The Australian Research Quality Framework: A Live Experiment in Capturing the Social, Economic, Environmental and Cultural Returns of Publicly Funded Research

Claire Donovan

Research funding organizations and science policy circles use the term 'research quality' to describe the measurable influence of academic research on the academic community. 'Research impact' denotes the benefits or returns from research which flow beyond the academic realm to 'end users' of research. These 'end users' are traditionally defined as industry, business, government, or, more broadly, the taxpayer.

As Donovan (forthcoming 2007) explains, indicators of research 'quality', such as research income and citation measures, have become part of the fabric of research evaluation; and in recent years, there has been growing interest in similarly evaluating research 'impact'. This has been spurred by the desire of governments to gauge the value of publicly funded research to 'end users' beyond academia. The reasons for accounting for impact vary: to justify expenditure on academic research in terms of its return on taxpayers' investment, or the creation of public value for society; to redirect national science foresight planning towards 'relevant' research; to enhance international industrial and economic competitiveness; and (in tandem with 'quality' assessment) to inform the performance-based distribution of block funding to universities.

Distinct phases in the evolution of impact evaluation have been noted (Donovan, 2007; Donovan forthcoming 2007; Martin, 2007):

Technometrics. The initial search for reliable quantitative measures sought to collate data on investment from industry,

commercialisation and technology transfer. However, these data were found to represent low-order levels of impact which did not extend to broader economic or societal benefits, and marked an unsophisticated approach to impact measurement confined to science, technology, engineering and medicine, and the concerns of industry and business (Donovan, 2006a).

Sociometrics. A second phase of impact evaluation sought more socially relevant measures in the form of 'sociometrics', which attempted to map research outcomes onto existing government social statistics (The Allen Consulting Group, 2005). Yet these impact indicators presented no credible link between academic research and macro-level social trends, and overlooked the cultural import of research.

Case studies. A third wave of impact evaluations acknowledged that quantification may conceal more than it reveals. Typically employed by dedicated research funding organisations to assess the outcomes of specific funding initiatives, these evaluations proceeded on a case study basis, and sought to combine quality and impact measurements using both quantitative, and qualitative or deliberative, approaches. The case studies generally attempted to gauge a more broadly conceived notion of impact which probed various dimensions of the economic, social and environmental returns from research. This approach demonstrates sensitivity to the definition of 'impact', which will vary according to the perspectives of different 'end-users' such as government, citizens, consumers, business, industry, community groups, NGOs and practitioners. These different perceptions affect what is valued and hence measured, and so case study methodology includes several impact dimensions, and encourages 'end-user' participation throughout the evaluation process (Wooding, Hanney, Buxton, and Grant, 2004; Spaapen, Dijstelbloem, and Wamelink, 2007).

Yet, this sophisticated case study approach has been largely confined to evaluating the impact of scientific and medical research, and has not been adapted to the assessment of a nation's whole research base.

The international context. In terms of national research assessment exercises, to date the most developed examples of 'impact' evaluation occur in the Netherlands and New Zealand. These evaluations sit alongside 'quality' assessment; and primarily focus on the economic value of publicly funded research, while measures of broader 'user engagement' are bound to low-level input and output indicators rather than tangible societal benefits. Yet the Netherlands seeks data on the influence of research on developments or questions in society at large, and New Zealand collects brief contextual descriptions of linkages beyond academia.

These innovations resonate with current concerns in the international research evaluation community, which has come to recognise the limited value of impact assessment tied exclusively to economic and quantitative concerns; and the latest movement in impact evaluation is towards developing richer qualitative and contextual approaches at the national level (Donovan, 2007; FWF/ESF, 2007, p. 45). In this vein, the prospective 2008 Australian Research Quality Framework (RQF) is the first national research assessment exercise to include a truly comprehensive and methodologically diverse 'impact' audit.

A Brief History of RQF Development

Australia's RQF came into being as a hybrid solution to academic concerns about research 'quality' and government interest in research 'impact'. The Australian academic community wanted the government to allocate university block funding on the basis of discipline-based

peer review of research 'quality' rather than the extant metrics-based formula (DEST, 2004b). The government wanted to boost Australia's innovation strategy through linking academic research to the concerns of industry and business, particularly in the context of broader economic, social and environmental benefits to society (DEST, 2004a). The RQF was proposed as a panel-based exercise to evaluate both research excellence and the wider benefits of academic research for the nation, and to allocate funds on the basis of outcomes. It is unsurprising, given the 'quality' push from academia and the 'impact' pull from government, that the RQF philosophy of impact evaluation has been contested and reshaped throughout its development and implementation.

There have been several phases of RQF development involving various advisory groups, technical working groups, and much consultation with the Australian higher education sector. In December 2004, Brendan Nelson, the Minister for Science, Education and Training, appointed an Expert Advisory Group (EAG) which launched a consultative exercise to determine the structure and features of the RQF (DEST, 2005a, p. 3). The EAG published its preferred RQF model in September 2005 (DEST, 2005c), and gave its final advice in December 2005 (DEST, 2005d).

A new Minister, Julie Bishop, took office, and in March 2006 a new Development Advisory Group (DAG) was created, chaired by the Chief Scientist, and tasked to refine the RQF model and detail its phases of implementation. In June 2006, the Minister announced that the first RQF would take place in 2008, and the DAG appointed several technical working groups to address in detail various RQF features in need of further development. This included a Technical Working Group on Research Impact, which reported its findings to the DAG during August 2006.

The Technical Working Group on Research Impact. The Impact Working Group comprised of senior academics, senior university managers, representatives from business and industry, experts in impact evaluation, and several DAG members. The membership also represented academic interests in science, technology, engineering, medicine, commerce, humanities, creative and performing arts, and social science. Its remit was to provide detailed advice to the DAG in the following areas:

Methodology. Recommend the optimal methodology to assess the impact of Australia's universities.

Indicators. Develop generic and discipline-specific quantitative and qualitative measures of research impact.

Assessment period. Establish the appropriate length of the assessment period required for effectively assessing research impact.

Evidence portfolios. Determine the necessary evidence for research groups to demonstrate impact, including: composition of impact statements; metrics to be provided in context statements; if 'four best outputs per researcher' are adequate to demonstrate research quality and impact.

Demonstrating impact. Advise how research groups are to demonstrate research impact; and how ratings of research impact are most effectively reported.

Verifying impact. Propose appropriate processes for assessment panels to evaluate research impact.

Various RQF features were fixed, and the Impact Working Group had to navigate around these. For example, the EAG had defined research impact as the "social, cultural, economic and/or environmental outcomes for industry, government and/or other identified communities regionally within Australia, nationally and/or internationally" (DEST, 2005d: 24). Other key characteristics include

the RQF being a panel-based peer and end-user review of the 'quality' and 'impact' of Australian university research. There are 13 panels which are clusters of disciplines that share similar assessment profiles (e.g. Physical, Chemical and Earth Sciences; Engineering and Technology; Social Science and Politics; Law, Education and Professional Practices; Humanities; Creative Arts, Design and Built Environment). The assessment will be conducted at the 'research group level', rather than at the individual level (as in the case of the New Zealand Performance Based Research Fund) or the discipline level (like the UK Research Assessment Exercise). The 'quality' assessment consists of panel judgements combining a peer review of the four 'best' outputs of per researcher with quality 'metrics' applied to research groups. In terms of impact assessment, set features are an 'impact scale' against which to report and judge levels of research impact; and research groups are to submit an 'impact statement' linking the group's research to claimed impact outcomes, the beneficiaries, the measurable difference made by the research, and the details of end-users who may confirm research groups' impact claims (DEST, 2005d: 20).

Key recommendations. The Impact Working Group met four times between June and August 2006. During this period the DAG provided feedback through its members within the Group, and via the Department of Education, Science and Training (DEST), sometimes suggesting that advice be modified – a demonstration of the academic 'push' and government 'pull' in action. At the request of the DAG, the Impact Group produced a short outline of its advice, highlighting changes or refinements to EAG recommendations. This was made public as a DAG "Guiding Principles" document in August 2006 (DEST, 2006a). The Impact Working Group presented its final report to the DAG in August 2006 (TWGRI, 2006); and the DAG published a revised version of this

advice in September 2006 (DEST, 2006b), and its final recommended RQF model in October 2006 (DEST, 2006c). The Impact Working Group's recommendations are summarised below, along with noteworthy deviations from EAG and DAG thinking:

Methodology. The optimum assessment methodology is a qualitative and contextual approach, mediated through the judgement of academic peers and end-users. Information is best derived from context statements, impact statements, case studies, and, where appropriate, relevant quantitative and qualitative indicators (TWGRI, 2006: 1).

Indicators. Quantitative metrics are underdeveloped and cannot be used as a proxy for determining impact ratings for research groups; but where appropriate, some qualitative and quantitative indicators may support impact claims (TWGRI, 2006: 1). The DAG decided that assessment panels would, nonetheless, be given generic impact indicators, and be asked to determine additional cluster-specific ones (DEST, 2006c: 18).

Assessment period. The EAG chose a six year assessment window (2000 to 2006) for quality and impact assessment, and that the impact to be assessed must be related to research conducted within that same six year period. The Impact Working Group proposed that while the research impact assessed should occur within the six year window, this may be derived from original research conducted earlier (TWGRI, 2006: 1); and assessment panels use their judgement to determine a reasonable timeframe from the original research to the impact claimed. These recommendations were endorsed by the DAG (DEST, 2006: 1-2). The Working Group believed cases where the original research is older than fifteen years will require additional supporting evidence (TWGRI, 2006: 7). The DAG limited the period for older research to an additional six years only (DEST, 2006c: 21)

Evidence portfolios. a) Impact statements should be the basis of assessing research impact. These should be evidence based, no more then 10 pages in length, and consist of: a statement of claims against impact criteria (see "*Demonstrating* impact" below), up to four case studies illustrating examples of impact, details of end-users who can verify the impact claims (DEST, 2006a: 1-2). b) No metrics are to be provided in the context statement, but may be used to support claims made in a research group's impact statement. c) The EAG proposed the same four 'best' outputs per researcher be used to assess both quality and impact claims. The Impact Working Group recommended that impact assessment should draw on a group's complete body of work, including non-traditional outputs such as reports to government (TWGRI, 2006: 1), a revision to the RQF model supported by the DAG (DEST, 2006b: 6).

Demonstrating impact. a) Research impact is best demonstrated by linking a group's impact claims to criteria set out in the 'impact rating scale'. Evidence should connect the group's original research to impact ratings (TWGRI, 2006:7). The Impact Working Group recommended clear guidelines be developed at the discipline level, a proposal endorsed by the DAG (DEST 2006, 2). b) Connecting impact claims to the 'impact rating scale' is the most effective way to report claims of research impact. The EAG recommended a simple three-point scale demonstrating degrees of public benefits derived from research (DEST, 2005d: 24). This preference was shared by the Impact Working Group, but the DAG directed it to develop a five-point scale with more attention to engagement with 'end-users'. The Impact Working Group's final scale was a blend of end-user interaction and public benefit, initially endorsed by the DAG (DEST, 2006a: 1), but later modified to reflect more commercial and industrial concerns (DEST, 2006b, 7). (See "Accounting for impact" below for a discussion of the contested nature of the impact scale).

Verifying impact. Assessment panels will review research groups' evidence portfolios, and will apply their collective expert judgement to determine the validity of the claims made against the impact criteria. Impact ratings will be assigned, and the rating process will be moderated between discipline panels to ensure consistency and fair treatment for multi-disciplinary research. The Working Group recommended 'Payback' consensus scoring model particularly suited for this purpose (TWGRI, 2006: 1: Wooding *et al.*, 2004).

Contested Themes in the RQF Philosophy of Impact Evaluation

> *"It is my view that if we are able to get right the measure of impact – in both its form and its recognition – then we will have created a research evaluation measure that will greatly surpass those of other nations." – Hon. Julie Bishop, Minister for Education, Science and Training (DEST, 2006d)*

The role of impact evaluation in the RQF came with high expectations from government. As the Minister elaborates, "It will ensure that not only do we, as a country, reward high quality research, but also we reward research which makes a demonstrable change to the way we live or enjoy our lives." However, the RQF philosophy of impact assessment has, at times, resembled the 'Pushmi-pullyu' of Dr. Doolittle fame: a two-headed llama which tries to travel in opposite directions. The government 'pull' towards impact is offset by a 'push' towards more scholarly concerns; and this 'pull' is sometimes forcefully directed towards the interests of industry and commerce, yet counterbalanced by an equally strong 'push' towards broader public benefits. It is within this context that the chapter now turns to examine central concepts in impact evaluation which display these inherent tensions:

defining research impact, communicating research beyond academia, and accounting for research impact.

Defining research impact. For RQF purposes, 'impact' was originally concerned with social, economic and environmental effects, reflecting a trend towards 'triple bottom line' accounting (see Donovan, forthcoming 2007). The EAG's consultation with the higher education sector led to introducing the 'cultural' as a fourth impact domain, and this 'quadruple bottom line' is unique in international impact assessment terms. When we turn to consider what, precisely, impact denotes, there are contradictory messages contained in the RQF deliberations which reflect a fragile balance of 'push-pull' interests.

Only the Impact Working Group supplies actual content for the four impact domains (TWGRI, 2006: 4; DEST, 2006b, 4), which is dropped by the DAG, but reintroduced in the 2007 submission specifications. Impact is described as adding to the social, economic, natural and cultural capital of the nation:

Social benefit. "Improving quality of life; stimulating new approaches to social issues; changes in community attitudes, and influence upon developments or questions in society at large; informed public debate and improved policy-making; enhancing the knowledge and understanding of the nation; improved equity; and improvements in health, safety and security."

Economic benefit. "Improved productivity; adding to economic growth and wealth creation; enhancing the skills base; increased employment; reduced costs; increased innovation capability and global competitiveness; improvements in service delivery; and unquantified economic returns resulting from social and public policy adjustments."

Environmental benefit. "Improvements in environment and lifestyle; reduced waste and pollution; improved management of natural

resources; reduced consumption of fossil fuels; uptake of recycling techniques; reduced environmental risk; preservation initiatives; conservation of biodiversity; enhancement of ecosystem services; improved plant and animal varieties; and adaptation to climate change."

Cultural benefit. "Supporting greater understanding of where we have come from, and who and what we are as a nation and society; understanding how we relate to other societies and cultures; stimulating creativity within the community; contributing to cultural preservation and enrichment; and bringing new ideas and new modes of experience to the nation."

We find that the Impact Working Group and early DAG documents define impact in terms of public benefit within these domains (DEST, 2006a: 1; TWGRI, 2006: 2; DEST, 2006b: 3). On the other hand, the EAG and the DAG's recommended RQF are concerned with direct practical utility and more targeted groups of end-users: for example, impact is interchanged with the word "usefulness" (DEST, 2005b, 11; 24), and is "the recognition by qualified end-users that quality research has been successfully applied to achieve social, cultural, economic and/or environmental outcomes." (2005d, p. 12; DEST, 2006c: 10), and is found in "short-term...outcomes for industry, government and/or other identified communities" (DEST, 2005b, 24). We shall see that when considering impact domains and impact rating scales, these divergent views entail mixed messages about what constitutes legitimate impact, and how this may be measured and verified.

Finally, when looking at how 'impact' is defined, it is important to note what is excluded. a) The RQF immediately rejected the notion of impact as 'knowledge transfer', for example the commercialisation of other people's ideas (DEST, 2005c: 7). In this respect, 'impact' is limited to a research group's own original

research. b) a research group may apply for exclusion from impact assessment if its research is at an early stage of development, or if its research orientation means it would be inappropriate to be assessed in terms of impact (DEST, 2005d: 25; DEST 2006b, 5; DEST, 2006c: 22). c) Contrary to European developments (FWF/ESF, 2007), 'basic' research is exempted from impact assessment on the grounds that it is not devalued (DEST 2005d, 25; TWGRI, 2006: 2; DEST, 2006b: 3):

> "the fundamental research of today may yield the research impact of the future. In this respect, impact assessment must allow for progress from initial research through to eventual impact, and acknowledge that this is not a necessarily linear process, and that this development takes time." (TWGRI, 2006: 2)

However, the 'pull' of this sentiment is at odds with the DAG's decision to only allow an additional six year window to connect original research to impact: a counter-intuitive short-term 'push' that devalues basic research through excluding many significant and enduring research impacts.

Communicating research beyond academia. During RQF development there were differing views on what form of 'publication' should be used to link a group's original research to its impact claims. The EAG had recommended that the same 'best' four outputs per researcher be used for both quality and impact assessment (DEST, 2005d). However, this failed to recognise that vehicles for communication differ for academic and non-academic audiences. This also led to concerns that a linear ideal of scientific discovery underpinned the impact assessment model: that a group of scientists publish a journal article, the idea is taken up and developed, and impact for society is

then accrued in terms of technical or health benefits, for example. The Impact Working Group argued that "the types of research output one would submit to demonstrate quality and impact are often quite different because these publications are tailored for different audiences". It recommended that non-traditional outputs such as reports for government, public exhibitions, and media broadcasts were an essential link between original research and engagement with 'end-users', and so should be separately drawn upon for impact assessment. It also argued that impact which occurs within the six year assessment period is likely to be connected to traditional and non-traditional research outputs produced before the six year window for quality assessment, and thus the window for impact should be extended (TWGRI, 2006: 6). These sentiments were endorsed by the DAG (DEST, 2006b: 6; DEST, 2006c: 10).

In this instance the 'push' was led by the Impact Working Group's search for the optimum methodology for impact assessment, which polarised the university sector as this preference was supported by technical universities, and opposed by the 'pull' of elite academic institutions for an RQF which gave primacy to the peer review of 'high quality' publications.

Accounting for research impact. As has been noted, divergent views of what impact is entail different views of what should be measured and how. The 'push' to impact as industrial and commercial advance finds its ultimate expression in quantitative metrics tied to investment from business and industry, patents, and commercialisation; the 'pull' towards public value seeks to make previously intangible public benefits of research visible by employing a contextual approach, informed by qualitative and quantitative evidence, and judged by academic peers and end-users. The latter position was supported by the Impact Working Group, drawing on international best

practice in impact evaluation, and strongly favouring a case study approach to methodology (see Donovan, forthcoming 2007). However, we can easily imagine both approaches adopting a panel system informed by evidence supporting a scale of impact claims against the four impact domains presented in "defining research impact" above, albeit in a largely quantitative or more contextual manner (potential examples are provided in TWGRI, 2006: 9). Hence the principle of the case study approach was endorsed by the DAG, as indeed was the continued (and in vain) search for robust metrics of high-order impact (DEST, 2006d).

The impact scale is perhaps the most hotly contested aspect of RQF impact evaluation. As has been noted, during RQF deliberations the impact scale morphed from a simple three-point measure of degrees of wider benefit to a fine-grained five-point scale geared to end-user interaction. The actual RQF scale is presented in "Australia's Live Experiment" below, and matches the final DAG preference (DEST, 2006d: 22). It is a linear, progressive scale, premised upon a route to impact which begins with (1) engagement with end-users who recognise the importance of the research to a defined area, (2) the adoption of research, (3) adoption producing benefits for end-users, and (4) the magnitude of the benefit derived from the adoption. The Impact Working Group's alternative scale was non-linear, and preferred (1) reciprocal engagement with end-users, (2) significant uptake of research by the end-user community, (3) producing significant added-value or improvements, and (4) producing transformational benefits on a large scale. It was felt that while the language of 'adoption' was suited to an idealised model of practice in engineering with industry as the end-user, it alienated the humanities, arts, and social sciences. The 'pull' was towards a more inclusive scale which would embrace all disciplines, and the diffuse manner in which research has value beyond academia; the 'push' was concerned with targeted end-user

engagement and driving behaviours which would make Australia's science base more efficient.

Australia's Live Experiment

RQF development continued throughout 2007, including a series of discipline workshops which each devoted half a day to research impact, further sector consultations, RQF trials including testing mechanisms for assessing research impact, and the development of generic specifications and panel-specific guidance, which were released in September 2007 (DEST, 2007).

The generic specifications display a great deal of 'push' in that a repeated catchphrase is the usefulness of research for "government, industry, business and the wider community". The definition of impact is extended: "Impact refers to the extent to which research has led successfully to social, economic, environmental and/or cultural benefits for the wider community, or an element of the community" (DEST, 2007: 5), which allows for the inclusion of private value in addition to public value. In terms of defining research impact, there is an explicit request that research groups should include in their impact statements "identifiable and supportable impact-related indicators. This requires the impact statement to identify the beneficiaries of the research and the way in which they have benefited" (DEST, 2007: 30). The 'push' also dominates in the flavour of examples of impact given: "improved quality of products/services, cost-effectiveness, customer satisfaction, lives saved or productivity" (DEST, 2007: 33); "Policy impacts can also include changes to policies of corporations, councils, professional groups and non-government organisations" (DEST, 2007: 33); and a series of examples are given to illustrate outcomes which would match

impact ratings from D to A, which are restricted to industry or clinical psychology.

Exhibit 1: The Impact Rating Scale (DEST, 2007: 31)

Rating	**Description**
A	Adoption of the research has produced an outstanding social, economic, environmental and/or cultural benefit for the wider community, regionally within Australia, nationally or internationally.
B	Adoption of the research has produced a significant social, economic, environmental and/or cultural benefit for the wider community, regionally within Australia, nationally or internationally.
C	Research has been adopted to produce new policies, products, attitudes, behaviours and/or outlooks in the end-user community.
D	Research has engaged with the end-user community to address a social, economic, environmental and/or cultural issue, regionally within Australia, nationally or internationally.
E	Research has had limited or no identifiable social, economic, environmental and/or cultural outcome, regionally within Australia, nationally or internationally.

The panel-specific guidance, however, does provide tailored examples of engagement, uptake of research, and extent of benefit. Yet no examples of impact metrics are offered. In this sense, the RQF remains a 'Pushmi-pullyu', with the 'push' at the grand policy level, and the 'pull' at the research group and panel level, leaving scope

for contextual interpretation of the impact scale in discipline-specific terms.

Despite the publication of the generic specifications and panel-specific guidance, we find, therefore, that impact measurement in the RQF remains a 'live experiment' as (1) its fine detail continues to be refined at the panel level, although this lack of transparency is of vital concern for research groups which need guidance in effectively constructing their impact statements, and (2) the balance of quantitative indicators versus contextual evidence to inform the second RQF remains under review. The RQF is also a 'live experiment' as we are unsure of its future: there is a general election due in Australia, and the Labor Party, which was ahead in the polls the day the RQF specifications were released, has vowed that if it replaces the current Liberal coalition government it will abandon impact assessment. There are suggestions that the RQF may take place in 2009 rather than 2008, or that impact measurement should be a 'shadow exercise' in the RQF's first iteration.

We find that the RQF approach to impact evaluation is a world first as other countries have tended to focus on economic returns, or rely upon quantitative rather than contextual approaches to impact assessment. The consequence has been that impact measurements have proven unsatisfactory, largely because the public value of research has not been adequately addressed. The RQF has certainly gone a long way towards developing an optimal methodology for capturing the social, economic, environmental and cultural returns of publicly funded research.

The 'Pushmi-pullyu' aspect of implementing a pluralistic impact evaluation may be part of an inevitable compromise of government and academic interests. However, this runs the danger of presenting mixed messages about what, precisely, research impact is, and how best to

account for it within a national research assessment exercise. We wait to see if this 'live experiment' will come to fruition.

References

Austrian Science Fund [FWF]/European Science Foundation [ESF]. "Conference Report. Science Impact: Rethinking the Impact of Basic Research on Society and the Economy." Vienna: FWF/ESF, 2007. Retrieved September 18, 2007, from http://www.science-impact.ac.at/index.html.

Bishop, J. "Keynote Address." Paper presented at the Knowledge Transfer and Engagement Forum, Sydney, June 16 2006a. Retrieved September 18, 2007, from http://www.dest.gov.au/Ministers/Media/Bishop/2006/06/b001160606.asp.

Department of Education, Science and Training [DEST]. *Research Quality Framework: Assessing the Quality and Impact of Research in Australia. RQF Submission Specifications.* Canberra: Commonwealth of Australia, 2007.

Department of Education, Science and Training [DEST]. *RQF Guiding Principles: Research Impact. Report by the RQF Development Advisory Group.* Canberra: Commonwealth of Australia, 2006a. Retrieved September 18, 2007, from http://www.dest.gov.au/NR/rdonlyres/B44A08DF-9C17-4616-800D-7D34EFC14DA2/13524/ResearchImpact.pdf.

Department of Education, Science and Training [DEST]. *Research Quality Framework: Assessing the Quality and Impact of Research in Australia.*

Research Impact. Report by the RQF Development Advisory Group. Canberra: Commonwealth of Australia, 2006b.

Department of Education, Science and Training [DEST]. *Research Quality Framework: Assessing the Quality and Impact of Research in Australia. The Recommended RQF. Report by the RQF Development Advisory Group.* Canberra: Commonwealth of Australia, 2006c.

Department of Education, Science and Training [DEST]. "Australian Government Endorses Research Quality Framework." Media Release, November 14. Canberra: Commonwealth of Australia, 2006d. Retrieved September 18, 2007, from http://www.dest.gov.au/Ministers/Media/Bishop/2006/11/B002141106.asp.

Department of Education, Science and Training [DEST]. *Research Quality Framework: Assessing the Quality and Impact of Research in Australia. Issues Paper. RQF Expert Advisory Group.* Canberra: Commonwealth of Australia, 2005a.

Department of Education, Science and Training [DEST]. *Research Quality Framework: Assessing the Quality and Impact of Research in Australia. Advanced Approaches Paper. Report by the RQF Expert Advisory Group.* Canberra: Commonwealth of Australia, 2005b.

Department of Education, Science and Training [DEST]. *Research Quality Framework: Assessing the Quality and Impact of Research in Australia. The Preferred Model. Report by the RQF Expert Advisory Group.* Canberra: Commonwealth of Australia, 2005c.

Department of Education, Science and Training [DEST]. *Research Quality Framework: Assessing the Quality and Impact of Research in Australia. Final Advice on the Preferred RQF Model. Report by the RQF Expert Advisory Group.* Canberra: Commonwealth of Australia, 2005d.

Department of Education, Science and Training [DEST]. *Backing Australia's Ability: Building Our Future Through Science and Innovation*. Canberra: Commonwealth of Australia, 2004a.

Department of Education, Science and Training [DEST]. *Evaluation of Knowledge and Innovation Reforms Consultation Report*. Canberra: Commonwealth of Australia, 2004b.

Donovan, C. "The Qualitative Future of Research Evaluation." *Science and Public Policy*, forthcoming 2007.

Donovan, C. "Accounting for the Triple Bottom Line: A Robust Qualitative Measure of the Public Value of Research." Paper presented at the Austrian Science Fund/European Science Foundation Conference on Science Impact: Rethinking the Impact of Basic Research on Society and the Economy, Vienna, May 11 2007. Retrieved September 18, 2007, from http://www.science-impact.ac.at/documentation/pdf/Session_C_Donovan.pdf.

Donovan, C. "An Instrument too Blunt to Judge Sharp Minds." *Times Higher Education Supplement*, 11 August 2006a, p. 14.

Martin, B. "Assessing the Impact of Basic Research on Society and the Economy." Paper presented at the Austrian Science Fund/European Science Foundation Conference on Science Impact: Rethinking the Impact

of Basic Research on Society and the Economy, Vienna, May 11 2007. Retrieved September 18, 2007, from http://www.science-impact.ac.at/documentation/pdf/Session_C_Martin.pdf.

Research Quality Framework Technical Working Group on Research Impact [TWGRI]. *Final Report: Optimal Methodology for Assessing Research Impact. August 2006.* Canberra: TWGRI, 2006.

Spaapen, J., Dijstelbloem, H, and Wamelink, F. *Evaluating Research in Context: A Method for Comprehensive Assessment*. The Hague, Netherlands: COS, 2007.

The Allen Consulting Group. "Measuring the Impact of Publicly Funded Research." Report to the Australian Government Department of Education, Science and Training. Canberra: The Allen Consulting Group, 2005.

Wooding, S., Hanney, S., Buxton, M. and Grant, J. *The Returns from Arthritis Research Volume 1: Approach, Analysis and Recommendations*. Cambridge: RAND Europe, 2004.

Claire Donovan is a Research Fellow at the Research Evaluation and Policy Project in the Political Science Program, Research School of Social Sciences, The Australian National University. During 2006 she was Chair of the Department of Education Science and Training's (DEST) Research Quality Framework Development Advisory Group (RQFDAG) Technical Working Group on Research Impact. The views expressed in this paper do not necessarily reflect those of the Working Group, the Development Advisory Group, or DEST.