**Looking for Lazarus:
Rust or Renewal in the Australian Public Service**

***A submission to the Independent Review of the
Australian Public Service, 2018.***

**Angela O’Brien-Malone**

[**Mark R. Diamond**](#MRD)

Suggested citation: O’Brien-Malone, A., & Diamond, M. R. (2018). *Looking for Lazarus: Rust v Renewal in the Australian Public Service*. Submission to the Independent Review of the Australian Public Service. DOI: [10.6084/m9.figshare.6934634](https://doi.org/10.6084/m9.figshare.6934634)

# Background

We are both social scientists who have been employed within the Australian Public Service (APS). We each have a long record of employment as academics in Australian universities. Our research expertise lies in program evaluation and related matters within the APS.

Our submission is informed by our experience as experts within the APS, and particularly by our experience of being evidence‐makers and evidence‐assessors within the APS. We have made contributions to reviews that touch upon these matters.

Although the current *Independent Review of the Australian Public Service* is important, it is not the first of its kind. Several previous reviews asked questions similar to those posed by the current review, but their reports have faded into obscurity. If the current review is to be drive meaningful change, it must grapple with some of the issues that were raised in the previous reviews of the APS but were not successfully addressed. Unsurprisingly, the current Terms of Reference include mention of some of those issues.

# Structure of the submission

In this paper, we build on previous submissions we have made regarding matters that are highlighted in the current *Independent Review*. We have taken this approach because the past is, unfortunately, still with us. Despite previous reviews raising the *possibility* of improvements in the way the APS:

(a) deals with data and evidence,

(b) maintains key competencies within its staff, and

(c) ensures the transparent and effective use of taxpayers’ money,

there has been little progress towards overcoming several entrenched problems. Little is needed in the way of *new* argument regarding the problems. What is needed is determination, and change.

In the following sections, we make a large number of recommendations, without locally embedded argument or support, but with extensive forward-reference to the detailed argument and references in four other papers that we reproduce here as:

[Appendix 1](#_Appendix_1:_On): *On rational advocacy*

[Appendix 2](#_Appendix_2:_Evaluation—A): *Evaluation—A neglected but vital reform*

[Appendix 3](#_Appendix_3:_Research): *Research to government and back again*

[Appendix 4](#_Appendix_4:_Reform): *Reform of non-financial reporting.*

The recommendations, which cover the issues listed as (a), (b) and (c) above, are covered under two broad headings: “Acquiring, developing and retaining key competencies to ensure that the APS can discharge its responsibilities”, and “Ensuring the transparent and effective use of taxpayers’ money.”

# Acquiring, developing and retaining key competencies to ensure that the APS can discharge its responsibilities

**Recommendation**: That the APS be encouraged to establish and use career paths based on expertise. (“The APS must develop a research culture and value ‘in-house’ expertise”, page 5 of Appendix 1; “Becoming an APS generalist is not conducive to developing domain-specific expertise”, page 4 of Appendix 1; “Creating career paths based on expertise”, page 4 of Appendix 3).

**Recommendation**: That the careers of experts within the APS be structured so as to facilitate both the maintenance of expertise and the maintenance of a public claim to expertise. (“Creating career paths based on expertise”, page 4 of Appendix 3).

**Recommendation**: That APS agencies publish, and encourage staff to develop and contribute to, individually authored (not agency authored) research-based discussion papers modelled on the Research Discussion Papers of the Reserve Bank of Australia. (“The APS must develop a career path that rewards expertise”, page 6 of Appendix 1; “Creating career paths based on expertise”, page 4 of Appendix 3; “Comment”, page 5 of Appendix 3).

**Recommendation**: That candidates for higher research degrees be encouraged to consider undertaking their research with the context of the APS; and that the APS be encouraged to house and support such undertakings. (“Attracting research expertise to the APS”, page 4 of Appendix 3).

**Recommendation**: That APS agencies work to improve the exchange of ideas *and people* between the APS, universities and other research institutions. (“Other approaches to gaining in-house expertise”, page 7 of Appendix 1).

# Ensuring the transparent and effective use of taxpayers’ money

**Recommendation**: That the evaluation of programs be given the same status as audit of programs. Consideration might be given to amending the *Public Governance, Performance and Accountability Act* *2013* to give evaluation functions the same status within Agencies as audit functions. (“Money spent ineffectively is money wasted”, page 1 of Appendix 2).

**Recommendation:** That program evaluations be designed by professional evaluation experts not by generalist staff. Evaluation functions with each Agency should be centralized to maximize the use of limited evaluation expertise. (“Evaluation requires expertise”, page 2 of Appendix 2).

**Recommendation**: That evaluation functions within each Agency be overseen by an SES member with appropriate professional qualifications in evaluation. The SES member should have status equivalent to that of the Agency auditor. (“Evaluation requires expertise”, page 2 of Appendix 2; “What maintains poor evaluation?”, page 8 of Appendix 4).

**Recommendation:** That evaluation consultancies be overseen by centralized evaluation expertise within each Agency, with a view to assuring the quality, honesty and integrity of program evaluations. (“ Ensuring honesty and independence in evaluation”, page 2 of Appendix 2).

**Recommendation:** That the SES officer responsible for the centralized evaluation function within an Agency also be responsible for maintaining the independence and integrity of evaluations. (“ Ensuring honesty and independence in evaluation”, page 2 of Appendix 2; “What maintains poor evaluation?”, page 8 of Appendix 4).

**Recommendation:** That the assessment of tenders for evaluation consultancies be free of undue influence by program staff. (“ Ensuring honesty and independence in evaluation”, page 2 of Appendix 2).

**Recommendation:** That evaluation reports be delivered to an independent officer within each Agency such as the SES officer responsible for the centralized evaluation function. (“ Ensuring honesty and independence in evaluation”, page 2 of Appendix 2; “What maintains poor evaluation?”, page 8 of Appendix 4.

**Recommendation**: That Agencies be required to table a list of evaluation reports and evaluation consultancies with the Parliament. (“Most evaluations are buried”, page 3 of Appendix 2; “Design of the performance evaluation process”, page 11 of Appendix 4).

**Recommendation**: That Senate Estimates Committees be encouraged to seek advice from evaluation experts in framing questions to Agency heads regarding claims of program effectiveness. (“Most evaluations are buried”, page 3 of Appendix 2).

**Recommendation**: That the reporting structure for non-financial outcomes under the *Public Governance Performance and Accountability Act 2013* be tied to activities that have causal agency. (“Assessment must be tied to the agents of change”, page 3 of Appendix 4; “Non-financial reporting: What does government need to know?”, page 5 of Appendix 4).

**Recommendation:** That the Department of Finance require all proposals for new policy initiatives to describe explicitly the assumed causal chain between the activities that are to be funded and the anticipated outcomes. (“What should be measured and why? The role of ‘program’ logics”, page 5 of Appendix 4).

**Recommendation:** That prior to implementation, program logic be used to decide what variables will index short-, medium- and long-term effectiveness of the activity. That funding be tied to the requirement that the proposed indices actually be measurable. (“What should be measured and why? The role of ‘program’ logics”, page 5 of Appendix 4).

…oooOOOooo…

# Appendix 1: On Rational Advocacy

Paper submitted 30 November 2009 to the Review of Australian Government Administration

**On Rational Advocacy**

A response to the discussion paper entitled

*Reform of Australian Government Administration: Building the world’s best public service*

**Angela O’Brien-Malone *BSc, BA(Hons), PhD***

**and**

**Mark R. Diamond *MPsych, PhD***

# Background of the Authors

We are both social scientists who are currently employed within the APS. We each have a previous long record of employment as academics in Australian universities. Our research expertise has largely focussed on the use of various methodologies in the social sciences, particularly the methodologies which are used within the disciplines of psychology and education. Since these methods also form the foundation for much of program evaluation, our research expertise is also the basis for our roles working on program evaluation and related matters within the APS. This submission is informed by our experience as experts within the APS, and particularly by our experience of being evidence-makers and evidence-assessors within the APS.

# Abstract

In this submission we argue that significant in-house expertise is required by the APS if it is to fulfill its responsibilities as a professional and rational advocate of ideas that are in the best long-term interests of Australia. Currently, the APS has limited capacity to evaluate essential evidence from the social sciences and it relies largely on non-experts to do so. If the APS is to recruit, retain, and effectively utilize in-house expertise then it must: develop a research culture, value ‘in-house’ expertise, and actively engage in helping in-house experts to maintain their expertise. It must also develop a career path that rewards expertise. We suggest strategies by which the APS might achieve these important goals.

# Table of contents

[Background of the Authors ii](#_Toc247358573)

[Abstract iii](#_Toc247358574)

[Table of contents iv](#_Toc247358575)

[1 Introduction 1](#_Toc247358576)

[2 Rational Advocacy 1](#_Toc247358577)

[2.1 Is expertise really necessary? 1](#_Toc247358578)

[2.2 The APS has limited capacity to evaluate essential evidence from
the social sciences 3](#_Toc247358579)

[2.3 The APS needs in-house expertise to evaluate evidence 4](#_Toc247358580)

[2.4 Becoming an APS generalist is not conducive to developing domain-
specific expertise 4](#_Toc247358581)

[2.5 The APS must develop a research culture and value ‘in-house’
expertise 5](#_Toc247358582)

[2.6 The APS must actively engage in helping in-house experts to maintain
their expertise 6](#_Toc247358583)

[2.7 The APS must develop a career path that rewards expertise 6](#_Toc247358584)

[2.8 Other approaches to gaining in-house expertise 7](#_Toc247358585)

[3. Summary 8](#_Toc247358586)

[4. References 9](#_Toc247358587)

# 1 Introduction

The *Reform of Australian Government Administration* (ROAGA) discussion paper describes a fundamental purpose of the APS as being to serve “the government of the day, including by striving *to be a professional and rational advocate of ideas* that are in the best long-term interests of Australia” (p. 1, emphasis added).

What does rational advocacy, in the context of evidence-based policy, actually mean? What is the basis of rational advocacy? What are the prerequisites for it to occur within the APS? What consequences are likely to ensue for the Australian Government and for the Australian people if rational advocacy in the APS fails? It is to these questions that this submission is directed.

Rational advocacy has two aspects: first, domain-relevant knowledge and expertise; second, skill in the formulation of a rational argument. Without the former, *any* argument will fail for want of a solid foundation. Without the latter, it will be impossible to present a defensible case for a particular position. Without a combination of expertise and argument, it will be impossible to develop credible, defensible, evidence-based policy.

Since knowledge and expertise play a central role in rational advocacy and the development of evidence-based policy, they form the subject of this submission. In particular, we address:

* the necessity for expertise
* current capacity within the APS to judge the quality of evidence
* barriers to collaboration between academics and other experts and the APS.

We note that the KPMG report[[1]](#footnote-1) entitled *Benchmarking Australian Government Administration Performance* uses the UK Civil Service as a comparator for the APS. We do likewise in this submission. We also note that the KPMG report concluded that the APS performs “comparatively poorly” in:

* *its capability for coordinated, informed and strategic policy;*
* *its tools, methods and institutions for integrating external expertise...* (p. 2)

This, broadly, is also our conclusion, but our discussion of the problems and possible solutions goes beyond that contained in the KPMG report.

# 2 Rational Advocacy

## 2.1 Is expertise really necessary?

During the past three decades, the UK has experienced several crises that have highlighted the need that the civil services of all developed countries have for

* the capacity to evaluate knowledge from the sciences and the social sciences
* the capacity to understand the nature of scientific expertise.

At least two of those crises—namely the emergence of Bovine Spongiform Encephalopathy (BSE) in the 1980s and the outbreak of Foot and Mouth Disease in pigs, sheep and cattle in 2001—led to the establishment of formal Inquiries. The reports that eventually issued from those Inquiries included the conclusions that Government Departments must:

*… retain ‘in house’ sufficient scientific expertise to enable them to understand and review advice given by advisory committees.*

*… review advice given by advisory committees to ensure that the reasons for it are understood and appear to be sound.* ([Inquiry into BSE and variant CJD in the United Kingdom](#UKBSEI), 2000, Vol 1, Chapter 14, p 256)

[b]ase policy decisions on best available science *and ensure that the processes for providing scientific advice are widely understood and trusted.* ([Foot and Mouth Disease 2001: Lessons to be Learned Inquiry](#UKFAMD), 2002, p. 11, emphasis in original)

In November 2005 a broad inquiry into the UK Government’s handling of scientific advice, risk, and evidence in policy making was launched by the House of Commons Science and Technology Committee. The inquiry addressed the social, physical and natural sciences and described the situation in the UK in 2005 in these terms:

*Many of the most high profile policy issues are critically dependent on the input of scientists. These include: securing the economic development of the UK through the knowledge economy; protecting the population of the country against an avian influenza pandemic and other infectious diseases; mitigating and adapting to climate change; safeguarding the UK’s energy supply; detecting and averting potential terrorist threats; and tackling obesity. In each case, effective policy development requires both an effective scientific advisory system and appropriate use of evidence and advice on science and risk by Government. This Government has repeatedly stated its commitment to pursuing an evidence based approach to policy making and placed great emphasis on the role of science in informing policy.* ([Scientific Advice, Risk and Evidence Based Policy Making](#UKHCSTC), paragraph 2)

The comments also represent a fair summary of the current position in Australia—we are critically dependent on the input of scientists and experts from many disciplines to meet the Australian Government’s objective of developing and using evidence to address significant policy issues.

What was the outcome of the 2005 inquiry by the House of Commons Science and Technology Committee? Detailed recommendations from the report which address, for instance, structural issues, might not translate straight forwardly into Australian circumstances but the tenor of comments in the report is nonetheless revealing:

*Experts in the civil service (and from outside) need to work closely with civil service generalists if scientific advice and evidence are to be effectively incorporated into policy. William Solesbury, Senior Visiting Research Fellow at the Centre for Evidence Based Policy and Practice, told us that it was in this area—the competence of the generalist staff—that the civil service’s weakness lay: “I do not think there is, as yet, very much, or at least not a very sophisticated understanding of the occasions when evidence is useful, the sort of evidence to be obtained, how to evaluate evidence when it is available, how to interpret it, and how to weigh it.”* ([Scientific Advice, Risk and Evidence Based Policy Making](#UKHCSTC), paragraph 48)

So, the conclusion of the House of Commons Science and Technology Committee, together with the earlier inquiries in the UK, was that knowledge and expertise from the social sciences was essential for evidence-based policy development. They further concluded that the UK Civil Service had significant problems with regard to understanding or evaluating evidence, and valuing or incorporating expertise within the Service.

Is the APS better-positioned than the UK Civil Service? No. Australia has simply benefited from luck in that no similar crisis has yet happened here. The APS currently suffers from many of the same weaknesses as those that placed the national good of the UK at such grave risk during those periods of crisis, and it would behove the APS to learn from the UK experience.

## 2.2 The APS has limited capacity to evaluate essential evidence from the social sciences

Folk explanations abound in the social sciences. Ask anyone in the street how to prevent domestic violence, reduce the crime rate, improve literacy, stop truancy, reduce traffic accidents, decrease unemployment, prevent child-abuse, or put an end to problem gambling, and you are sure to have found someone with “the answer”. People are willing to tout their solution to social policy problems in the absence of any serious evidence,[[2]](#footnote-2) yet those same people would be most unwilling to repair their microwave oven without employing an expert or seeking expert advice—even though social phenomena are considerably more complex than a microwave oven.

Unfortunately, the same mindset is found within the APS. The call for evidence-based policy has resulted in the proliferation of what is best termed “advocacy research”.[[3]](#footnote-3) A viewpoint about how to solve a particular policy problem seems “obviously” reasonable,[[4]](#footnote-4) and evidence is then accumulated to provide a case for supporting that solution.[[5]](#footnote-5) In its most blatant forms, it is the activity of advocacy research that has come to be facetiously referred to as “policy based evidence”, but in most instances, advocacy research is simply a consequence of (a) failure to recognize the immense complexity of human behaviour, (b) failure to understand that, despite its complexity, human behaviour and societal behaviour is open to empirical inquiry, (c) failure to understand that evidence and opinion are not the same and that appropriate empirical inquiry leads to evidence rather than opinion, and (d) a lack of expertise in evaluating evidence.

*An evidence-based approach requires a policy-making process that is receptive to evidence; a process that begins with a question rather than an answer, and that has institutions to support such inquiry.… The joke about ‘policy-based evidence’ has not been made in abstract — we have long observed such an approach in operation through the lens of regulation-making in Australia.* ([Banks](#Banks), 2009, p 18)

Unfortunately, the same causes as have led to the proliferation of policy solutions based on “obviousness”, also lead to problems with evaluating evidence. If you are not well-versed in the social sciences then you might believe that evaluating evidence is easy. Just read what’s on the web, listen to what “practitioners on the ground” are saying, and draw a conclusion. Would that it were so simple—but it is not. Evaluating evidence in the social sciences is difficult, as is reflected in the suggestion that that the soft sciences be renamed the *hard* sciences.[[6]](#footnote-6) But, despite the difficulty of evaluating social science evidence, the capacity to do so is vital to the proper functioning of the APS and to the realisation of the government’s desire to formulate robust, evidence-based policy.

## 2.3 The APS needs in-house expertise to evaluate evidence

One approach to the problem of obtaining expert advice has been to use scientific advisory committees. Although the solution is not without value, it is, at best, only a partial solution.

What can be achieved using in-house experts is different from what can be achieved using advisory committees. Put bluntly, for advisory committees, it’s not their main job. As a consequence, the greater part of the intellectual endeavours of the members are not directed towards fulfilling the needs of the APS. Further, advisory committees do not generally leave—nor even have the opportunity to leave—a legacy of increased understanding in the APS. Nor do they have significant influence on internal agency processes.

In contrast, in-house experts would

* have an integrated understanding of the needs and culture of their agency
* understand the nexus between their APS agency and the broader political process
* allow all levels of agency staff to access expertise, as needed
* leave a legacy of understanding in the generalist staff
* be in a position to build relationships of influence within their agency
* contribute to building a culture of evidence-based policy, and a research culture, within the APS.

## 2.4 Becoming an APS generalist is not conducive to developing domain-specific expertise

Flexibility and agility are the buzzwords of the modern APS. They form the title of Chapter 7 of ROAGA and are touted as being essential attributes of the APS organization. However —

*[w]hile it is, of course, understandable that the Government seeks to develop civil servants with transferable skills, it is short-sighted if that precludes highly-skilled experts who wish to remain experts within their field from progression to the upper echelons of the civil service or if its view of which transferable skills are valuable is too narrow. The tendency for civil servants to rotate between jobs on a regular basis (often in order to broaden their experience) can also be detrimental in specialist areas where accumulation of knowledge and experience is particularly important.* ([Scientific Advice, Risk and Evidence Based Policy Making](#UKHCSTC), paragraph 52)

Expertise takes a long time to develop and a great deal of effort to maintain.

On what basis do we make that assertion? Answer: evidence. A universal conclusion from studies of expertise in a wide variety of domains is that the development of expertise requires around 10 000 hours, or ten years, of practice.[[7]](#footnote-7) After 10 000 hours of attentive practice within the APS one can expect that a person will approach being an expert in the workings of the public service. But it is only with a similar degree of experience in acquiring knowledge and skill in the social sciences, together with practice in the exercise of that knowledge and skill, that expertise in the evaluation of social scientific evidence is likely to develop.

Consequently, the APS needs to acquire well integrated, in-house expertise—but that requires that it be able to attract and retain experts. There are a number of barriers which the APS must address if it is to make itself an attractive place for experts to work.

## 2.5 The APS must develop a research culture and value ‘in-house’ expertise

A member of the APS Senior Executive Service once suggested to us that we move out of doing evaluation-related work. When we remarked that it was in the area of program evaluation that our expertise lay, their response was to say that they felt strongly that no one in the APS should lay claim to any form of expertise!

It might be astonishing that such an opinion could be held by someone in a position of influence within the APS but it simply reflects the fact that *in-house* expertise is neither well recognised nor particularly valued within the Service. In situations where expertise might be found in-house or procured through a consultancy, the consultant is frequently perceived as having superior skills and knowledge, whatever the objective reality. The in-house problem has long historical antecedents[[8]](#footnote-8) and is certainly not unique to Australia. It was commented on, for example, by the House of Commons Science and Technology Committee in the following terms:

*Determining which expertise should be retained in-house and which sought externally is of critical importance but there are many compounding factors that render this process extremely challenging. Most departments have not been particularly effective at collecting data on the skills and experience of their employees in a way that facilitates assessment of the overall expertise held in-house—a problem exacerbated by the fact that… employees may not advertise their specialist skills in a generalist oriented civil service.* ([Scientific Advice, Risk and Evidence Based Policy Making](#UKHCSTC), paragraph 60)

As Gary Banks, Chairman of the Productivity Commission, put it in his address to the Australia and New Zealand School of Government in February 2009:

*Any agency that is serious about encouraging an evidence-based approach needs to develop a ‘research culture’. Establishing dedicated evaluation units, achieving a critical mass of researchers, strengthening links with academic and other research bodies, are all integral to this.* ([Banks](#Banks) 2009, p 21)

## 2.6 The APS must actively engage in helping in-house experts to maintain their expertise

If the APS is to be an attractive employer for social scientists then it must enable them to: (a) maintain and develop their expertise, and (b) maintain a *claim* to recognition of that expertise within the wider community of scholars.

These two ingredients are linked but are not identical. An in-house expert who fails to maintain and develop their expertise will be of ever decreasing value to the APS. Their knowledge will become dated, their skills will decline, and their real contribution to policy formation will plummet. On the other hand, an expert who cannot maintain an ongoing claim to expertise within their wider community of scholars will continue to be of value to the APS but will run the very real risk of being unwanted by any other employer!

For most scientists and social scientists, it is the process of researching, writing, and ultimately publishing which significantly contributes to sustaining and developing their expertise[[9]](#footnote-9)—but APS agencies are renowned for not publishing. The reasons for not publishing are as creative as a schoolboy’s excuses for unfinished homework—the Minister needs to make a decision, the information is confidential, the material is being considered by Cabinet, the report is commercially sensitive, the research isn’t relevant to policy, the research is *too* relevant to policy, the research is too old, too new, and so forth. The end result is that little is published, and the APS can turn out to be an efficient machine for turning experts into non-experts, or, at the very least, into experts who cannot substantiate their claim to expertise.

Needless to say, that is not the most desirable of outcomes.

## 2.7 The APS must develop a career path that rewards expertise

Without a career path which values and rewards their expertise, why would scientists and academics undertake periods of work within the APS? They won’t—or not sufficient numbers of them.

*We remain concerned that despite the introduction of PSG* [Professional Skills for Government]*,* *the civil service is still likely to be a place where generalist skills are valued more highly than specialist ones.* ([Scientific Advice, Risk and Evidence Based Policy Making](#UKHCSTC), paragraph 52)

*In policy-making, scientific literacy must be given equal value to economic knowledge and drafting ability…* ([Scientific Advice, Risk and Evidence Based Policy Making](#UKHCSTC), paragraph 53)

*… reform of the civil service is essential to bring about a cultural shift: specialist skills must be given equal value to generalist skills and this should be reflected in rewards structures.* ([Scientific Advice, Risk and Evidence Based Policy Making](#UKHCSTC), paragraph 53)

*It is also essential that more opportunities are created for scientists to progress to the most senior positions without being required to sideline their specialist skills.* ([Scientific Advice, Risk and Evidence Based Policy Making](#UKHCSTC), paragraph 53)

Reward is more than remuneration. One glaringly obvious problem that the APS must overcome if it is to succeed in attracting experts into its workforce is the lack of commensurate structures of remuneration *and* responsibility in universities and the APS.

The education of a social scientist is directed towards developing independent mastery of their domain of expertise—the ability to think independently, and to make and be responsible for their own judgements, is central to their expertise and effectiveness. Independence of thought is also highly valued by researchers. But while the remuneration offered to EL1s[[10]](#footnote-10) is broadly commensurate with the remuneration offered to the top end of the Senior Lecturer range or the bottom of the Associate Professor range,[[11]](#footnote-11) the duties and *responsibilities* of EL1s are not commensurate with this academic comparison group. Thus there is a disjunction in the way that remuneration and responsibility are interlinked in the APS and the way they are interlinked in universities. Transferring into the APS from a university at the same level of remuneration will considerably reduce individual responsibility—and independence. This is likely to be a severe disincentive for many experts who might otherwise consider joining the APS.

The APS must develop career paths which recognise and reward expertise. This requires creating the circumstances where internal experts can influence decisions, make judgement calls, and think independently. It will not be simple and it might require some experimentation. However, there are relevant Australian models which could provide a place to begin, the Reserve Bank Service[[12]](#footnote-12) being one such example.

## 2.8 Other approaches to gaining in-house expertise

In addition to building expertise in the APS by direct recruitment, another avenue would be to improve the exchange of ideas and people between the APS, universities and other research institutions. Academics could be seconded into the APS—as Professor Andrew Leigh of the ANU Research School of Social Sciences was seconded to Treasury in 2008. Similarly, postgraduate researchers could be encouraged to undertake their PhD research on topics of importance to the APS thus yielding doctoral graduates who are ‘tailor-made’ to fit the APS’s need for expertise.[[13]](#footnote-13)

# 3. Summary

Solving many of the current policy problems that Australia faces will require the input of scientists and experts from many disciplines. If the APS is to fulfil its responsibility to act as a professional and rational advocate of ideas that are in the best long-term interests of Australia and thereby develop the evidence-based policy which is the objective of the Australian Government, then the APS must have the capacity to interpret and evaluate evidence from all branches of the sciences.

Currently, the APS has limited capacity to evaluate essential evidence from the social sciences and it relies largely on non-experts who do not have the requisite skills and knowledge. The process of developing domain-specific expertise is diametrically opposed to the processes that are involved in acquiring the depth of generalist skills which the APS requires in the majority of its staff. Consequently, the APS needs the advice of experts. One approach to the problem of obtaining expert advice has been to use scientific advisory committees but that approach provides only a partial solution to the problem. A much more complete solution would be for the APS to attract and retain, in-house, the expertise it needs but there are a number of barriers that it must overcome if it is to succeed.

If the APS is to make itself an attractive place for experts to work, then it must

* nurture a culture of research that values expertise
* enable experts to maintain and develop their expertise and to maintain a claim to recognition of their expertise outside of the APS
* develop career paths that recognise and reward expertise.

These would be significant changes. They will not be easy to achieve and some experimentation will be required. The reward will be a much improved APS with the capacity to achieve its fundamental purpose.

# 4. References

Banks, G. (2009). *Evidence-based policy-making: What is it? How do we get it?* Chairman, Productivity Commission. Australian and New Zealand School of Government (ANZSOG)/Australian National University (ANU) Lecture Series on 4 February 2009. [[web-link](http://www.pc.gov.au/__data/assets/pdf_file/0003/85836/cs20090204.pdf)]

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100, 363–406.

Firebaugh, G. (2008). *Seven Rules for Social Research*. Princeton, NJ: Princeton University Press. [[web-link](http://press.princeton.edu/titles/8593.html)]

Head, B. (2008). Three lenses of evidence-based policy. *Australian Journal of Public Administration*, 67, 1–11. DOI: 10.1111/j.1467-8500.2007.00564.x [[web-link](http://dx.doi.org/10.1111/j.1467-8500.2007.00564.x)]

KPMG. (2009, November). *Benchmarking Australian Government Administration Performance*. [[web-link](http://www.pmc.gov.au/consultation/aga_reform/docs/benchmarking_australian_government_KPMG.pdf)]

Leigh, A. (17–18 August, 2009). *Evidence-Based Policy: Summon the Randomistas?* Paper presented at Strengthening Evidence-Based Policy in the Australian Federation, Productivity Commission Roundtable, Canberra, ACT. An early version is available at [[web-link](http://econrsss.anu.edu.au/~aleigh/pdf/PC_Randomistas.pdf)]: accessed 17 November 2009.

Murdoch University. (2006). Murdoch University Academic Staff Collective Workplace Agreement 2006. Murdoch, WA: Author. [[web-link](http://www.murdoch.edu.au/hr2/information/muascwa2006.pdf)]

Petrosino, A., Turpin-Petrosino, C., & Buehler J. (2002). “Scared Straight” and other juvenile awareness programs for preventing juvenile delinquency. Cochrane Database of Systematic Reviews, Issue 2. Art. No.: CD002796. DOI: 10.1002/14651858.CD002796. [[web-link](http://www.cochrane.org/reviews/en/ab002796.html)]

Reserve Bank of Australia. (2008). Reserve Bank of Australia Annual Report 2008. Sydney: Author. [[web-link](http://www.rba.gov.au/PublicationsAndResearch/RBAAnnualReports/2008/Pdf/2008_annual_report.pdf)]

United Kingdom. House of Commons Science and Technology Committee. *Scientific Advice, Risk and Evidence Based Policy Making*. Seventh Report of Session 2005–06, Volume I. London: The Stationery Office Limited. Available 27 October 2009, [[web-link](http://www.publications.parliament.uk/pa/cm200506/cmselect/cmsctech/900/90002.htm#evidence)].

United Kingdom. The BSE Inquiry. Lord Phillips of Worth Matravers. *The inquiry into BSE and variant CJD in the United Kingdom*. Volumes 1–14. House of Commons. [[web-link](http://www.bseinquiry.gov.uk/report/index.htm)]. Retrieved 27 October 2009.

United Kingdom. Foot and Mouth Disease 2001: Lessons to be Learned Inquiry. (2002) Report HC888. London: HMSO. [[web-link](http://archive.cabinetoffice.gov.uk/fmd/fmd_report/report/index.htm)]

# Appendix 2: Evaluation—A neglected but vital reform

Paper submitted 25 July 2010 to the Senate Finance and Public Administration Committee Inquiry into the Ahead of the Game Blueprint for the Reform of Australian Government Administration.

**Evaluation: a neglected but vital reform**

Submission to the Parliament of Australia

*Senate Finance and Public Administration Committee*

 *Inquiry into the Ahead of the Game Blueprint for the*

*Reform of Australian Government Administration*

[**Mark R. Diamond *MPsych, PhD***](#MarkDiamond)

**and**

[**Angela O’Brien-Malone *BSc, BA(Hons), PhD***](#AngelaOBrienMalone)

# Table of contents

[Background 1](#_Toc267773625)

[Evaluation: essential to evidence-based policy 1](#_Toc267773626)

[Money spent ineffectively is money wasted 1](#_Toc267773627)

[Evaluation requires expertise 2](#_Toc267773628)

[Ensuring honesty and independence in evaluation 2](#_Toc267773629)

[Most evaluations are buried 3](#_Toc267773630)

[Summary of recommendations 4](#_Toc267773631)

[References 5](#_Toc267773632)

[About the authors 6](#_Toc267773633)

# Background

In the discussion paper entitled *Reform of Australian Government Administration: Building the World’s Best Public Service*, the Government’s vision for the APS is described as including “evidence-based policy making processes as part of a robust culture of policy contestability” ([p. 18](#AGRAGA1)).

Although the importance of evidence-based policy was emphasized in twenty seven submissions to the Advisory Group, it is sidelined in the Advisory Group’s report.Program evaluation is an essential mechanism underlying evidence-based policy. *Ahead of the Game* fails the APS and the Australian public by neglecting to grapple with the integrity and quality of evaluation processes within the APS by which evidence of program effectiveness is obtained.

In this submission, we address critical reforms that are needed to support a robust culture of policy contestability and so make the APS more effective in delivering evidence-based policy. We include 12 recommendations for facilitating and embedding these reforms in the APS.

# Evaluation: essential to evidence-based policy

## Money spent ineffectively is money wasted

Section 44 of the Financial Management and Accountability Act 1997 requires that Commonwealth resources, including public money, be used ethically, *effectively* and efficiently.

When public money is spent on programs that do not work, the money is lost. Audit is the means to detect and prevent loss through theft or inefficiency. Evaluation is the means to detect and prevent waste due to ineffectiveness.

Evaluation should not be confused with program monitoring. Program monitoring is largely an audit function concerned with how much money was spent by a program, how it was spent, when it was spent, what it was spent on (people, widgets, services, etc.), and whether the expenditure was lawful. Evaluation, in contrast, is concerned with whether a program, and hence whether the expenditure, achieved the policy objective to which it was directed.

It is only by means of thorough, honest and well designed evaluations that Agencies can inform themselves, and their Ministers, about what works, and equally importantly, what does not work.

**Recommendation 1**: Evaluation of programs should be given the same status as audit of programs.

**Actions to make it happen**: The Financial Management and Accountability Act 1997 could be amended to give evaluation functions the same status within Agencies as audit functions.

## Evaluation requires expertise

An examination of the database of government contracts on the AusTenders website[[14]](#footnote-14) shows that millions of dollars of public money are spent each year on evaluation consultancies. Most of that money is wasted. It is wasted because, by the time an evaluation consultant is hired, the evaluation has already been seriously compromised.

Most evaluation activities within the APS are shaped by people who have no professional technical expertise in evaluation. Poor evaluation planning, poor evaluation design, and poor drafting of the evaluation request-for-tender mean that it will be impossible (even in principle) for a consultant to discover whether a program is, or is not, effective.

Authentic evaluation requires professional expertise, particularly in evaluation design and in the assessment of evidence. Notwithstanding the proliferation of evaluation sections within APS agencies, very few APS employees have expertise in evaluation. Most evaluation sections are staffed by generalist APS employees whose background is in the workings of the public service, not in the technical areas of evaluation design and evidence assessment.

**Recommendation 2**: The APS should recruit and retain staff with professional qualifications in social science research methodology or evaluation. Agencies should be encouraged to create career paths for evaluation experts.

**Recommendation 3**: Program evaluations should be designed by professional evaluation experts not by generalist staff. Evaluation functions with each Agency should be centralized to maximize the use of limited evaluation expertise.

**Recommendation 4**: Evaluation functions within each Agency should be overseen by an SES member with appropriate professional qualifications in evaluation. The SES member should have status equivalent to that of the Agency auditor.

**Recommendation 5**: The Auditor-General should be asked to review Commonwealth program evaluation practices and consultancies with a view to reporting on their effectiveness.

## Ensuring honesty and independence in evaluation

External evaluation consultants rely for their incomes on repeat-business from Agency program areas. Program areas rely on favorable reports from evaluation consultants if their programs are to continue. These circumstances do not favour independence and honesty in evaluations. If governments are to be properly informed about the effectiveness of programs—that is, whether the expenditure of public money is achieving the policy objective that was intended—then program evaluation must be unbiased in its design and execution, and untainted by vested interests.

Although some of the structural changes needed to promote independence in evaluation are addressed in Recommendations 3 and 4 above, additional changes are needed.

**Recommendation 6**: Evaluation consultancies should be overseen by centralized evaluation expertise within each Agency, with a view to assuring the quality, honesty and integrity of program evaluations.

**Recommendation 7**: The SES officer who is responsible for the centralized evaluation function within an Agency should also be responsible for maintaining the independence and integrity of evaluations.

**Recommendation 8**: The evaluation of tenders for evaluation consultancies should be free of undue influence by program staff.

**Recommendation 9**: The report of an evaluation should be delivered to an independent officer within each Agency such as the SES officer responsible for the centralized evaluation function.

## Most evaluations are buried

*Even the best evidence is of little value if it’s ignored or not available when it is needed. (Banks* [*2009*](#Banks)*, p. 18)*

The majority of program evaluation reports do not see the light of day. Although some Agencies have a good record for publishing evaluation reports, most do not. The excuses for non-publication are numerous but the effect is the same: the Parliament and the Australian public are denied the opportunity to discover what is working, and what is not working—to discover where public money is being spent effectively, and where it is being wasted on programs that do not work.

A reading of Hansard reveals that even Senate Estimates Committees have difficulty discovering: what evaluations have been done; how well designed they were; whether they could, in principle, have determined the effectiveness or otherwise of programs; whether they were well executed; and what they have shown. Although the existence of some, otherwise hidden, evaluation reports might be revealed by close examination of the online database of government contracts[[15]](#footnote-15), the site does not report many smaller contracts. Furthermore, the delay between initiation of an evaluation and reporting of the contract on the AusTender website is significant.

**Recommendation 10**: That Senate Estimates Committees inquire more assiduously into the existence and findings of evaluation reports of Agency programs.

**Recommendation 11**: That Agencies be required to table a list of evaluation reports and evaluation consultancies with the Parliament.

**Recommendation 12**: That Senate Estimates Committees be encouraged to seek advice from evaluation experts in framing questions to Agency heads regarding claims of program effectiveness.

# Summary of recommendations

**Recommendation 1**: Evaluation of programs should be given the same status as audit of programs. [[link to relevant section](#_Money_spent_ineffectively_is money )]

**Recommendation 2**: The APS should recruit and retain staff with professional qualifications in social science research methodology or evaluation. Agencies should be encouraged to create career paths for evaluation experts. [[link to relevant section](#_Evaluation_requires_expertise)]

**Recommendation 3**: Program evaluations should be designed by professional evaluation experts not by generalist staff. Evaluation functions with each Agency should be centralized to maximize the use of limited evaluation expertise. [[link to relevant section](#_Evaluation_requires_expertise)]

**Recommendation 4**: Evaluation functions within each Agency should be overseen by an SES member with appropriate professional qualifications in evaluation. The SES member should have status equivalent to that of the Agency auditor. [[link to relevant section](#_Evaluation_requires_expertise)]

**Recommendation 5**: The Auditor-General should be asked to review Commonwealth program evaluation practices and consultancies with a view to reporting on their effectiveness. [[link to relevant section](#_Evaluation_requires_expertise)]

**Recommendation 6**: Evaluation consultancies should be overseen by centralized evaluation expertise within each Agency, with a view to assuring the quality, honesty and integrity of program evaluations. [[link to relevant section](#_Ensuring_honesty_and_independence i)]

**Recommendation 7**: The SES officer who is responsible for the centralized evaluation function within an Agency should also be responsible for maintaining the independence and integrity of evaluations. [[link to relevant section](#_Ensuring_honesty_and_independence i)]

**Recommendation 8**: The evaluation of tenders for evaluation consultancies should be free of undue influence by program staff. [[link to relevant section](#_Ensuring_honesty_and_independence i)]

**Recommendation 9**: The report of an evaluation should be accepted by an independent officer within each Agency such as the SES officer responsible for the centralized evaluation function. [[link to relevant section](#_Ensuring_honesty_and_independence i)]

**Recommendation 10**: That Senate Estimates Committees inquire more assiduously into the existence and findings of evaluation reports of Agency programs. [[link to relevant section](#_Most_evaluations_are_buried)]

**Recommendation 11**: That Agencies be required to table a list of evaluation reports and evaluation consultancies with the Parliament. [[link to relevant section](#_Most_evaluations_are_buried)]

**Recommendation 12**: That Senate Estimates Committees be encouraged to seek advice from evaluation experts in framing questions to Agency heads regarding claims of program effectiveness. [[link to relevant section](#_Most_evaluations_are_buried)]

# References

Advisory Group on Reform of Australian Government Administration (2009). *Reform of Australian Government Administration: Building the world’s best public service*. Canberra, Australia: Department of the Prime Minister and Cabinet. [[web-link](http://www.dpmc.gov.au/publications/aga_reform/aga_discussion/index.cfm)]

Advisory Group on Reform of Australian Government Administration (2010). *Ahead of the Game: Blueprint for the Reform of Australian Government Administration*. Canberra, Australia: Department of the Prime Minister and Cabinet. [[web-link](http://www.dpmc.gov.au/publications/aga_reform/aga_reform_blueprint/index.cfm)]

Banks, G. (2009). *Evidence-based policy-making: What is it? How do we get it?* Chairman, Productivity Commission. Australian and New Zealand School of Government (ANZSOG)/Australian National University (ANU) Lecture Series on 4 February 2009. [[web-link](http://www.pc.gov.au/__data/assets/pdf_file/0003/85836/cs20090204.pdf)]

O’Brien-Malone, A., & Diamond, M. R. (2010). *On Rational Advocacy: A response to the discussion paper entitled “Reform of Australian Government Administration: Building the world’s best public service”.* Submission Number 0172 at http://www.dpmc.gov.au/consultation/aga\_reform/index.cfm . Canberra, Australia: Department of the Prime Minister and Cabinet. [[web-link](http://www.markdiamond.com.au/download/Submission-to-ROAGA.pdf)]

# About the authors

Dr Mark Diamond is an evaluation expert with a long record of publication in peer reviewed journals on matters related to evaluation and research. He has been a research and evaluation consultant to the private sector and to government. He is currently employed part-time within the APS and at the University of Tasmania. He maintains a blog on matters related to ethics, evaluation and research at [www.markdiamond.com.au](http://www.markdiamond.com.au) .

Dr Angela O’Brien-Malone has extensive experience as a senior academic with specialist expertise in evaluation and research methodology. She is currently employed part-time within the APS and holds appointment as an Honorary Research Associate in the School of Psychology at the University of Tasmania.

# Appendix 3: Research to government and back again

Paper submitted 2 August 2010 to the Australian Government Department of Innovation, Industry, Science and Research.

 **Research to government and back again: addressing cross‐sectoral mobility**

Response to the consultation paper

*Meeting Australia’s research workforce needs: A consultation paper to inform the development of the Australian Government's research workforce strategy*

**Angela O’Brien‐Malone *BSc, BA(Hons), PhD***

**and**

**Mark R. Diamond *MPsych, PhD***

**Summary**

In this submission we argue that the APS needs to become significantly better at using research expertise than it is at present if it is to fulfill its responsibilities as a professional and rational advocate of ideas that are in the best long‐term interests of Australia. Currently, the APS has limited capacity to evaluate essential evidence from the social sciences yet it does not present itself as an employer of choice for the research experts that it so desperately needs. Three significant problems for research experts who might consider working for the APS are that: (a) there are limited career opportunities for experts within the APS, (b) experts will find it difficult to maintain their expertise, and (c) they will find it difficult to maintain a reputation as an expert within the wider community of researchers. We make recommendations addressing these difficulties.

**Table of contents**

Summary ii

1 Introduction 1

2 Background 1

3 Why aren’t researchers rushing into the APS? 2

4 Why so little researcher mobility? 2

 Publications, track record 2

 Remuneration and reward 3

 Research expertise is not understood or valued 3

 Creating career paths based on expertise 4

 Attracting research expertise to the APS 4

5 Recommendations 5

6 References 6

Background of the Authors 7

**1 Introduction**

The consultation paper, *Meeting Australia’s research workforce needs: A consultation paper to inform the development of the Australian Government's research workforce strategy,* covers a number of issues which impinge on Australia’s research workforce. One significant issue which is addressed in the consultation paper is that of mobility:

*5.311 Available indicators suggest that Australia’s research workforce (like that of other countries) is highly mobile, with research careers in many cases spanning national (and in some cases sectoral) boundaries and multiple organisations of employment.*

*5.312 There is also evidence to support strong positive returns to Australia arising from this mobility.…*

*5.313 In spite of these benefits and positive existing activity, however, Australian researchers cite challenges in maintaining desired levels of mobility both within Australia and internationally. Challenges fall into two areas: disincentives to transitions between sectors and into or out of the research workforce in Australia, and barriers to international mobility….*

The focus of discussion in this submission is on mobility between the public sector and research institutions. Because we are social scientists, our discussion is particularly focussed on researchers in the social sciences although the issues we raise are also relevant to other research domains.

**2 Background**

The public sector has the potential to be the largest consumer in Australia of research related to the social sciences. It is easy to see why. Consider the following policy areas and pressing research questions—questions which fall wholly or in part into the realm of the social sciences:

* ***Education***: How can we improve literacy and numeracy outcomes for Indigenous children in remote Australia?
* ***Early childhood***: What is the best curriculum for early childhood education?
* ***Justice***: Can we predict who will reoffend and in what circumstances?
* ***Employment***: How can we help adolescents to make a successful transition from school to work or further study?
* ***Health***: Is there an obesity epidemic and, if so, what can we do to stop it?
* ***Environment***: How can we change people’s patterns of car usage?
* ***Infrastructure***: How can we predict the future demographics of our cities so that we can plan for future infrastructure needs?
* ***Science***: How can we identify breaking developments in science that might be important for Australia?

Collecting evidence within the social sciences is notoriously hard to do well—prompting Leigh’s (2009) observation that the “soft‐sciences” might more accurately be termed the hard sciences—and even the very best research comes with many caveats and difficulties in interpretation. Skill in interpreting evidence is what a researcher in the social sciences will have acquired, and ready access to such skill is essential if the APS is to provide good policy advice to governments. One might, therefore, expect that expertise in research would be highly sought‐after in the public sector, and that there would be significant movement of social scientists between research institutions, such as universities, and the public sector. But this is not so. Why not?

One answer is that, although the public sector would welcome such an exchange with open arms, it can’t get researchers to engage with them. Another answer is that the converse is true—researchers have found the public sector unwilling or unable to engage with them. The truth is considerably more complicated, as we argue in the remainder of this paper.

**3 Why aren’t researchers rushing into the APS?**

In principle, there are several potential gains for researchers from spending some part of their career in the APS. They include:

* the opportunity to work on substantial and interesting problems that require the translation of research knowledge into practice
* the prospect of enhancing one’s reputation by personally contributing to lasting improvements in problems of national importance
* the chance to form a network of contacts within the APS that might allow one to have an ongoing influence on public policy even after one’s departure from the public service
* facilitating the possibility of gaining future consultancies by developing an understanding of both government and the APS
* the opportunity to gain specific knowledge of government procurement processes—highly desirable knowledge both from the point of view of a modern university and the individual consultant researcher.

**4 Why so little researcher mobility?**

Why do so few researchers spend a part of their careers within the APS? The consultation paper touches on some of the reasons:

*5.314 With respect to disincentives to workforce transitions, a frequently cited obstacle to transition out of the academic workforce (whether for a temporary period of unemployment or movement to another sector) is publishing track record — a key factor for researchers in keeping pace with research activity in their chosen field and maintaining required levels of output to compete for grants and career progression opportunities. More broadly, a lack of portability of superannuation arrangements and prohibitive pay differentials are perceived to encourage unidirectional rather than more fluid transitions between areas of research employment, such as between academia and the business enterprise sector.*

These are fair points, but, with regard to ease of bidirectional movement between research institutions and the APS, they are not the only ones and they are open to being misconstrued.

***Publications, track record***

If social scientists are to be attracted into the APS then they must be able both to: (a) maintain and develop their expertise, and (b) maintain a claim to recognition of that expertise within the wider community of scholars. Although these two points are interlinked they are not identical, and both are related to publishing.

For most social scientists the process of researching, writing, and publishing is one that significantly contributes to them sustaining and developing their expertise—but APS agencies are renowned for not publishing. The reasons for not publishing are many and various but the end result is the same: little is published, and working in the APS can turn out to be an efficient method for transforming expert researchers into non‐experts, or into experts who cannot substantiate their claim to expertise. An in‐house expert who fails to maintain and develop their expertise will be of ever decreasing value to the APS. On the other hand, an expert who cannot maintain an ongoing claim to expertise will continue to be of value to the APS but will run the very real risk of being unwanted by any other employer. Understandably, researchers who are considering a move to the APS see this as a non‐trivial problem.

***Remuneration and reward***

Reward is more than money. One glaringly obvious problem that the APS must overcome if it is to succeed in attracting researchers into its workforce is the lack of commensurate structures of remuneration and responsibility in universities and the APS.

The education of a social scientist is directed towards developing independent mastery of their domain of expertise. The ability to think independently, and to make and be responsible for their own judgements, is central to their expertise and effectiveness. Independence of thought is also highly valued by researchers. But while the remuneration offered to APS staff at Executive Level 1 (EL1) is broadly commensurate with the remuneration offered to the top end of the academic Senior Lecturer range, the duties and responsibilities of EL1s are not commensurate with those of their academic colleagues; they are very much less significant. Transferring into the APS from a university at the same level of remuneration will considerably reduce the individual responsibility and independence of a social scientist. This difference in the way that remuneration and responsibility are interlinked in the APS and the way they are interlinked in universities is likely to be a severe disincentive for many researchers who might otherwise consider joining the APS.

***Research expertise is not understood or valued***

Gary Banks, Chairman of the Productivity Commission, noted in his address to the Australia and New Zealand School of Government in February 2009 that:

*Any agency that is serious about encouraging an evidence‐based approach needs to develop a ‘research culture’. Establishing dedicated evaluation units, achieving a critical mass of researchers, strengthening links with academic and other research bodies, are all integral to this. (Banks 2009, p 21)*

Banks is correct, but there are significant impediments to developing a research culture, not least of which is that APS generalist staff often do not understand what research is, or what evidence is, or why either are necessary. The APS is not alone in having this problem.

In November 2005 an inquiry into the UK Government’s handling of scientific advice, risk and evidence in policy making was launched by the House of Commons Science and Technology Committee. The tenor of comments in the report of the Inquiry is revealing. For example:

*Experts in the civil service (and from outside) need to work closely with civil service generalists if scientific advice and evidence are to be effectively incorporated into policy. William Solesbury, Senior Visiting Research Fellow at the Centre for Evidence Based Policy and Practice, told us that it was in this area—the competence of the generalist staff—that the civil service’s weakness lay: “I do not think there is, as yet, very much, or at least not a very sophisticated understanding of the occasions when evidence is useful, the sort of evidence to be obtained, how to evaluate evidence when it is available, how to interpret it, and how to weigh it.” (Scientific Advice, Risk and Evidence Based Policy Making, paragraph 48)*

The conclusion of the Inquiry was that while scientific knowledge and expertise was essential for evidence‐based policy development, the UK Civil Service had significant problems in understanding and evaluating evidence, and in valuing and incorporating expertise within the Service.

Is the situation any different in the APS? We think not. Although the APS could gain from having more in‐house research expertise, it is in the paradoxical position of lacking the very expertise that would enable it to understand its own need.

***Creating career paths based on expertise***

If Australia’s public service is to benefit from in‐house expertise, and Australia’s research workforce is to benefit from greater options for mobility into and out of the public sector, then the APS must develop career paths which recognise and reward expertise. Such a development would require creating the circumstances where internal experts can use and maintain their expertise, can continue to make a public claim to expertise, can influence decisions, can make judgement calls, and can think independently. It will not be straightforward but there are relevant, successful, Australian models, the Reserve Bank Service[[16]](#footnote-16) being one such example.

In addition to building expertise in the APS by direct recruitment, another avenue would be to improve the exchange of ideas and people between the APS, universities and other research institutions. Academics could be seconded into the APS. Similarly, postgraduate researchers could be encouraged to undertake their PhD research on topics of importance to the APS thus yielding doctoral graduates who are ‘tailor‐made’ to fit the APS’s need for research expertise.

***Attracting research expertise to the APS***

The issues we have discussed in this submission have generally been addressed as part of researcher mobility. However, the opportunity to work on applied problems and in organizations other than universities might well prove attractive to potential candidates for higher degrees by research (HDR) candidates, particularly given the demographic changes noted in the consultation paper:

*4.31 The changing nature of its HDR student cohort also presents some challenges to Australia. The average age of HDR candidates in Australia has increased over the last decade, with a significant proportion of students now commencing their degrees in the 30‐39 year age bracket and above. These demographic changes are reflected in the changing circumstances under which students are approaching studies, with many students having family and financial responsibilities that require more flexible modes of engagement with their degree...*

We note that there are already some resources available to support mobility of researchers between research organizations, including universities, and business.

*2.37 Through the Enterprise Connect Researchers in Business Program* [ECRBP]*, the government supports the placement (through support for up to 50 per cent of salary to a maximum of $50,000 for each placement) of researchers from universities or public research organisations in businesses for a period of two to twelve months. The scheme’s objectives include breaking down cultural barriers between sectors and speeding the dissemination of expertise.*

We suggest that something similar to the ECRBP be set up to encourage researchers to spend some part of their careers in the APS. This would foster the development of a research culture within the APS and engage researchers in directing their research efforts towards addressing significant policy issues. Such engagement between researchers and the APS would further the development of evidence‐based policy in Australia.

**5 Recommendations**

Fostering greater mobility of researchers in the social sciences between universities and the APS would be beneficial to researchers and universities and would enhance the APS’s ability to provide evidence‐based policy advice to government. These are potentially great gains for Australia, but there are significant impediments to this mobility. We make the following five recommendations.

**Recommendation 1**

That the APS be encouraged to establish and use career paths based on expertise.

**Comment**: Section 35 of the Public Service Act 1999 (Cth) establishes the Senior Executive Service of the APS, the function of which is to provide a group of APS employees, each of whom, within his or her Agency, provides one or more of the following at a high level: professional expertise, policy advice, management. Although some APS agencies (such as the Australian Bureau of Statistics) clearly recruit staff because of their professional expertise and then utilise that expertise, such is not the case with many other agencies where the norm appears to be that SES members are appointed on the basis of management skills.

**Recommendation 2**

That the careers of experts within the APS be structured so as to facilitate both the maintenance of expertise and the maintenance of a public claim to expertise.

**Comment**: By permitting Agency staff to publish research reports under their own name and under the banner of the Agency (but disclaiming representation of the Agency’s views), APS Agencies could greatly increase the likelihood that professional staff could retain their public claim to expertise. The Research Discussion Papers of the Reserve Bank of Australia (which commenced in 1969) provide a good model for such publications.

**Recommendation 3**

That the APS be encouraged to foster a culture of research and to develop a genuinely evidence‐based approach to policy development.

**Recommendation 4**

That candidates for higher research degrees be encouraged to consider undertaking their research with the context of the APS; and that the APS be encouraged to house and support such undertakings.

**Recommendation 5**

That an equivalent program to the Enterprise Connect Researchers in Business Program be set up to facilitate mobility of researchers into the APS.

**6 References**

Australia. Department of Innovation, Industry, Science and Research. (2010). *Meeting Australia’s research workforce needs: A consultation paper to inform the development of the Australian Government's research workforce strategy*. Canberra: Author. [web‐link]

Banks, G. (2009). *Evidence‐based policy‐making: What is it? How do we get it?* Chairman, Productivity Commission. Australian and New Zealand School of Government (ANZSOG)/Australian National University (ANU) Lecture Series on 4 February 2009. [web‐link]

Diamond, M. R., & O’Brien‐Malone, A. (2010). *Evaluation: a neglected but vital reform. Submission to the Parliament of Australia Senate Finance and Public Administration Committee, “Inquiry into the Ahead of the Game Blueprint for the Reform of Australian Government Administration”.* Submission number 1 at http://www.aph.gov.au/Senate/committee/fapa\_ctte/govt\_admin/submissions.htm . Canberra, Australia: Parliament of Australia, Senate. [web‐link]

Leigh, A. (17–18 August, 2009). *Evidence‐Based Policy: Summon the Randomistas?* Paper presented at Strengthening Evidence‐Based Policy in the Australian Federation, Productivity Commission Roundtable, Canberra, ACT. An early version is available at [web‐link]: accessed 17 November 2009.

O’Brien‐Malone, A., & Diamond, M. R. (2010). *On Rational Advocacy: A response to the discussion paper entitled “Reform of Australian Government Administration: Building the world’s best public service”.* Submission Number 0172 at http://www.dpmc.gov.au/consultation/aga\_reform/index.cfm . Canberra, Australia: Department of the Prime Minister and Cabinet. [web‐link]

United Kingdom. House of Commons Science and Technology Committee. *Scientific Advice, Risk and Evidence Based Policy Making*. Seventh Report of Session 2005–06, Volume I. London: The Stationery Office Limited. Available 27 October 2009, [web‐link].

**Background of the Authors**

Dr Angela O’Brien‐Malone has extensive experience as a senior academic with specialist expertise in evaluation and research methodology. She is currently employed part‐time within the APS and holds appointment as an Honorary Research Associate in the School of Psychology at the University of Tasmania.

Dr Mark R. Diamond is an evaluation expert with a long record of publication in peer reviewed journals on matters related to evaluation and research. He has been a research and evaluation consultant to the private sector and to government. He is currently employed part‐time within the APS and at the University of Tasmania. He maintains a blog on matters related to ethics, evaluation and research at [www.markdiamond.com.au](https://www.markdiamond.com.au/) .

# Appendix 4: Reform of non-financial reporting

Paper submitted 29 September 2014 to Australian Government Department of Finance.

**Reform of non-financial reporting: A window of opportunity for the honest appraisal of policy effectiveness**

Submission to the Australian Government Department of Finance

*A response to the document entitled* Enhanced Commonwealth Performance Framework—Discussion Paper

[**Angela O’Brien-Malone *BSc, BA(Hons), PhD***](#AngelaOBrienMalone)

 **and**

[**Mark R. Diamond, PhD**](#MarkDiamond)

#

# Table of contents

[What this response addresses 3](#_Toc399794361)

[The problem 3](#_Toc399794362)

[Deciding what works 3](#_Toc399794363)

[Assessment must be tied to the agents of change 3](#_Toc399794364)

[Non-financial reporting: What does government need to know? 5](#_Toc399794365)

[What should be measured and why? The role of ‘program’ logics 5](#_Toc399794366)

[Disengaging the factors that maintain inadequate evaluation 7](#_Toc399794367)

[There is a problem with evaluation quality 7](#_Toc399794368)

[What maintains poor evaluation? 8](#_Toc399794369)

[Australian Public Service staff 8](#_Toc399794370)

[Consultants 9](#_Toc399794371)

[What can be done? 10](#_Toc399794372)

[Design and analysis—problems at several levels 11](#_Toc399794373)

[Design of the performance evaluation process 11](#_Toc399794374)

[Designing evaluation for the life of an activity 12](#_Toc399794375)

[Implementation design 13](#_Toc399794376)

[Examples of performance measurement tools and approaches 13](#_Toc399794377)

[Proportionality 14](#_Toc399794378)

[Nomenclature 14](#_Toc399794379)

[Summary of recommendations 15](#_Toc399794380)

[References 16](#_Toc399794381)

[About the authors 17](#_Toc399794382)

# What this response addresses

The *Enhanced Commonwealth Performance Framework Discussion Paper* (the ‘discussion paper’, Australian Government Department of Finance, 2014) proposes a structure in which the non-financial outcomes of government-funded activities could be reported. In this response we analyse a serious problem that lies at the heart of the proposed approach. The problem is of sufficient magnitude to undermine the good results that the discussion paper is aimed at achieving. We make recommendations about how to avoid the problem and how to facilitate the aims of the *Public Governance Performance and Accountability Act 2013* (Cth) that are outlined in the discussion paper.

## The problem

The discussion paper presents a structure for reporting the non-financial outcomes of Commonwealth entities and ties the structure of that reporting to the same financial structures implied in Portfolio Budget Statements (PBS). Unfortunately the financial reporting structure of the Portfolio Budget Statements does not map cleanly onto either non-financial key-performance indicators (KPIs) or non-financial outcomes. It is frequently the case that the administered items listed under a PBS-level program do not all address the same (set of) KPIs or the same (set of) outcomes. Although the hierarchical organisation of activities into PBS-level programs provides a convenient way of indicating what appropriations will provide the money for each activity, and therefore also provides a framework for financial reporting, it obscures what is required for non-financial reporting.

**Recommendation 1**: That the Department of Finance divorce the reporting structure for non-financial outcomes from that used for financial reporting.

# Deciding what works

## Assessment must be tied to the agents of change

The ambiguous use of the word program[[17]](#footnote-17) within the discussion paper and more broadly might easily lead to the impression that Commonwealth entities deliver PBS-level-programs that bring about PBS-level-outcomes. But this is not so. Commonwealth entities use PBS-level program appropriations to fund and deliver various kinds of interventions (sub-programs, activities) that act to bring about change. Put another way, providing the money to the Commonwealth entity does not of itself bring about desired social change but spending it might do so.

The most important question for non-financial reporting is “Did it work?”. In this context, *it* can only refer to something that has causal agency; that is, *it* must refer to something that has a capacity to alter a social or policy related variable. PBS-level programs are conceptual accounting structures. They are not agents of change. Agency for change, in a causal sense, attaches to activities.

The Higher Education Participation and Partnerships Program (HEPPP) provides a useful illustrative example. The HEPPP was first referred to, though not by name, in the Department of Education Employment and Workplace Relations Portfolio Budget Statements of 2009–10,[[18]](#footnote-18) where it was described in the following terms:

Funding to support low SES participation—a new partnership program will be established and an increased loading for low SES students will become payable to providers. The partnership program is intended to promote leading practice and increase aspirations of students to higher education and increase the rates of higher education attainment of low SES, Indigenous, regional and remote students. Partnerships will be developed with schools in low SES areas and vocational education and training providers. Further, a loading will be paid to universities based on the number of low SES students enrolled.

The description appeared as one of 16 administered items[[19]](#footnote-19) for Program 3.1 (Higher Education Support) that was itself intended to contribute to Outcome 3 of the portfolio, namely:

A growth in skills, qualifications and productivity through funding to improve teaching quality, learning, and tertiary sector infrastructure, international promotion of Australia’s education and training sectors, and partnerships with industry

A significant aspect of the hierarchical organisation of expenditures in the Portfolio Budget Statements—from the top-level Outcome 3, through Program 3.1, to the individually administered item (in this case the HEPPP)—is that it would make no sense to attempt a non-financial evaluation of Program 3.1 because it is only an accounting convenience and does not have causal agency for change. To ask whether Program 3.1 was successful would be as meaningless as asking whether an individual’s bank account is “successful” in contributing to the individual’s growth, prosperity, and hopes for the future.[[20]](#footnote-20)

While the hierarchical arrangement that appears within Portfolio Budget Statements is clearly of great utility from the perspective of financial reporting, something else is essential if any kind of meaningful assessment is to be made of what activities, directed to what policies, “work”.

Exactly what it is that is necessary for good non-financial reporting is the subject of the next section.

## Non-financial reporting: What does government need to know?

There is no point in throwing good money after bad—government needs to know whether the activities that it funds contribute to bringing about intended changes. As the discussion paper puts it:

To improve government resource decision-making, performance information needs to be effectively and efficiently integrated into the budget process. (p. 41)

We agree. But the quoted statement does not describe what performance information is needed, nor how it is to be integrated into the budget process. Describing *what*, and explaining *how*, will be critical to the success, or failure, of an Enhanced Commonwealth Performance Framework.

If governments are (a) to discover which activities are effective and which are not, and (b) to avoid the waste of public money that comes from funding ineffective activities, then non-financial performance information directly relevant to those activities must be incorporated into the budgetary decision process.

Deciding whether an activity has been effective requires answering two questions. First, was there a change? An assessment must be made of the extent to which one or more policy-relevant variables have changed during the pertinent period of time. That assessment will reflect on whether something—though not necessarily the activity of interest—has brought about a policy-relevant change.

Second, was it the activity that brought about the change? An assessment must be made of the extent to which argument and evidence support an attribution of causal agency for the observed change to the activity of interest.[[21]](#footnote-21) If there is no evidence that a change has occurred, or if it appears that the activity is not responsible for the observed change, then there will be no reason to conclude that the activity has been successful. Activity-level evaluations are essential if government is to have any chance of discovering what is working and what is not.

**Recommendation 2**: That the reporting structure for non-financial outcomes be tied to activities that have causal agency.

## What should be measured and why? The role of ‘program’ logics

To conclude that a particular social policy and the activities it mandates are effective in bringing about a relevant policy objective depends both upon measuring things that are expected to change as a consequence of the activities and upon being able to link the cause of the changes to the activities. Put differently, one must be able to provide a causal theory linking the activities to the outcomes.[[22]](#footnote-22) Intelligent decisions about what to measure must necessarily be informed by the causal theory—decisions that should usually be made before program implementation.[[23]](#footnote-23)

It will usually not be sufficient to measure only the relevant KPI presented in the annual report of a Commonwealth entity. Consider for example, a policy intervention such as tobacco control. The aim of the policy might be to decrease the prevalence of lung-disease in the population and one of the activities used to bring about this outcome could be to advertise on television the harmful effects of tobacco smoking. It would obviously not be useful in the early years of the intervention to rely on measures of prevalence to determine whether the intervention is worth continuing. The hypothecated causal chain of events might take the following form:

(Activity) Television advertising of harmful effects
🡒 increased knowledge of harmful effects in the viewing population
🡒 increased desire in some members of the viewing population to quit smoking
🡒 increased attempts to quit smoking
🡒 increased numbers of people who succeed in quitting smoking
🡒 decreased risk of developing lung cancer in those individuals who quit
🡒 decreased numbers of individuals with lung cancer in the television viewing population
🡒 decreased overall population prevalence of lung cancer (Policy objective).

Each step in the causal chain serves as a necessary precursor to the next step, and the inferential path from activity to policy objective not only spans many years but will collapse if any of the hypothesised links proves to be invalid. The obvious, but rarely invoked, solution is to use causal theory to inform the specification of short-term, medium-term and long-term objectives and associated sets of intermediate KPIs. Other examples where a causal chain from activity to outcomes could sensibly be developed include:

* To increase university enrolments (KPI) it might first be necessary to increase the number of students who continue their education to the end of secondary school at Year 12 (intervening causal variable).
* To improve the health of new-born babies in remote communities (KPI), it might first be necessary to increase the number of pregnant women who receive information about, or training in, how to care for a baby (intervening causal variable).
* To improve the nutritional status of children it might first be necessary to improve the nutritional knowledge of mothers (intervening causal variable).[[24]](#footnote-24)

Regrettably, most causal chains (or ‘program logics’[[25]](#footnote-25)) for government social programs are highly underdeveloped. To confirm the early effectiveness of an activity, and so too the usefulness of government expenditure, it might be necessary for the program logic to describes the short-, medium- and long-term indices on which change should be observable. A program logic of that kind will also serve to introduce the idea that the high-level KPIs alone will be insufficient, and that some additional is needed .

Also, problematically for the structure of non-financial performance reporting proposed in the discussion paper, a policy related activity can contribute to achieving more than one KPI—even if the overarching policy goal nominates only one such KPI as being relevant for the activity. For example, the Higher Education Participation and Partnerships Program (HEPPP) exists to improve the enrolment and retention in university studies of low SES students. The HEPPP supports a range of activities, some of which take place within schools (rather than universities) and are directed in the short-to-intermediate-term to increasing the number of low SES students who complete secondary school. The activities are funded through the HEPPP under the reasonable assumption that completing secondary school is usually a necessary step towards enrolling in university studies. So activities that occur within the HEPPP actually address *two* high-order KPIs, and indeed two top-level PBS-level outcomes (!), one that relates to schools and one that relates to higher education. In cases such as this, not only does the assumption fail that each activity (or ‘administered item’) maps onto only one KPI, but non-financial reporting that examines PBS-level outcomes in terms exclusively related to the PBS-level programs under that outcome will also be misdirected.

Overall, the lack of investment in developing or explicating the causal chain from funded activity to policy-relevant outcome means that government is denied opportunities to enhance policy outcomes. It also leads to a failure to establish what must be measured in order to determine activity effectiveness and the appropriateness of further funding.

**Recommendation 3:** That the Department of Finance require that all proposals for new policy initiatives to describe explicitly the assumed causal chain between the activities that are to be funded and the anticipated outcomes.

**Recommendation 4:** That prior to implementation, program logic be used to decide what variables will index short-, medium- and long-term effectiveness of the activity. That funding be tied to the requirement that the proposed indices actually be measurable.

**Recommendation 5:** That the Department of Finance encourage entities to map the connections between the intermediate KPIs of the activities they administer and to use the map as the structure against which to report non-financial outcomes.

# Disengaging the factors that maintain inadequate evaluation

## There is a problem with evaluation quality

It is uncontroversial that many, though not all, evaluations of social policy initiatives in Australia are poorly conducted. The discussion paper itself notes that:

Key stakeholders such as the Joint Committee of Public Accounts and Audit and the ANAO have been critical of the … value of information currently available about the performance of Commonwealth programmes . (p. 11)

The current quality of external non-financial performance reporting in the Commonwealth public sector has significant scope for improvement. (p. 12)

Australia is not alone; authorities in other countries have made similar comments. For example, when the National Audit Office of the UK[[26]](#footnote-26) examined the design of 34 published evaluations and the extent to which the impacts measured in the evaluations could be attributed to the policy or program, they found that:

The assessment of the fitness for purpose of a selection of 34 evaluations from four departments … [found that only] 14 evaluations were of a sufficient standard to give confidence in the effects attributed to policy (p. 7).

Evaluation reports that were weaker in identifying causality tended to be more positive in assessing what the intervention achieved (p. 8).

In other words, the evaluations most likely to conclude that a program was highly effective were the very evaluations whose designs were incapable of sustaining inferences about causality! The situation in Australia does not appear to have been subjected to the same formal examination but it is unlikely to be different from that in the UK since a number of the factors which serve to maintain poor quality in evaluations are the same in both countries.

## What maintains poor evaluation?

### Australian Public Service staff

Although poor evaluation practice by Commonwealth entities is frequently attributed to a dearth of public servants with expertise in evaluation, or to a lack of evaluation capacity, there are reasons to doubt that view. First, there are now university courses[[27]](#footnote-27) which are targeted at evaluation knowledge and skills and graduates from those courses form part of the potential workforce that is integrated into the Australian Public Service (APS). Second, there are significant numbers of people within the Australian Public Service with at least some knowledge of evaluation practice as is evidenced, for instance, by the numbers who attend the Canberra Evaluation Forum.[[28]](#footnote-28) Thirdly, even if there is a shortage of public servants with evaluation expertise, a large number of university-based and private-sector consultancies, albeit with varying degrees of knowledge and expertise, are recorded on the multi-use lists that Commonwealth entities use to outsource program evaluation.

But if a lack of evaluation capacity is not the root cause of poor evaluations of social policy activities, what is? In a paper we presented at the 2014 International Conference of the Australasian Evaluation Society (AES) we examined a range of reasons. However, because the full paper will shortly become available through the AES website,[[29]](#footnote-29) we address here only two of those causes, both related to the well-recognised importance of behavioural contingencies.

Evaluations of Australian Government social policy activities are sometimes conducted by entity staff but are also frequently commissioned from external evaluation consultancy businesses. There is a considerable body of evidence and theory[[30]](#footnote-30) to suggest that the behaviour of people in both these circumstances is likely to be motivated by the structure of incentives.[[31]](#footnote-31) In the absence of carefully constructed *compensatory* incentives, the interests of evaluation consultants, government, and APS managers are unlikely to be in complete alignment.[[32]](#footnote-32) Serving the long-run interests of government necessitates discovering which activities are effective and which are ineffective, inefficient, or actively damaging to a policy goal. But managers are generally rewarded for managing money and staff well, and are generally not rewarded for showing that the money they are managing might be better spent on some other more policy-effective activity overseen by another manager, rather than on the ineffective and wasteful activity that they themselves are unfortunate enough to have been delegated to watch over. Indeed, with regard to the good evaluation of government-funded activities, there appear to be no instances of the use of specific evaluation-directed incentives[[33]](#footnote-33) to bring the interests of entity officers into close alignment with those of government.[[34]](#footnote-34)

The problem of staff career interests has existed for a while. In its report on policy making in the United Kingdom, the Institute for Government[[35]](#footnote-35) found that:

Civil servants often feel that their incentives are geared towards looking forwards to the next big policy issue. As one put it: “[…] looking back isn’t necessarily encouraged, and it’s not necessarily celebrated... people don’t see it as a part of their cycle of work.”

This isn’t new—around 440 BC Sophocles wrote:

None love the messenger who brings bad tidings(*Antigone*).

### Consultants

Evaluation consultants also face competing demands that create disincentives for behaviours that are in the long-run interest of government. Evaluation consultancy is a business that must deal with the necessity of making a profit and obtaining further work after the current contract finishes.

As the Institute for Government19 noted:

Since many evaluators depend on repeat contracts, they have powerful incentives to acquiesce in self-censorship. (p. 8).

The fact that an officer of a Commonwealth entity might be interested in promotion, salary advancement, and low-stress employment, and the fact of an evaluation consultancy being a profit-making enterprise, are not of themselves adverse to the interests of the Commonwealth. If the interests of the three parties have been correctly aligned by the arrangement of compensatory incentives,[[36]](#footnote-36) then the varied and sometimes hidden motivations of the parties will be unproblematic. It is when the interests of the consultant evaluator interact with those of the commissioner of the evaluation in a way that brings them into mutual alignment at the expense of the long-run interest of government that the risk of poor evaluation products is highest.

As long as organizational culture conflates the effectiveness of managed activities with the quality of the manager,[[37]](#footnote-37) the risk of poor evaluations will continue. Government social policies, and their related activities, are usually gambles, their outcomes rarely certain. To decide that an activity is ineffective, and to recommend that it be either redesigned or cease, should be seen (although it rarely is) as laudable program management and a solid evaluation outcome.

### What can be done?

A popular but false belief is that internally conducted program evaluations necessarily lack independence and must result in compromised reports, whereas evaluations undertaken by external consultants will necessarily be independent, uncompromised and veracious. As the previous section suggests, the dichotomy is unreliable; it is incentives that motivate reliable versus unreliable evaluation reporting.

Unless a way can be found to short circuit the hidden nexus between the business needs of external consultants and the career advantages of the staff of Commonwealth entities, then we will not bring about systematic improvements in the evaluation culture of those entities and non-financial reporting will suffer. Efforts to produce genuine evidence-based policy will be compromised just as they have been in the past.

Fortunately, there are still reasons for hope. Breaking the nexus requires having evaluators—whether external or internal—not be accountable to program management staff. This idea is simply a carry-over from the lessons learned in the audit profession. In our 2010 Senate submission[[38]](#footnote-38) we recommended that government should treat evaluation within agencies in a broadly parallel fashion to the way that audit is treated. The reporting lines and decision-making powers related to evaluation and evaluation staff should be divorced from those of the area that manages the program to be evaluated. Having entity audit staff report out-of-line from the managers of the activities that they audit helps to ensure the impartiality of internal auditors. Similarly, having entity evaluation staff report out-of-line would help to ensure their impartiality, especially if solid incentives are created for robust evaluation reporting. If evaluation consultancies are commissioned by out-of-line internal evaluation staff then consultants are more likely to be rewarded for producing sound and unbiased evaluations. Finally, entity evaluation staff would have the possibility of career advancement based on producing impartial evaluation work.

Two final points. First, the skill sets involved in reporting and investigating financial performance and in evaluating non-financial performance are very different. The amalgamation of evaluation tasks into the roles and responsibilities of existing audit staff will not bring about improvements in evaluating non-financial performance. The temptation to amalgamate should be avoided.

Second, social policy evaluation is an applied social science. The accurate appraisal of the effectiveness of social policy requires methods that are not ad-hoc or founded on intuition.[[39]](#footnote-39) To assign evaluation tasks to entity staff who are ignorant of good evaluation practice is to cripple an evaluation from the outset. Nor is it the way to build an evaluation culture, establish evaluation capacity, improve evaluation outcomes, or build the reputation of an entity. Evaluation expertise is required within entities as well as outside of them.

**Recommendation 6:** That evaluation within agencies be treated in a broadly parallel fashion to the way that audit is treated; that the reporting lines and decision-making powers related to evaluation and evaluation staff be divorced from those of the area that manages the program to be evaluated.

**Recommendation 7:** That evaluation tasks and roles *not* be amalgamated into the responsibilities of existing audit staff.

**Recommendation 8:** That entities be encouraged to assign staff with evaluation knowledge and skills to perform internal evaluation functions.

# Design and analysis—problems at several levels

## Design of the performance evaluation process

Government needs to know whether the money it is spending to fund activities is making a difference. An essential part of improving the quality of information that government gets is to make sure it gets the full picture.

As the discussion paper makes clear, transparency is necessary and is facilitated by publication:

Performance information, including commentary on and acknowledgement of factors that may have limited performance, should be publically available. Information should be made available in a way that enables users to track the contribution an activity’s performance makes to achieve the government’s strategic priorities (p. 40).

But, as the Institute for Government noted, it is unfortunately the case that:

Departments have the incentives and opportunity to tone down [or, suppress—see p. 50] unfavourable findings (p.8).

One solution to this problem of image management is to require that all evaluations be registered with a central agency *prior* to being undertaken and all evaluation reports be published.[[40]](#footnote-40)

**Recommendation 9:** That the Department of Finance require all evaluations, prior to being undertaken, to be entered on a public register maintained by a central agency.

**Recommendation 10:** That the Department of Finance require all evaluation reports to be published.

## Designing evaluation for the life of an activity

A major limitations that plagues the evaluation of social-policy activities is that, in practice though not in principle, many activities simply cannot to be evaluated for impact or attribution. By that we mean that it is impossible to know whether any change has taken place and, if it has, equally impossible to know whether the activity of interest was responsible for it. Frequently:

* baseline data was not collected
* the logic of the activity was never specified so appropriate intermediate indices on which to measure change were never identified
* relevant data was never collected over the life of the activity
* changes —sometimes many over many years—were made to the program without an adequate record having been made of the changes
* the activity was implemented in a way which meant that, long before any evaluator had contact with it, it would be impossible to address attribution in any strong sense.

Evaluation planning should begin at the beginning, before an activity is implemented. It should begin before a single dollar is spent on the activity. Once money is out the door, it is already too late.

Evaluation should be planned for the life of an activity so that the (potential) wave of change can be tracked. An evaluation plan that is established prior to roll-out of the activity, and which is informed by an analysis of the causal logic for the activity, is the only possible way of achieving this. The short, medium- and long-term indicators of change identified in the evaluation plan should be integrated into each entity’s non-financial performance reporting and mapped to relevant PBS-level KPIs and PBS-level outcomes.

**Recommendation 11:** That the evaluation of activity effectiveness be planned prior to implementation and be integrated into the non-financial performance reporting of entities.

## Implementation design

The necessity of *designing* the implementation of activities is vastly underrated. Implementation designs should include the development of the activity logic and subsequent identification and collection of baseline data on the KPIs and other relevant variables in the causal chain that leads from the action of the activity to the desired outcome. Implementation design should (1) include a consideration of piloting (in the technical sense) the activity,[[41]](#footnote-41) and (2) design the national roll-out of the activity in a way that allows for assessment of its effectiveness.

Contrary to what is frequently thought, piloting an activity is not the same as ‘trying things out’. A technical piloting of the activity comprises the collection of data using a specified design, with a corresponding analysis, that allows for alternative explanations for change (alternative to the activity, that is) to be excluded. Such technical piloting of activities should be undertaken whenever possible since they provide the strongest, and sometimes the only, evidence that change on the outcome of interest is likely due to the action of the activity. Technical piloting needs to be undertaken prior to national roll-out of an activity because that is the only opportunity to construct control groups against which to make comparisons.

A well-designed national roll-out can add substantially to the evidence of the activity’s effectiveness. For example, a staged roll-out of the activity, say across geographical areas, can allow the tracking of change in parallel across these areas.

**Recommendation 12:** That the implementation of activities be designed so as to allow the evaluation of the effectiveness of the activities.

## Examples of performance measurement tools and approaches

At Attachment J of the discussion paper, the Department of Finance has listed examples of performance measurement tools and approaches. In our view, the list is relatively light-on for tools and approaches that have a basis in evaluation design. The list should be extended to include:

* implementation design, e.g. staged roll-outs (see the report of the NAO)
* regression discontinuity designs
* technical piloting of activities using designed data collections
* randomized control trials
* cohort-sequential designs[[42]](#footnote-42)
* the clear specification of the chain of causation that is hypothesized to lead from funded activities to policy-related outcomes
* evaluation that is informed by the chain of causation and is directed at assessing progress on short-, medium- and long-term impacts of activities
* an evaluation plan that is established prior to roll-out and encompasses a whole-of-life approach to evaluating the success of the activity
* the need for rigorous data analysis of both quantitative and qualitative data

**Recommendation 13:** That the Department of Finance extend the list of performance measurement tools and approaches contained in Attachment J of the discussion paper.

# Proportionality

The impost of assessing the effectiveness of activities is considerable. The principle of proportionality should be used to decide which activities should be more thoroughly evaluated. In determining the issue of proportionality, it is usual to consider two factors: cost and risk. Opportunity should be added to the list.

* The greater the funding commitment the greater the need to establish effectiveness, and the greater the extent of funding which should be made available to evaluate the activity
* Higher-risk activities should be subject to more thorough evaluation
* Truly new policies or delivery approaches provide rare opportunities to learn something about what is efficacious; the greater the opportunities for learning what works, the greater the potential return on evaluation, and hence the greater the need to pursue evaluation.

**Recommendation 14:** That, in the context of the principle of proportionality, cost, risk and opportunity be the factors used to decide which activities should be more thoroughly evaluated.

# Nomenclature

There are a number of words used in the discussion paper that are either ambiguous, or used to mean something different from the meaning usually given to them at different levels of management within the Australian Public Service. The words *program*, *measure*, *outcome*, and *evaluation* are all problematic. It is not sufficient that the words be understood by accountable authorities;[[43]](#footnote-43) if the performance framework is to be a success then they must be understood by officers more generally, and particularly by those officers who are responsible for the management of activities.

One of the most serious sources of confusion relates to the use of the word ‘program’ (or programme) to refer to both PBS-level programs (e.g. Program 3.1—Higher Education Support) and administered items (e.g., the HEPPP). This lack of clarity in nomenclature appears to have contributed to confusion about the Outcomes and Programs Framework at the level of program managers (i.e., has contributed to confusion at the activity level) and to a failure of a clear line-of-sight between program management, monitoring and evaluation processes and entity performance-reporting. Program managers do not necessarily understand how their program fits into something else called a program—something else that isn’t at all like what they, the program manager, would call a program. This is a recipe for communication failure.

A smaller, but highly avoidable confusion lurks in the apparently-but-perhaps-not (!) interchangeable use of ‘activity’, ‘sub-program’ and ‘major activity’. Then there are ‘business activities’, ‘major business areas’. The point is probably clear. The rules need to be couched in a single, unambiguous set of terms which are used in a consistent manner throughout the document.

**Recommendation 15:** That the Department of Finance develop a consistent and unambiguous nomenclature which can be used throughout all communications regarding non-financial reporting and when talking about activities; that this nomenclature be maximally suited for communication with program managers.

# Summary of recommendations

1. That the Department of Finance divorce the reporting structure for non-financial outcomes from that used for financial reporting.
2. That the reporting structure for non-financial outcomes be tied to activities that have causal agency.
3. That the Department of Finance require that all proposals for new policy initiatives to describe explicitly the assumed causal chain between the activities that are to be funded and the anticipated outcomes.
4. That prior to implementation, program logic be used to decide what variables will index short-, medium- and long-term effectiveness of the activity. That funding be tied to the requirement that the proposed indices actually be measurable..
5. That the Department of Finance encourage entities to map the connections between the intermediate KPIs of the activities they administer and to use the map as the structure against which to report non-financial outcomes.
6. That evaluation within agencies be treated in a broadly parallel fashion to the way that audit is treated; that the reporting lines and decision-making powers related to evaluation and evaluation staff be divorced from those of the area that manages the program to be evaluated.
7. That evaluation tasks and roles *not* be amalgamated into the responsibilities of existing audit staff.
8. That entities be encouraged to assign staff with evaluation knowledge and skills to perform internal evaluation functions.
9. That the Department of Finance require all evaluations, prior to being undertaken, to be entered on a public register maintained by a central agency.
10. That the Department of Finance require all evaluation reports to be published.
11. That the evaluation of activity effectiveness be planned prior to implementation and be integrated into the non-financial performance reporting of entities.
12. That the implementation of activities be designed so as to allow the evaluation of the effectiveness of the activities.
13. That the Department of Finance extend the list of performance measurement tools and approaches contained in Attachment J of the discussion paper.
14. That, in the context of the principle of proportionality, cost, risk and opportunity be the factors used to decide which activities should be more thoroughly evaluated.
15. That the Department of Finance develop a consistent and unambiguous nomenclature which can be used throughout all communications regarding non-financial reporting and when talking about activities; that this nomenclature be maximally suited for communication with program managers

# Additional references

Cartwright, N., & Hardie, J. (2012) *Evidence-based policy: A practical guide to doing it better*. New York, NY: Oxford University Press.

Bovaird, T. (2014). Attributing Outcomes to Social Policy Interventions – ‘Gold Standard’ or ‘Fool's Gold’ in Public Policy and Management? Social Policy & Administration, 48(1), 1–23.

Pawson, R. (2013). *The Science of Evaluation: A Realist Manifesto*. London: Sage.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. Journal of Financial Economics, 3(4), 305–360.

# About the authors

Dr Angela O’Brien-Malone has extensive experience as a senior academic with specialist expertise in evaluation and research methodology. She has made numerous research-based submissions to government including to the Inquiry into the Reform of Australian Government Administration. She is currently employed within the APS.

Dr Mark Diamond is an evaluation expert with a long record of publication in peer reviewed journals on matters related to evaluation and research. He has been a research and evaluation consultant to the private sector and to government. He is currently employed within the APS.

1. [KPMG](#KPMG) (2009). [↑](#footnote-ref-1)
2. “Science is a method to keep yourself from kidding yourself” — Edwin Land [↑](#footnote-ref-2)
3. For a detailed discussion of the distinction between advocacy research and scientific research, see [Firebaugh](#Firebaugh) (2008). [↑](#footnote-ref-3)
4. One well-known example is the “Scared Straight” programs run with adolescents whose delinquent behaviour places them at risk of ending up in juvenile detention. The Scared Straight approach is based on the idea that if these juvenile delinquents visit a juvenile prison of some kind it will make them aware of how unpleasant it is and scare them out of their delinquent behaviours. Sounds like common sense. But juveniles who are put through these programs show a pattern of an *increase* in reoffending ([Petrosino et al](#Petrosino)., 2002). Far from helping, the programs make things worse. [↑](#footnote-ref-4)
5. As Brian Head from Queensland University has noted: “Relatively few research and consulting projects are commissioned without some expectation that the reports may assist in upholding a certain viewpoint.” ([Head](#Head), 2008, p. 1) [↑](#footnote-ref-5)
6. [Leigh](#Leigh) (17–18 August, 2009). [↑](#footnote-ref-6)
7. [Ericsson, Krampe & Tesch-Römer](#Ericsson) (1993). [↑](#footnote-ref-7)
8. “A prophet is not without honour except … in his own house,” (Matthew 13:57, Revised Standard Version). Curiously, the opposite problem—[Not Invented Here](http://en.wikipedia.org/wiki/Not_Invented_Here)—besets many private sector organizations. [↑](#footnote-ref-8)
9. Publication is a critical step because both the *process* and *fact of* publication opens the research to peer review and comment. [↑](#footnote-ref-9)
10. Our choice of EL1 is motivated by the fact that EL2 and SES members in almost all departments are employed as generalists, not experts. [↑](#footnote-ref-10)
11. See, for example, [Murdoch University](#Murdoch) (2006). [↑](#footnote-ref-11)
12. The Reserve Bank Service, established under Part VII of the Reserve Bank Act (1959), is the employer of staff at the Reserve Bank of Australia. [↑](#footnote-ref-12)
13. The Reserve Bank Service utilises a variety of similar approaches ([RBA](#RBA), 2008). [↑](#footnote-ref-13)
14. [www.tenders.gov.au](http://www.tenders.gov.au/) [↑](#footnote-ref-14)
15. [www.tenders.gov.au](http://www.tenders.gov.au/) [↑](#footnote-ref-15)
16. The Reserve Bank Service, established under Part VII of the Reserve Bank Act (1959), is the employer of staff at the Reserve Bank of Australia. [↑](#footnote-ref-16)
17. *Program* is used in the context of Portfolio Budget Statements to mean a financial aggregation of administered items but elsewhere refers to such things as the Higher Education Participation and Partnerships Program which is an administered item. [↑](#footnote-ref-17)
18. <[www.education.gov.au/portfolio-budget-statements-2009-10](http://www.education.gov.au/portfolio-budget-statements-2009-10)> [↑](#footnote-ref-18)
19. The items included a Superannuation Program to provide assistance to certain higher education institutions to meet expenses associated with unfunded superannuation schemes [↑](#footnote-ref-19)
20. Reviews of PBS-level programs can give useful financial information, as exemplified by the Strategic Review of Indigenous Expenditure (see [www.finance.gov.au/foi/disclosure-log/2011/foi\_10-27\_strategic\_reviews.html](http://www.finance.gov.au/foi/disclosure-log/2011/foi_10-27_strategic_reviews.html)) but cannot indicate what policy-related changes are attributable to what funded activity. [↑](#footnote-ref-20)
21. The critical role that argument, and a causal theory, play in the evaluation of non-financial outcomes is highlighted by Pawson (2013) and Cartwright and Hardie (2012), amongst others. [↑](#footnote-ref-21)
22. See for example, Achinstein, P (2001) *The Book of Evidence*, Oxford: Oxford University Press; but see also, Cartwright, N., Goldfinch, A., & Howick, J. (2009). Evidence-based policy: Where is our theory of evidence? *Journal of Children's Services*, 4(4), 6–14. doi: [10.5042/jcs.2010.0017](http://dx.doi.org/10.5042/jcs.2010.0017). [↑](#footnote-ref-22)
23. The fact that the causal theories for activities are rarely developed prior to implementation is one of the greatest hurdles facing attempts to develop useful non-financial reporting. [↑](#footnote-ref-23)
24. Establishing complete casual claims is not a trivial undertaking, see Rol, M., & Cartwright, N. (2012). Warranting the use of causal claims: a non-trivial case for interdisciplinarity. *Theoria*, 74, 189–202. doi: [10.1387/theoria.4075](http://dx.doi.org/10.1387/theoria.4075). [↑](#footnote-ref-24)
25. See the Wikipedia entry for [Logic Model](http://en.wikipedia.org/wiki/Logic_model) for one explanation of program logic. [↑](#footnote-ref-25)
26. National Audit Office. (2013). Evaluation in government. Available at: <http://www.nao.org.uk/wp-content/uploads/2013/12/10331-001-Evaluation-in-government_NEW.pdf> [↑](#footnote-ref-26)
27. E.g., <http://education.unimelb.edu.au/cpe> [↑](#footnote-ref-27)
28. <http://www.act.ipaa.org.au/CEF-Canberra-Evaluation-Forum/cef-canberra-evaluation-forum.html> [↑](#footnote-ref-28)
29. <http://aes.asn.au/professional-learning/international-conference/previous-conferences.html> [↑](#footnote-ref-29)
30. e.g., in the academic literature relating to management, behavioural economics and psychology. [↑](#footnote-ref-30)
31. Incentives include public recognition, money, job security, more time with the kids, statutory imperatives such as those contained in ss 10, 13 and 35(c) of the *Public Service Act*, and promotion opportunities. [↑](#footnote-ref-31)
32. The Australian Public Service has long recognised this and seeks to influence the behaviour of officers of Commonwealth entities by paying performance related bonuses to more senior members of the service. Good performance, in this context, equates with behaviour that serves the interests of government. [↑](#footnote-ref-32)
33. Audit, as distinct from evaluation, has a respected history of such incentives. Auditors are rewarded not only for adhering to audit standards but for uncovering instances of malfeasance. [↑](#footnote-ref-33)
34. ss 10 and 13 of the Public Service Act 1999, notwithstanding. [↑](#footnote-ref-34)
35. Hallsworth, M., Parker, S., & Rutter, J. (2011). *Policy Making in the Real World: Evidence and Analysis: Institute for Government*. Retrieved from <http://tinyurl.com/PolicyMakingInTheRealWold> on 17 July 2014. [↑](#footnote-ref-35)
36. Genuine transparency in the procurement and contracting of evaluation services can serve as a very powerful social incentive to promote alignment. Regrettably, side-channel communication between parties to an evaluation procurement can subvert transparency; see <http://en.wikipedia.org/wiki/Covert_channel>. [↑](#footnote-ref-36)
37. i.e., as long as it does not recognise that managers can be placed in the unfortunate position of managing policy-related activities which, through no fault of the manager, are misconceived or simply ineffective [↑](#footnote-ref-37)
38. Diamond, M. R. & O’Brien-Malone, A. (2010). *Evaluation: a neglected but vital reform*. Submission to the Senate Standing Committee on Finance and Public. Canberra: Senate, Parliament House. [[web-link](https://senate.aph.gov.au/submissions/comittees/viewdocument.aspx?id=0240c19c-a1a1-45c1-8b0f-ce335f69d0c6)] [↑](#footnote-ref-38)
39. There have been approximately 150 years of research into the methods that allow us to discover the impact of interventions designed to affect human behaviour. [↑](#footnote-ref-39)
40. See Simes, R. J. (1986) [PubMed-ID [3760920](http://www.ncbi.nlm.nih.gov/pubmed/3760920)]. The United States government register of clinical trials can be found at <http://www.clinicaltrials.gov> [↑](#footnote-ref-40)
41. We have used the term ‘technical’ piloting here to distinguish these designed data collections from simply ‘trying out’ aspects of the intervention, e.g., the administrative approach or timetable. [↑](#footnote-ref-41)
42. See for example, Prinzie, P., & Onghena, P. (2005). Cohort Sequential Design. In B. Everitt, B. S. & D. C. Howell, *Encyclopedia of Statistics in Behavioral Science*: John Wiley & Sons, Ltd. doi: [10.1002/0470013192.bsa110](http://dx.doi.org/10.1002/0470013192.bsa110) [↑](#footnote-ref-42)
43. The meaning of *aaccountable authority* is defined in s 12 of the *PGPA Act*. [↑](#footnote-ref-43)