangement and description, reference and disposal. In 2006 she worked with the Regional Assistance Mission to Solomon Islands to develop recordkeeping guidelines for the Solomon Islands government. She returned to Australia in 2007 to manage the first phases of the PARBICA Recordkeeping for Good Governance Toolkit. Danielle has degrees in journalism, librarianship, archives and records management and museums and collections, and is a past Managing Editor and Councilor of the Australian Society of Archivists.

Recordkeeping professionals generally agree that recordkeeping is a shared responsibility throughout any organisation. It requires senior management support, competent administrators and compliant staff to ensure that full and accurate records of business are created and maintained. In contrast, performance measurement in an institutional setting normally requires organisations to report on their contribution to outcomes over which they have full control.

How can archivists and records managers demonstrate their impact on recordkeeping outcomes, when so much depends on other actors? How can recordkeeping initiatives be made sustainable when their use and implementation is largely handed over to others?

This paper uses the PARBICA Recordkeeping for Good Governance Toolkit as a case study to address issues of measuring the performance and sustainability of recordkeeping initiatives.

Recordkeeping professionals agree that recordkeeping is a shared responsibility. It requires management support, competent administrators and compliant staff to ensure that full and accurate records are created and maintained. In contrast, performance measurement in an institutional setting normally requires organisations to report on their contribution to outcomes over which they have full control. How can archivists and records managers demonstrate their impact on recordkeeping outcomes, when so much depends on others? How can recordkeeping initiatives be made sustainable when their use and implementation is handed over to others?

In reporting on the development of the Recordkeeping for Good Governance Toolkit, the National Archives of Australia and the Pacific Regional Branch of the International Council on Archives (PARBICA) have had to explain, like any other publicly-funded body, how their work has met the goals set for it. These things can be difficult to measure in any context. How can PARBICA, with its limited resources, reliably measure the impact of its work not just in one country, but across the Pacific?

The Recordkeeping for Good Governance Toolkit has been developed by PARBICA, with assistance from the National Archives of Australia and, more recently, Archives New Zealand. It is a growing set of tools that should enable the governments of Pacific Island countries to progressively improve their performance in a range of key recordkeeping areas. Phase one of the toolkit provided promotional and diagnostic tools, which sought to gain the attention of senior government officials and to help them to understand their governments' own recordkeeping needs. Phase two concentrated on the first of a range of practical tools for meeting recordkeeping standards. It provides a model recordkeeping policy and a suggested record plan or file-titling system for common administrative records. Phase three – being led by Archives New Zealand – provides suggested disposal coverage for administrative records, and guidance on appraisal for other records. Phase four – back with Australia – is a

consolidation phase and provides train-the-trainer skills for PARBICA members to assist them in 'selling' and implementing the toolkit.

Phases one, two and four of the PARBICA toolkit have been funded by AusAID through its Pacific Governance Support Program (PGSP).[1] The PGSP has now been replaced by a slightly different funding program, but during its existence it aimed to develop public sector expertise in Pacific Island countries, build institutional capacity and strengthen regional approaches to shared problems.[2]

AusAID requires a full acquittal of the money it provides under its grant programs, and also asks for interim and final reports. Of course, the application process requires agencies seeking funding to demonstrate their eligibility for the program and to outline what they plan to use the money for. The PGSP application form asks for a description of the proposed activity and what it is intended to achieve, along with an assessment of how this will contribute to broader goals such as those outlined in the Pacific Islands Forum's Pacific Plan, and AusAID's own policies and country strategies.

Successful applicants sign a funding agreement with AusAID, which outlines the objective, outcome and outputs of the activity. At the end of the activity the final report asks for an assessment of achievements against the agreed objective, outcomes and outputs, and of the sustainability of the project's impacts.

The National Archives of Australia's application for phase two funding addressed these requirements in the following ways. The objective of the project was:

> to develop an easily understood file-titling system, or record plan, for common administrative files that can be adapted and used by governments around the Pacific.

The outcome envisaged was:

> to help government agencies to improve their recordkeeping and thus operate more efficiently and accountably.

Finally, the project's outputs were:

- a model recordkeeping policy;
- a record plan for common administrative functions; and
- skills development for Pacific Island archivists.[3]

In making a final report on the phase two of the project to AusAID, it was fairly easy to demonstrate that the National Archives of Australia (NAA) and PARBICA had met the objective and produced the outputs required by the funding agreement. Although some aspects of these measures are somewhat subjective, such as the ease of use of the file-titling system, and the degree of skills development achieved for the Pacific Island archivists, generally speaking it is not too hard to demonstrate that the project met these requirements.

In assessing what impact the project had on the agreed outcome – to improve government recordkeeping and through it efficiency and accountability – the NAA pointed to a number of developments in participating Pacific Island countries which demonstrate that the toolkit is beginning to be implemented:

* phase one products have been translated into French by the French Association of Archivists, and into Samoan by the Samoan Ministry of Education;

* products from both phases are being used as the basis for a records management training program being developed by the Public Service Commission in Papua New Guinea (PNG);

* the National Archives of Fiji is in discussions with its Public Service Commission to develop an initiative similar to the one in PNG, and to adopt the model recordkeeping policy;

* PARBICA was invited by the United Nations Development Programme (UNDP) to speak about the toolkit at its recent freedom of information workshops in Nauru and the Solomon Islands, and recordkeeping has been emphasised in the outcomes statements of both events;

* PARBICA members in Nauru and Palau successfully had their governments raise recordkeeping as a priority issue at the Pacific Islands Forum meeting in August 2009; and

* the National Archivist of Palau is seeking formal government endorsement and adoption of the Model Recordkeeping Policy and other PARBICA tools. Work has already commenced to tailor the model record plan for local needs.[4]

All of these achievements also help to demonstrate the sustainability of the work begun by PARBICA with the toolkit. PARBICA members are proving that the toolkit is something they can work with, and is relevant to their recordkeeping needs. The NAA and PARBICA also noted, in the discussion of sustainability, that support from senior people within government is crucial. In all of these examples of work being done with the toolkit, the PARBICA member who is behind these initiatives is either influential within government in his or her own right, or has been successful in attracting influential supporters. In the case of the UNDP project, the invitation to speak came to PARBICA after the Pacific Islands Forum secretariat brought the toolkit to the UNDP's attention. In countries where PARBICA members have not been getting sufficient support from government leaders we see significantly less evidence of progress.

All of these achievements are commendable, but are they really reliable measures of impact and sustainability? Do they really demonstrate that, as a direct result of the PGSP-funded work, Pacific Island governments are now able to improve their recordkeeping and thus operate more efficiently and accountably, which our outcomes statement claims we will enable? The answer at this stage really must be 'no'. So, how *can* we measure the performance of the toolkit, and what is it that we really want to measure anyway?

> When God had finished the creation of the world, on the sixth day of Creation, he looked at it, and he 'saw every thing he had made, and, behold, it was very good' (Gen. 1: 31). Having heard him say so, the archangel Lucifer came and asked God: 'How can you judge it? What are your criteria?' The following day, on the seventh day of creation, God assembled the archangels in heaven and he sent Lucifer to hell, punishing him for his nasty questions he was not able to answer.[5]

What are we measuring here? What are our criteria? The toolkit itself is an output, and it would be appropriate to make some kind of assessment of the quality of that output. In 2007, a team of British researchers proposed a set of criteria for assessing records management capacity and compliance toolkits, and compared four

such tools. The researchers suggested nine areas of evaluation:

- provenance, including the producer and its 'track record';
- audience, including applicability to a wide range of sectors, organisations and staff categories;
- coverage of the full records continuum;
- basis in legislation, standards or best practice;
- format and the processes required to use it;
- resource requirements for using the toolkit;
- accessibility and compatibility issues;
- usability, in terms of clarity, support materials and ease of understanding, as well as software-related issues; and
- evaluation approach, which considers customisation, flexibility, and depth of analysis.[6]

These kinds of measures are likely to influence the extent to which a toolkit might be taken up and successfully used by an organisation. If the toolkit is not developed by a respected organisation, if it has narrow applicability, requires specialised software or other significant resources, or is unclear or inflexible, users are unlikely to attempt to use it, and even less likely to use it successfully.

These assessments could also help us with evaluating the sustainability of the PARBICA toolkit. AusAID's guidance on promoting sustainability focuses on the key factors of:

- the policy environment in donor and partner countries;
- participation by stakeholders in the development project;
- administrative and personnel capacity;
- financial impacts, particularly ongoing costs;
- information and training issues;
- use of appropriate technology; and
- social, cultural, gender, environmental and political issues.[7]

Without going into a detailed analysis, there does appear to be some overlap between the criteria developed for assessing recordkeeping

toolkits and these key factors for development sustainability. In particular, both emphasise the need for taking account of the expertise of recipient organisations, and the ongoing financial, personnel and technological issues which may influence the toolkit's use. We could probably quite favourably match the toolkit achievements reported to AusAID against these factors.

While these kinds of criteria can tell us if we are on the right track in making our toolkit attractive to people who might want to use it, they do not tell us very much about what kind of success they will have in using it. They measure the quality of the toolkit as an output, and its sustainability as a product. They do not give us very much to go on in assessing the extent to which that output might contribute to the outcome of improved government recordkeeping.

The toolkit does have its own internal set of measures for assessing recordkeeping performance, in the form of the recordkeeping capacity checklist. The checklist consists of ten questions designed to give organisations an idea of the kinds of things they need to have in place to manage their records. It asks questions about senior management support, the existence of recordkeeping policies and procedures, assignment of responsibilities, resourcing, training, storage and access issues. Incidentally, question ten asks whether or not organisations 'set performance indicators for [their] records management unit, report on those indicators, and make improvements when indicators are not met'.[8]

The questions are rooted in the international records management standard. Unsurprisingly, they are similar to other ISO 15489-derived diagnostic tools, such as the National Archives of Australia's *Check-Up*.[9] The PARBICA checklist is, however, a very slimmed-down version of tools like *Check-up*, designed primarily as a prompt to action for decision-makers and resource allocators, rather than as a comprehensive evaluation tool.

There is no shortage of tools or agreed standards through which we can measure agency recordkeeping and its quality. Such measures, benchmarked over a period, would certainly begin to answer our question of whether or not governments have improved their recordkeeping. They do not, however, measure the impact of that recordkeeping on government efficiency and accountability.

Such an evaluation is not generally onerous when each agency is performing a self-assessment. The costs to and logistics for a national archives in assessing agencies within its government are much greater, although State Records NSW, for one, sees the coordination of self-assessment surveys as its 'primary means of monitoring recordkeeping'.[10] PARBICA is, however, a step further removed from the agencies it seeks to evaluate, and coordinating a self-assessment survey across the Pacific would be no mean feat. As Larmour notes in his assessment of anti-corruption measures in the Pacific, international initiatives are a two-step process:

> First, the international organization had to persuade the island government to adopt the new policy. Second, the island government had to implement it. It might fail the second step for all sorts of reasons, such as lack of political commitment, resistance among those affected, under-resourcing or because it was a bad idea in the first place.[11]

The same is true of PARBICA's efforts, both in implementing the toolkit and in finding ways of measuring that implementation.

The cost of measurement should also be a factor in developing performance frameworks. Not only monetary costs, but also compliance burdens for those supplying the information need to be considered. 'The cost of producing performance information should be balanced against the use of the information and how it will improve performance.'[12]

Would a whole-of-Pacific survey against the checklist demonstrate achievement of our outcome? Benchmarking, or assessing a product against an agreed standard, like the international standard which the checklist is derived from, is an accepted performance measure.[13] We have already seen that such benchmarking is usefully employed in a range of jurisdictions, and there is no shortage of tools to guide us in doing it. Are we not, though, merely measuring another organisation's inputs, rather than the real outcome to which we are hoping to contribute? Our impact on our real goal – efficient and accountable government – is still elusive.

Measurement of outcomes has been the practice in the Australian Government for close to ten years,[14] and is intended to ensure that

outputs are 'better aligned with desired policy outcomes, [and] to overcome the inherent problem whereby output production tends to become an end in itself.'[15] I imagine that not many governments really care about whether or not they have improved recordkeeping. They may, though, care about a policy outcome of increased efficiency and accountability which good recordkeeping enables.

Measuring outcomes rather than outputs allows government to make assessments about the worth of the products they are commissioning. For example, in seeking PGSP funding from AusAID, PARBICA was competing against many other important projects, such as helping meteorology organisations to issue tsunami warnings and assessing staffing levels for nurses in the Tongan Ministry of Health.[16] It is much more meaningful for AusAID to choose between objectives such as improved accountability, earlier tsunami warnings and more efficient nursing services than to compare the worth of a recordkeeping policy, a tsunami warning communication plan and a nursing roster.

The British Standards Institute has published a part three to its version of the International Standard on Records Management ISO 15489, which deals explicitly with performance management. The fact that, at the time of writing, it appears that not a single copy of this standard exists in an Australian library may say something about our focus to date on performance measurement.[17] The introductory material to part three of the British Standard emphasises the need for effectiveness measures: 'it is not enough to simply be performing well – there is an imperative to have demonstrable evidence that the program adds value to the organization'.[18]

In her assessment of the impact of the changed freedom of information (FOI) legislation in the UK, Elizabeth Shepherd noted that some agencies assessed a failure to comply with FOI as an acceptable risk, and quotes Ennion, who warned that 'it is always necessary to make a business case for spending on any improvements, simply saying "it's the law" is not enough'.[19] For recordkeeping, our business case is the outcome we have described for our work – its impact on efficiency and accountability.

One consensus of the performance measurement literature seems to be that measuring effectiveness against outcomes is difficult to do. Bob

McLean reminds us that '[t]he tendency is to focus on what is easy to measure [such as costs and efficiency] and provides a very introspective view.'[20] Graf notes that program evaluation 'is a time-, cost- and energy-consuming task,' which 'should not be done for its own sake.'[21] Lonti and Gregory warn that measuring outcomes raises 'questions about the feasibility of demonstrating outcome success and failure in ways that can be fairly sheeted home to managers'.[22] The Australian National Audit Office's (ANAO) review of the use of outcome measures in the Australian Government found that 'many indicators were impacted by factors outside the control of the agency including, for example, general economic conditions, or the activities of other levels of government or industry. This reduced agencies' ability to determine and report on their own performance'.[23] However, the UK government warns that it is particularly important to measure effectiveness – the relationship between outputs and outcomes – when there are many outside influences on outcomes.[24]

It is these last points that seem particularly relevant to recordkeeping initiatives. Even if we had the resources to measure over time the changes in recordkeeping performance of government agencies, how much of this performance can we really take responsibility for? AusAID's advice on sustainability factors gives us a clue as to the other issues that may be influencing recordkeeping performance, such as the political environment and cultural issues. Recordkeeping professionals know for themselves which stakeholders might influence their success. Is it PARBICA's fault if records managers in the Pacific cannot get access to any training? Can PARBICA really take all the credit for improved accountability if the Auditor-General's office has recently been reconstituted and is demanding better recordkeeping?

The Australian Auditor-General's guidance on measuring outcomes acknowledges these difficulties, and advises using intermediate outcomes,[25] or explaining contributions to shared outcomes.[26] In these scenarios, outcomes are defined for a shorter time period than the ultimately desired objective, or the organisation acknowledges the contribution of other players to a shared outcome. Both of these techniques would seem to be relevant to the recordkeeping field, where outcomes are often long-term, and stakeholder interactions complex.

Determining a feasible outcome is, however, only part of the solution. We should ideally be able to measure our contribution to the achievement of this outcome. The Australian Auditor-General acknowledges the difficulty of this task, and suggests a need for 'approximate indicators'.[27] The Australian Tax Office (ATO), which we might expect to have a much better chance of measuring its effectiveness against outcomes than PARBICA, says '[e]vidence of effectiveness should be defensible rather than definitive' because 'it is difficult to reach a conclusive statement about the causal relationships between compliance activities and the results produced.'[28]

In fact, coming up with meaningful measures of effectiveness against outcomes is apparently so difficult that the Australian Government no longer requires agencies to commit to specific effectiveness indicators for their outcomes. Instead, agencies give a broad statement of how their outputs, for which there are detailed performance indicators, contribute to the delivery of the stated outcome.[29]

Agencies have, nevertheless, made assessments in the past about how effective they have been in achieving the outcomes government sets for them. The NAA's designated outcome for 2007–08 was to ensure that:

> Commonwealth institutions have access to recordkeeping products and services that enable them to account for their actions and decisions through full, authentic and reliable records.

The NAA gauged its effectiveness in achieving this outcome by making an assessment of the extent to which 'authentic and reliable records are produced by Commonwealth institutions'. It measured this through assessing trends over time in reports about agency recordkeeping by the ANAO and the Australian Public Service Commission (APSC).[30] The ATO – another agency with a job of getting people to do important stuff they don't want to do – reported on a range of effectiveness measures, such as the extent to which it maintains community confidence in the taxation system. This indicator was assessed through surveys of taxpayer satisfaction with the ATO's professionalism, analysis of complaints data and technical and corporate benchmarking.[31]

For both of these agencies, the measures are what the ATO might call defensible rather than definitive. The NAA's use of another agency's data (both the ANAO and the APSC) would tend to indicate that we are looking at a shared outcome between these agencies and the NAA. In the ATO's case, it is not community confidence itself that is being measured, but the community's perception of the ATO's professional capabilities. These are approximate indicators which would tend to suggest that the ATO is achieving its outcome without measuring that outcome definitively.

Where does this leave PARBICA, the toolkit, and the rest of us who sometimes feel the need to justify our existence? We can and should be measuring the quality of outputs such as the toolkit against relevant standards such as ISO 15489. As a profession, we set ourselves these standards in order to be measured against them, to give – as far as possible – an objective account of the quality of our work.

If we wish, however, to demonstrate more conclusively our relevance to government and society, we need to do more than measure our products against the expectations of our peers. We need to make some sort of statement of the outcome we are trying to achieve. PARBICA did not develop the toolkit – whatever its quality – for its own sake, but to help its members to influence good recordkeeping in each of their countries. The extent to which PARBICA can measure our impact may be limited, but if we are to demonstrate our worth alongside other aid priorities, the attempt needs to be made.

There are many purposes of performance evaluation, measurement and reporting. Sometimes we just do it because we must. We are better served, however, if we do it both to learn from our past record, and to demonstrate the meaning of that record to others. If we can find reliable and convincing ways of demonstrating the impact our work has on wider government and social outcomes, we have a much better chance of convincing others of the value of our work. Given the time, effort and cost that goes into performing measurements, we really should be making sure that we are actually measuring performance.

Endnotes

1   Phase three of the toolkit has been funded by NZAid.

2   See the AusAID Pacific Governance Support Program webpage, available at *<http://www.ausaid.gov.au/business/other_opps/pgsp.cfm>*, accessed 26 August 2009.

3   National Archives of Australia, *Funding agreement between AusAID and the National Archives of Australia*, 18 July 2007, internal document, p. 2.

4   National Archives of Australia, *Activity Completion Report*, 22 July 2009, internal document, pp. 1–2.

5   Christoph Graf, 'Program Evaluation in Archives: an essay on a research agenda', *Archival Science*, vol. 4, nos 1–2, 2004, p. 127.

6   Julie McLeod, Sue Childs and Susan Heaford, 'Records Management Capacity and Compliance Toolkits: a critical assessment', *Records Management Journal*, vol. 17, no. 3, pp. 20–23.

7   Australian Agency for International Development, *Promoting Practical Sustainability*, Australian Agency for International Development, Canberra, 2000.

8   Pacific Regional Branch of the International Council on Archives, *Guideline 1: Recordkeeping Capacity Checklist*, available at *<http://www.parbica.org/Toolkit%20pages/Documents/Checklist/Checklist%20html.htm>*, accessed 31 August 2009.

9   National Archives of Australia, *Check-up: A tool for assessing your agency's information and records management*, available at *<http://www.naa.gov.au/records-management/check-up/index.aspx>*, accessed 31 August 2009.

10   State Records Authority of New South Wales, *Monitoring Tools*, available at *<http://www.records.nsw.gov.au/recordkeeping/topics/monitoring-records-management/monitoring-tools>*, accessed 23 September 2009.

11   Peter Larmour, *Evaluating International Action Against Corruption in the Pacific Islands*, State Society and Governance in Melanesia Discussion Paper 2007/1, Australian National University Research School of Pacific and Asian Studies, Canberra, 2007, p. 9.

12   HM Treasury et al, *Choosing the Right FABRIC: A Framework for performance information*, HM Treasury, London, 2001, p. 16.

13   Australian National Audit Office, *Better Practice Guide: Better Practice in Annual Performance Reporting*, Australian National Audit Office, Canberra, 2004, p. 50.

14   Senator Andrew Murray, *Review of Operation Sunlight: Overhauling budget transparency*, June 2008, p. 85, available at *<http://www.finance.gov.au/financial-framework/financial-management-policy-guidance/operation-sunlight/docs/budget-transparency-report.pdf>*, accessed 31 August 2009.

15   Zsuzsanna Lonti and Robert Gregory, 'Accountability or Countability? Performance Measurement in the New Zealand Public Service 1992-2002', *The Australian Journal of Public Administration*, vol. 66, no. 4, p. 468.

16   Australian Agency for International Development, *Annual Report 2007–2008*, p. 58.

17   References to part 3 of the standard are made via sample pages available through *Google Books*, available at *<http://books.google.com.au/books?id=vQOorgNwO7YC&printsec=frontcover&source=gbs_v2_summary_r&cad=0#v=onepage&q=&f=false>*, accessed 23 September 2009.

[18]   Philip Jones, *Effective Records Management: Part 3: Performance management for BS ISO 15489-1*, BSI Standards, London, 2003, p. 4.

[19]   E Ennion, 'How Has the Implementation of the Freedom of Information Act 2000 Affected Archive and Records Management Services in Public Bodies in the UK?', MA dissertation, University of London, 2005, p. 44, quoted in Elizabeth Shepherd, 'Freedom of Information and Records Management in the UK: What has been the impact?', *Journal of the Society of Archivists*, vol. 28, no. 2, October 2007, p. 128.

[20]   Bob McLean, 'Effective Records Management Part 3: Performance Management for BS ISO 15489', *Records Management Journal*, vol 15, no. 1, 2005, p. 59.

[21]   Christoph Graf, 'Program Evaluation and Archives: An Essay on a research agenda', *Archival Science*, vol. 4, nos 1–2, 2004, p. 133.

[22]   Lonti and Gregory, 'Accountability or Countability?', p. 471.

[23]   Australian National Audit Office, *Application of the Outcomes and Outputs Framework*, Australian National Audit Office, Canberra, 2007, p. 25.

[24]   HM Treasury, *Choosing the Right FABRIC*, p. 9.

[25]   Australian National Audit Office, *Better Practice Guide: Better Practice in Annual Performance Reporting*, Australian National Audit Office, Canberra, 2004, p. 11.

[26]   ibid., p. 10.

[27]   ibid., p. 13.

[28]   Australian Taxation Office, *Measuring Compliance Effectiveness – our methodology*, Australian Taxation Office, Canberra, August 2008, p. 6.

[29]   Compare, for instance, the National Archives of Australia's *Portfolio Budget Statements 2008–09* (available at <http://www.pmc.gov.au/accountability/budget/2008-09/>) with those for 2007–2008 (available at <http://www.archive.dbcde.gov.au/2008/05/federal_government_budget_2007-08_department_of_communications_information_technology_and_the_arts>), accessed 16 November 2009.

[30]   National Archives of Australia, *Portfolio Budget Statements 2007–08*, p. 8, available at <http://www.archive.dbcde.gov.au/2008/05/federal_government_budget_2007-08_department_of_communications_information_technology_and_the_arts>, accessed 31 August 2009.

[31]   Australian Taxation Office, Annual Report 2007–2008, Australian Taxation Office, Canberra, 2008, p. 22.