



Council for the Humanities, Arts & Social Sciences



**Australian Government**

**Department of Education,  
Science and Training**

# Measures of quality and impact of publicly funded research in the humanities, arts and social sciences





---

## OCCASIONAL PAPERS

---

**Measures of quality and  
impact of publicly funded  
research in the humanities,  
arts and social sciences**

---

November 2005

---

ISBN: 0-9757701-2-8

© Council for Humanities, Arts and Social Sciences (CHASS). This publication is copyright. Apart from any fair dealing for the purpose of private study, research, criticism or review, as permitted under the Copyright Act, no part may be reproduced by any process without written permission. Enquiries should be addressed to the publisher.

This project was carried out with the financial assistance of the Australian Government, through the Department of Education, Science and Training.

The views expressed in this report do not necessarily reflect the views of the Department of Education, Science and Training.



**Australian Government**  
**Department of Education,**  
**Science and Training**

---

## Preface

---

With nearly two-thirds of Australia's university students, a majority of its practising researchers and around forty-five per cent of its research students now studying in the humanities, arts and social sciences, these fields are vital to Australia's research effort. Any system of research support or training not fully attuned to this sector risks neglecting our nation's talent and seriously underplaying its innovative capacity.

The project's report outlines an innovative model for future assessment of research in this sector, a model that is equally suitable for assessment of research in the science, technology, engineering and medical sectors.

The most novel aspects of the report concern the variety of ways in which research impact might be assessed and the need for a further, separate category—research capability.

The report is also significant in the way in which it argues for an ultimate and integrated qualitative judgment by experts of an evidence-based case. It finds that data and measures are only truly meaningful when placed within the context of a case, which is best made by a department or research group.

The project was conducted between December 2004 and June 2005 by the Council for the Humanities, Arts and Social Sciences to a commission from the Department of Education, Science and Training.

The council is grateful to the hundreds of scholars in the humanities, arts and social sciences, and to many from other areas, including the Department of Education, Science and Training, who have contributed to the work through the case studies, focus groups and reference group. In particular, it is grateful to Claire Donovan, Malcolm Pettigrove, Jonathan Powles and Aliya Steed, who have authored various of the appendices.

I express appreciation to Toss Gascoigne and the staff of CHASS, who managed the project; to Jonathan Powles, who authored most of the final report; and to Bryan Horrigan and Linda Butler, who powerfully contributed to several key aspects of the study.

### **Malcolm Gillies**

President

Council for the Humanities, Arts and Social Sciences

September 2005



---

# Table of contents

---

<b>Executive summary</b>	<b>7</b>
<b>1 The research project</b>	<b>8</b>
<b>2 Findings</b>	<b>10</b>
<b>3 Valuing the humanities, the arts and the social sciences</b>	<b>13</b>
<b>4 A ‘map’ of research in the humanities, the arts and the social sciences</b>	<b>18</b>
4.1 Origin	20
4.2 Production	20
4.3 Outputs	21
4.4 Outcomes	23
4.5 The humanities	26
4.6 The creative arts	28
4.7 The social sciences	30
<b>5 Indicators: quality, impact and capability</b>	<b>32</b>
5.1 Quality	33
5.2 Impact	38
5.3 Capability	43
5.4 Results from the case-study trial	46
<b>6 The proposed model</b>	<b>48</b>
6.1 Criteria	48
6.2 Assessment panels	51
6.3 Flexibility of weighting	52
6.4 Unit of assessment and reporting	52
6.5 Broader applicability of the model	54
<b>7 Conclusion</b>	<b>55</b>
<b>Appendix A: Setting the scene</b>	<b>56</b>
A1 Introduction	56
A2 Current measures of the quality and impact of publicly funded HASS research in Australia	56
A3 Current international measures of quality and impact of publicly funded HASS research	70
A4 Potential measures of quality and impact of publicly funded HASS	72

<b>Appendix B: The value to the end-user</b>	<b>82</b>
B1 Executive summary	82
B2 Summary of main responses to interview and questionnaire questions	83
B3 Multiple end-users, multiple values, and individual and community capability development: a case in point	94
B4 Conclusion	95
B5 Interview format	96
B6 Informants	98
<b>Appendix C: The case-study trial</b>	<b>100</b>
C1 Organisation	100
C2 Outcomes	101
C3 Documentation for the trial process	108
C4 A postscript	113
<b>Appendix D: Acknowledgments</b>	<b>114</b>
<b>Appendix E: Abbreviations and acronyms</b>	<b>120</b>

---

## Executive summary

---

For decades, researchers in the humanities, arts and social sciences (HASS) have endured a funding system that assesses their work through a narrow prism of quantity: How many papers? How much grant money? How many research students?

This project set out to find a fairer way to evaluate research, by focusing on the quality and impact of the diverse outputs of HASS research, such as books, performances, reports to government, films, libretti and professional advice.

The consultation process was wide-ranging and intense, involving over two hundred submissions and consultations, a survey of literature and practice, case studies, and focus groups. The project examined the evidence from Australia and abroad, and looked for distinctly new approaches, including proposals being developed by science agencies. Draft proposals were tested across seven disciplinary areas in six different institutions. In the end, a new model for assessing the quality and impact of research emerged.

This model applies not only to the humanities, arts and social sciences, but can be applied with equal validity across the whole research spectrum. It has the flexibility to work with all disciplines and with interdisciplinary research.

The model has achieved an uncommon level of acceptance among participants in the inquiry. Its acceptability results largely from its familiarity: the model owes much to the university promotions system, which typically evaluates a case for quality and impact against evidence of academic achievement, community standing, professional outreach, and peer regard.

In all, the model meets five key criteria for acceptability by governmental, academic and other research stakeholders: simplicity, familiarity, reliability, flexibility and universality.

The model identifies three factors that may be adjusted to suit the disciplinary balance and focus of the research unit in question: research quality, research impact and research capability. The report emphasises the importance and inextricability of these three factors.

**Quality** The report finds that any assessment of research quality is best achieved by peer review of an evidenced-based case, informed by key metrics. It finds that potential indicators of quality are broadly comparable between science and non-science sectors.

**Impact** By contrast, the key indicators of the impact of research on other researchers and the broader community are diverse and highly discipline-specific. They need to be evaluated by a panel of experts drawn from academe and users and beneficiaries of publicly funded research.

**Capability** The model introduces a further category, research capability, which reflects the capacity of a research unit to contribute to future goals of research and research training; that is, to ensure the vitality and diversity of Australian research in the years to come. The case for capability, like the case for impact, should be evaluated by a panel of experts and beneficiaries of research.

The clear conclusion from the report's survey of literature and practice is that standard bibliometric measures alone are not enough for such assessment. This position was supported by voluminous feedback from the sector, indicating quantitative data are best used as supporting evidence in a nuanced case.

The new model for research assessment is based on a case, supported by evidence, about which judgment is made by experts. Importantly, it introduces a menu of suitable tools for assessment, from which the research unit can choose in compiling its case. This flexibility enables the model to be used right across the research spectrum.

The project commends the new model, which affords Australia the opportunity to claim international best practice in assessing research quality, impact and capability.



# 1 The research project

The Council for the Humanities, Arts and Social Sciences (CHASS) was established in 2004 as a peak representative body. The disciplines and practices it represents encompass the majority of Australian researchers. The function of CHASS is not limited to the academic sector: it plays a key promotional and advocacy role for the humanities, arts and social sciences (HASS) sectors as a whole and as a coordinating forum for academics, students, business and the broader community.

Announcing the formation of CHASS, the Australian Government Minister for Education, Science and Training, the Hon. Dr Brendan Nelson, observed:

*The humanities, arts and social sciences are critically important to the future development of Australia. Not only do they play a key role in supporting the national innovation system, they make a significant contribution to the development of our society, culture and individual identity. It is from this sector especially that the soul is passed from one generation to the next.*

In December 2004, CHASS commenced work on a research project, commissioned by the Department of Education, Science Training (DEST), to examine measures of the quality and impact of publicly funded research in the HASS sector. This report and its appendices are the outcome of that project.

The exercise has been essentially consultative, with the goal of representing the views of all stakeholders in HASS research. Stakeholders include not only academics, but also the users of research. The study had three major phases:

- 1 A survey of the literature, practice and current developments in the assessment and evaluation of research, with regard to:
  - HASS sector research and the problems associated with its evaluation
  - recent attempts, both within Australia and internationally, to measure the quality and impact of research, and specifically of HASS research
  - recent developments in evaluation practices that may prove useful in the evaluation of HASS research.

The survey of literature and practice informed the questions posed to the community during the consultative phase of the project, and the nature of the indicators suggested later in this report.

- 2 Consultations with stakeholders, including:
  - the formation of a broadly representative reference group to oversee the study
  - a web-based survey, which received more than 160 submissions, and a more detailed follow-up questionnaire to selected respondents
  - several consultative forums with academics, researchers and research managers
  - a series of focus groups with representative end-users of research.

Views expressed in the submissions have been taken into account in the report; many have been quoted verbatim, as they encapsulate some of the central issues, questions and problems in the assessment of HASS research.

- 3 A trial exercise to test aspects of the research assessment process.

The process was tested by inviting seven representative academic departments from different types of university and from different HASS disciplines to compile a 'case for excellence'. Each case was to provide evidence of the quality and impact of the department's research and of the research capability of the organisation as a whole. The intention was to determine which particular indicators provided the best evidence of research quality, impact and capability.

Two other studies fed into the final report. The first was a testing of the 'payback' model for assessing the outcomes and outputs of research funding by applying that model to the creative and performing arts; the second examined the value, to the end-user, of research in the HASS sector.

The three main strands to the research project—survey of literature and practice, consultations, and trial—have been incorporated into the final report. Four appendices give greater detail of the study and the participants:

- Appendix A is a detailed review of the literature and recent practice in the quantitative and qualitative assessment of research quality and impact in the HASS sector.
- Appendix B is a detailed account of the consultative exercise with the end-users of HASS research.
- Appendix C gives details of the processes and outcomes of the case-study trial.
- Appendix D lists the individuals and organisations—researchers, policy makers and end-users of HASS sector research—who contributed to the project.

## 2 Findings

This section summarises the findings of the study.

**A Standard bibliometrics are not an appropriate way to assess research in HASS, because they fail to capture so many important outputs of this sector.**

No quantitative indicators can be relied upon exclusively. Standard bibliometric practices do not capture the variety of research outputs (for example, books, documentaries and policy reports) in the humanities and social sciences, and are clearly absurd in the creative arts. Indicators derived from success in research funding fail to capture the very different sources of such funding across the sector, and the very different costs of conducting research in different disciplines. However, such indicators should not be completely discarded. They can provide useful evidence to help inform the judgments of appropriately constituted panels.

**B The best way to evaluate research is through assessment by peers and experts.**

Most evidence points to the fact that evaluation of quality is best achieved through appropriately informed qualitative judgment. The overwhelming majority of respondents to the consultative processes favoured research assessment based on peer assessment. International examples of judgments currently made on peer review include the United Kingdom Research Assessment Exercise; in Australia, examples are the awarding of Australian Research Council (ARC) grants and academic promotions.

**C Research groups or departments are the most suitable unit of assessment for HASS sector research.**

Although there are attractions in assessing research at the level of an individual researcher or by discipline within a faculty or university, the prevailing view in this study was that the research department or group was the appropriate size. Assessment at this level gives greater transparency to the assessment process, and allows a more precise alignment between research planning and performance and national priorities.

The HASS sector felt that research success should be judged in the national or international context, rather than the institutional context. Research is increasingly collaborative, but the collaboration is between cognate research groups in different institutions, and between such groups and industry, rather than between different units of the same university or organisation.

**D The model for research assessment should be based on a case, supported by evidence, about which judgment is made by experts.**

The case put by a research unit should be constructed around quality, impact and capability. The case should be assessed in a two-stage process: first by academics with specific disciplinary expertise, and then by a conglomerate panel of experts including academic peers, experts from industry and experienced research managers. Figure 1 shows the process.

**E In assessing a research unit, three components should be examined: quality, impact and capability.**

- *Quality* refers to the academic or artistic standing and significance of the research, and its direct influence on academic practice in the field. Indicators of quality are broadly comparable across all fields of research (including the sciences).
- *Impact* refers to the influence the research has had more widely in society. This includes impact not only on the academic discipline, but also on policy, public awareness, institutional and artistic practice, and commercial, economic, social and cultural benefit as appropriate to the specific discipline. Indicators of impact are largely specific to the field of research.
- *Capability* refers to the research group's ability to contribute effectively to future research locally and nationally, and includes the number of graduates and postgraduates the unit produces and an assessment of the skills they add to the workforce.

**F Research units should choose the most appropriate measures to assess the quality, impact and capability of their research from a menu of options.**

These measures will take into account disciplinary differences. The report recommends that a range of measures be listed on a menu of indicators, to be used as appropriate in the assessment process. It is not always possible or desirable to distinguish or prescribe which indicators apply primarily to ‘quality’, which to ‘impact’ and (to a lesser extent) which to ‘capability’.

**G Research units should have flexibility in nominating how they are to be assessed.**

In preparing its case, each unit should have the flexibility to specify, within defined limits, the weighting given to each of the three categories of quality, impact and capability:

Quality: 40–70%

Impact: 20–50%

Capability: 10–30%

Once the minimum levels of 40% (quality), 20% (impact) and 10% (capability) are met, each unit should have the freedom to allocate the remaining 30% at its discretion. This approach allows flexibility in the assessment framework, and accommodates the differing natures of specific disciplines and the differing strategic goals of research units.

**H Broadly, quality can be measured the same way across all disciplines.**

The assessment of research quality is best achieved by peer review of an evidence-based case. Most disciplines have forums for the dissemination of work regarded as having high intrinsic standing, including journals and galleries. Most disciplines have professional or academic bodies, with elected membership a hallmark of quality. While precise quantitative comparison between these is problematic, researchers generally accept broad qualitative judgments about quality made on the basis of such evidence. The evaluation of quality is best achieved through appropriately informed qualitative judgment.

**I Impact should be measured by a panel of experts drawn from academe and users and beneficiaries of research, who examine an evidence-based case prepared by the research unit.**

The key indicators of the impact of research on other researchers and the broader community are diverse and highly discipline-specific. The assessment of impact should be qualitative, and construed differently from the assessment of quality. Specifically, the impact should be assessed using an ‘audit’ model. Institutions, research groups or departments can compile an evidence-based case for the impact of their research. Appropriately qualified experts—both academic and, crucially, experts from industry, public institutions and other research ‘end-users’—should then audit some of these cases to determine whether the evidence presented supports the case made.

**J Capability reflects the capacity of a research unit to contribute to future goals of research and research training, ensuring the vitality and diversity of Australian research.**

The case for capability should be evaluated by a broadly drawn panel of experts, including experienced research managers—for example, deputy vice-chancellors or pro vice-chancellors (research). Assessment should include an examination of the strategies in place to build the unit’s research capability and of the number of graduates and postgraduates the unit produces, including an assessment of the skills they add to the workforce.

**K The model recommended by this study applies not only to the humanities, arts and social sciences, but can be applied with equal validity across the whole research spectrum.**

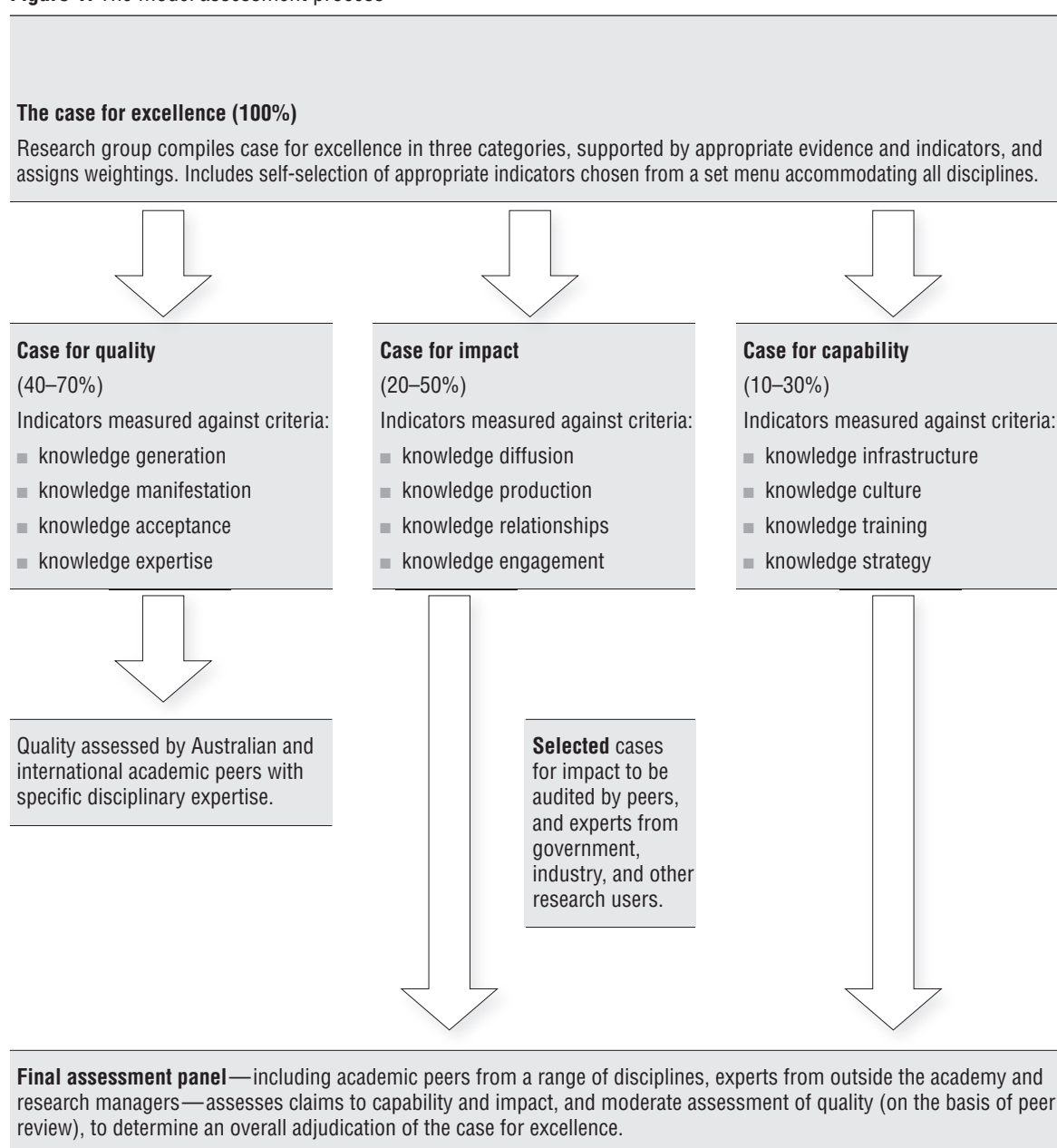
The model can accommodate the diversity of research practices and outcomes in different disciplines; the diversity of types of unit (department, faculty, centre) that undertake research; and the ways in which research quality and impact contribute to the disciplines and to society. It has the flexibility to work with all disciplines and with interdisciplinary research.

**L Existing measures of research quality and impact can be used to advantage in the construction of a case.**

A number of evaluation processes used for other purposes can also be used to assess research. These include the processes for assessing competitive grant applications; the processes for assessing academic promotions; and the processes for selecting articles for publication. They provide good models for research assessment on a wider scale and, crucially, already have the confidence and understanding of the research community.

**M While standard quantitative measurements are both inadequate and unfair for evaluations in the humanities and the arts, international experiences with newly conceived non-standard bibliometrics show interesting potential.**

Section 3 of Appendix A explores a number of non-standard metrics being employed in Europe and New Zealand.

**Figure 1:** The model assessment process

### 3 Valuing the humanities, the arts and the social sciences

Perceptions of the value of research in the humanities, the arts and the social sciences are changing. As the Minister for Education, Science and Training, Dr Brendan Nelson, stated in a National Press Club speech on 8 March 2005:

*We need to make sure that humanities and social sciences are considered to be no less important in building our future, a future on science and technology. Because in the end, if all of the scientific problems of life were ever solved in an applied sense, the most important questions would remain unanswered.*

There is an increasing appreciation not only of the contribution that HASS research makes in its own right, but of the importance of cross-disciplinary research. This is coupled with a growing awareness that researchers in HASS can make a significant contribution to solving intractable problems previously seen as lying solely in the province of science, technology, engineering and medicine (STEM). Examples include public attitudes to and behaviour in the use of water; non-engineering factors contributing to the road toll; human factors in sustainable transport and development; coherent medico-legal and socio-ethical approaches to the critical care of patients; and the health of Indigenous Australians. Another example is the contribution HASS makes to Australia's defence and anti-terrorism measures through language and cultural expertise in Asia.

What, then, are these 'important' questions to which Minister Nelson refers? What are the practical outcomes of HASS research? Does this research have economic benefits? And does a combined approach by HASS and STEM researchers provide solutions to intractable problems?

In absolute terms, the important questions are those that strike at the heart of who we are as individuals and as a society. The humanities ask us to account for ourselves, our history, our stories and our human values. The social sciences draw on rigorous investigation to help us make informed choices about the sort of society we wish to live in—how we organise education, health, wealth and security. The arts ask the same questions in different ways, providing not scientific knowledge but insight—the flashes of inspiration that illuminate, and encapsulate, our place in the world.

These questions by no means belong to the world of academia. Researchers in the humanities and social sciences, and artists, carry an enormous responsibility in return for the privilege of being able to practise their disciplines: they must not only inform society but listen to it, with attention, rigour and sensitivity. If practised well, 'research' in the arts is inextricably linked with creative arts all over Australia, from internationally renowned musicians and actors, through to local community events and galleries. Social sciences research happens not only, and perhaps not even principally, in universities but also in decision-making, informing the policy and practice of governments, charities, corporations and public organisations of all types. And although the millions of Australians and international guests who visit our museums and libraries, who celebrate our history or question its legacy, and who engage with public debate on issues ranging from the ethics of health to the freedom of religious practice might not consciously consider themselves as being involved with the 'humanities', this is exactly what they are doing.

The 'important questions' were echoed in the 2005 Research Quality Framework Issues Paper, which referred the 'broader impact for society, through economic, environmental and social benefits'.<sup>1</sup>

<sup>1</sup> Research Quality Framework: assessing the quality and impact of research in Australia. Issues paper, March 2005, p. 10

The second issue concerns the practical outcomes and economic benefits of HASS research. The outcomes were encapsulated in a May 2005 report to DEST, *Commercialisation of research activities in the humanities, arts and social science in Australia*, which found an enormous diversity of ways in which HASS research leads to immediate and practical benefits to society.<sup>2</sup> Many of these lead to direct economic benefits. For example, the report found thriving and active professional consultancies drawing on research in disciplines as wide-ranging as history, archaeology, editing and publishing, philosophy, social science, information services and information technology. Much of this work was with government: heritage, tourism, information services, health, publication management, developing strategies and plans, filmmaking and training. The report gave a dozen examples of projects and their value in dollar terms, with amounts ranging up to \$5 million a year.

While HASS research can lead to new jobs and new industries, it can also generate savings. This can be by introducing management efficiencies (better approaches to teaching and to administration), and by changing public attitudes to the way society consumes resources such as water. One view is that the economic benefits flowing from research which generates savings equal those that flow from the creation of new jobs.

The May 2005 report found that HASS researchers were attracted by the idea of being relevant, influential and connected to their communities. Such qualities allowed them to become engaged with important social and community problems. One respondent to that report typified this community engagement which characterises HASS research:

*An urban economist is linking with a senior academic in architecture through an ARC Linkage grant involving Mirvac (a large property developer) to explore the design needs of baby boomers. Mirvac is not particularly interested in IP as it already has a leading place in the market, so the research results are largely for the public good—looking at housing futures for Australia’s largely urban population.*<sup>3</sup>

In describing a program to explore the reasons young people were removing themselves prematurely from the education system in north Queensland, the report said:

*Regional universities, in particular, believed they had a special mandate to support communities. Engaging with communities was seen as a priority activity. They expressed concern at the ‘publish or perish’ mode of thinking, and were very critical of the detrimental effect this may have on engagement with the community.*<sup>4</sup>

Such studies are beginning to put paid to the outworn notion that HASS research is divorced from the real world, and instead show it to be as relevant and practical in its orientation as research in the STEM disciplines. The practical outcomes that the work generates often spring from collaboration, including combined approaches with STEM researchers, and cover many areas:

- Many of the most significant causes of mortality in Australia, such as road deaths and injuries, suicide and some diseases (heart disease, lung cancer, skin cancer), have their origins in human behaviour and understanding, and can be linked to individual, social and cultural beliefs, expectations and norms. HASS research can complement STEM research to contribute to responses by developing better knowledge of why people, as individuals or groups, act as they do, and how they might act differently. Mortality rates for road deaths might be reduced by changes to policy and legislation informed by HASS research. There is potential for mortality rates for a range of diseases to be reduced by relatively simple changes in social beliefs and behaviour (e.g. dietary changes), to which HASS research can make a valuable and comparatively cost-effective contribution. Community awareness and understanding of the ethical, philosophical and social dimensions of scientific, medical and technological innovations involving ‘life and death’ matters, such as stem-cell research and end-of-life medical decisions, depend much on HASS research.

<sup>2</sup> Gascoigne T and Metcalfe J (2005). *Commercialization study on the humanities, arts and social sciences sectors*. Department of Education, Science and Training, Canberra.

<sup>3</sup> *ibid.*, p.14

<sup>4</sup> *ibid.*, p. 20



- As Australia deals with problems of water availability and supply, simple changes in consumer understanding and behaviour can make a massive difference, without the need to resort only to high-technology and science solutions and expensive infrastructure (e.g. desalinating seawater, building new dams).
- The HASS sector makes substantial contributions to research that informs the education, tourism and entertainment industries, which have well-recognised social and economic benefits for Australia.
- HASS research helps to inform understanding of and behaviour in Australia's government, economy and international relationships—in other words, a modern civil society's essential preconditions for stable government, law and order, a functioning economy, business certainty, and successful R&D investment and commercialisation.
- As the Tertiary Education Commission of New Zealand noted in its 2003 report on performance-based research funding and research excellence, the very processes and methods for designing and implementing an assessment of a nation's research quality, impacts and benefits fall within the purview of the social sciences.<sup>5</sup>

In fact, it is impossible to quantify the total 'contribution' of the humanities, the arts and the social sciences to society. This is because, in these disciplines, *society itself is the subject of the research*. There is—or should be—an inseparable relationship between research as it is conducted in the academy; research in a broader sense as it is practised by professionals working in the social, cultural and artistic professions; and the everyday lives of Australians.

The greater the diversity of research in the humanities, arts and social sciences that circulates through any given society the more cultured, critical and aware that society is.

Assessing the value of creative research, however, is not as obvious or accessible as assessing the value of other types of research. The value of creative research has a culturally cumulative effect ... Thus, 'genuine assessment of the value of research in the sector' implies assessing the cultural 'health' of Australian society, specifically, and the awareness our society has of the diversity of works produced by creative researchers/practitioners.

*Submission, Dr Eril Baily, Associate Dean Research,  
Sydney College of the Arts, University of Sydney*

In the current political climate it seems problematic for us not to be able to argue that HASS research generates critical public debate ...

*Submission, Dr Fiona Martin, Lecturer in Journalism and Media Production,  
Southern Cross University*

One of the intriguing tasks of this research project has been to attempt to distinguish, for practical purposes, between 'research' as conducted by universities and publicly funded research agencies (PFRAs), and the wider practice and activity within the humanities, the arts and the social sciences. The boundary is impossible to pinpoint—and this is a good thing. The difficulty indicates, on the one hand, a flourishing and relevant HASS research culture in our universities and other research organisations, and on the other a vital, responsible and self-aware culture and community in Australia.

Indeed, a point of difference between the HASS disciplines and the STEM disciplines emerges here. Most STEM researchers in Australia work either in universities or PFRAs, or in the private sector, but the converse is true in the HASS sector. Most artists, writers and musicians who receive public funding to assist their creative practice are not staff of universities. Most people who undertake professional research in economics do not do so as academics, but as researchers working for private or—significantly—public institutions outside the academy. Even in the humanities, a significant amount of genuine research, much of it publicly

<sup>5</sup> For example, see <http://www.ms.govt.nz/publications/journal/24-April-2005/24-pages55-84.html>



funded, takes place outside ‘research organisations’. For example, more professional historians in Australia work outside universities than inside them.<sup>6</sup>

One indicator of the current and future strength of the research effort in the various HASS disciplines is the number of students enrolled in research doctorates in each broad discipline area. Table 1 gives the numbers for 2004.

**Table 1:** Doctorate by research student load, 2004, by broad field of study

Natural and physical sciences	5,308	Architecture and building	231
Information technology	1,090	Education	1,471
Engineering and related technologies	2,169	Management and commerce	1,310
Agriculture, environmental and related studies	861	Society and culture	5,960
Health	2,842	Creative arts	870
<b>Total STEM:</b>	<b>12,270</b>	<b>Total HASS:</b>	<b>9,842</b>

Source: Selected higher education statistics, DEST:

<[www.dest.gov.au/sectors/higher\\_education/publications\\_resources/profiles/students\\_2004\\_selected\\_higher\\_education\\_statistics.htm](http://www.dest.gov.au/sectors/higher_education/publications_resources/profiles/students_2004_selected_higher_education_statistics.htm)>

By these recent figures, 55% of current PhD student load is in the STEM sector, and 45% is in the HASS sector.

It is illuminating to compare these figures with the figures for total numbers of all students (including undergraduates and postgraduate coursework students). These are given in Table 2.

**Table 2:** Total numbers of students (load), 2004, by broad field of study

Natural and physical sciences	77,406	Architecture and building	12,971
Information technology	50,741	Education	54,039
Engineering and related technologies	39,563	Management and commerce	127,867
Agriculture, environmental and related studies	9,242	Society and culture	176,855
Health	63,543	Creative arts	48,393
<b>Total STEM:</b>	<b>240,495</b>	<b>Total HASS:</b>	<b>420,125</b>
<b>Total student load in all identified HASS/STEM courses</b>			<b>660,620</b>

When undergraduate and postgraduate coursework students are taken into account, the HASS sector comprises fully 64% of the total, while STEM students make up 36%.

These two tables indicate more than simply the strength and size of the two sectors. The difference between the two proportions—PhD students compared with all students—highlights the point made above about the connection of the HASS sector with the real world. Large numbers of individuals with education—often research-led education—in HASS practise their professions, and their discipline, in the broader society. Understanding the intricate web between graduates, research students and researchers inside and outside the academy is essential in understanding the way HASS research, and graduates informed by it, are diffused throughout society. By contrast, STEM sector research is far more concentrated around its essential infrastructure in industry, universities and PFRAs.

Under Australian Government arrangements for publicly funded research, HASS researchers receive a much smaller piece of the research funding pie than do STEM researchers. For example, in ARC Discovery

<sup>6</sup> Submission, Professional Historians Association (NSW) Inc.

Grants funding for research projects commencing in 2005, the four science-based panels awarded a total of \$214 million in research funding (approximately 72.5% of total Discovery Grant funds) and the two HASS-based panels awarded a total of \$81 million (the remaining 27.5%).

Similarly, in DEST's reporting of gross domestic expenditure on research and development (GERD) by field of research, recent figures show that approximately 33% is spent on engineering and technology; 15% on information, computing and communication sciences; 14% on medical and health sciences; and 8% on biological sciences. HASS research spending is likely included in the 10% that DEST describes as going to 'other research fields, courses and disciplines'.

The STEM sector has taken to heart the need to explain the value of its work to the community, to such an extent that there is a common perception in the community that 'research' means exclusively scientific research, and that the economic and social benefits of research derive mostly from scientific research.

The humanities, arts and social sciences have a strong case, but they have not made this case with strength and conviction. It is important to identify and make widely known the real impacts and benefits of HASS research, not just its quality as measured against scholarly standards. The community needs to be reminded that HASS research has social value, commercial worth, and community and cultural benefits—and people working in the HASS community are the ones to make these arguments.

Until this argument is articulated clearly and convincingly, HASS research remains in danger of being inadequately valued, despite the fact that it deals with 'the most important questions'.

It is important to acknowledge the diversity [of HASS research] and to resist shortcuts which might reduce what is taken into account for the sake of tidier administration and the self deception that it is simple after all—or should be. Having acknowledged that, one can approach the challenge of assessing value. This immediately raises the issue of value: whose value? What value? What is the currency? This needs to include a range of questions: are there immediate or short term financial benefits, such as we might find in some applied science research?, but also what are the long term benefits, which may be quite broad. The axis: 'short term—long term' and the axis: 'financial—other' are both important.

Short term financial gain might extend to include such things as multi-media productions which then become potential money spinners at the box office. Longer term gains need to encompass the issue of the role of culture, including religion and the arts, in contributing to the well being of society and its members and, on the financial side, the costs incurred when that wellbeing is undermined. Depression, aggression, antisocial behaviours, self-harm have complex causes. Those causes include the extent of an individual's or a community's capacity to create meaning in life which sustains them. To the extent to which this fails, the community faces major costs in the welfare budget in a wide range of areas. Most people need myths to live by and frameworks in which they value themselves. This is as fundamental for human creativity as a secure and nurturing environment is for early human development.

If this is not to remain simply a plausible hypothesis to which many implicitly give assent, it needs to be backed up by studies identifying the correlation between these factors which could then support what is still too often a widely held hunch. It should be possible to show the correlation between diminished social and cultural enrichment and social and economic costs to the community.

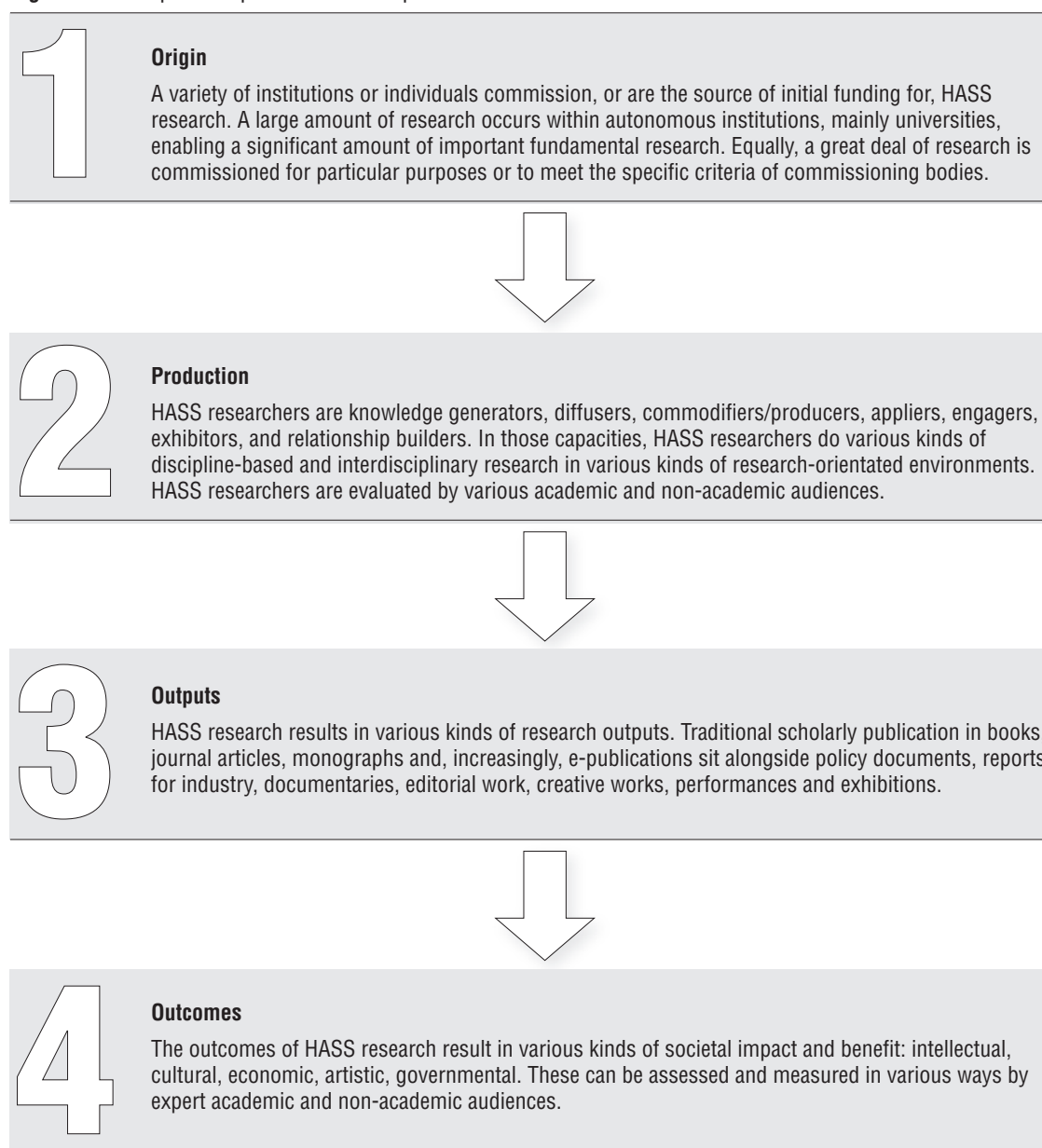
What has this got to do with research? It bears directly on research which forms part of the cultural/social/religious enrichment of a community and contributes to people having frames within which to find meaning and value for their existence without which they become a cost to others and themselves.

*Submission, Professor William Loader, FAHA,  
President of the Australian and New Zealand Society for Theological Studies*

## 4 A 'map' of research in the humanities, the arts and the social sciences

The first task in outlining a sophisticated approach to the evaluation of research quality and impact is to outline the 'lie of the land'. A 'map' of research activity in the HASS sector (Figure 2) captures the diverse sources of funding for HASS research; the diversity of its outputs, from scholarly publications to creative artworks; and the huge variety of individuals and organisations that make use of this research. Who funds the research, and how? Who conducts it? What types of publications or other outputs result from the research? Who uses the research? What is the impact of the research on academia and on society more generally?

**Figure 2:** Conceptual map of the research process



An important feature of this map is its identification of existing measures, indicators and processes for the assessment of research impact and quality. A number of evaluation processes used for other purposes can also be used to assess research, including those for assessing competitive grant applications, those for assessing academic promotions, and those for selecting articles for publication. These processes provide good models for research assessment on a wider scale and, crucially, already have the confidence and understanding of the research community. Such confidence and understanding is vital for the effectiveness, transparency and acceptability of an assessment process that could be employed for a number of purposes, including a Research Quality Framework (RQF).

It is difficult to overemphasise the importance of any RQF accessing and taking into account such existing measures. Three important criteria for an RQF are transparency, effectiveness and acceptability. Relying on current processes, which are already well understood and trusted by researchers, for the assessment of quality guarantees acceptability. Where an RQF can draw on existing measures to reduce the resource burden—including, crucially, the time that researchers need to spend in preparing their case—effectiveness will follow. The model for research assessment presented in this report allows academics to draw largely on such extant data. The success of seven academic units in compiling a case for excellence, often from existing data, in the very short time allowed for the case-study trial demonstrates the value of using these measures.

At its most general level, the conceptual framework given in Figure 2 can be used to illustrate the research process, from origin to eventual outcomes in benefits to knowledge and society. At this level, the stages shown are common to all HASS disciplines, and indeed to all research.

The map shows four broad domains or stages of the research process. In each, there is a wide range of existing processes for the evaluation of research quality and, in some cases, impact. One of the purposes of the mapping exercise is to document all the varieties of *existing* research quality assessment, so that any new recommendations are informed by existing practice—as several submissions to this project pointed out, there is no sense in re-inventing the wheel. The following sections outline some of these existing measures, indicators and processes for the assessment of research quality and impact at each stage—origin, production, outputs and outcomes.

While books and articles form a measure of personal achievement, they do not measure impact or benefit for either the scholarly or wider communities. Policy reports too often fall on deaf ears or are relegated to dead files. Opinion and analysis in the popular press has the limited lifespan of 24 hours. If this is indeed the case then it implies the need to determine new and hitherto unapplied criteria for evaluating the impact and outcomes of research in the HASS arena. I realize that this is precisely what you are seeking to identify. My point is that what does not so far figure in the research evaluations are measures that encompass ‘benefit’ and ‘impact’ appropriate to this arena.

Perhaps a genuine assessment of the value should be predicated upon assessing the potential (and actual) outcomes of projects based on the extent to which they address such concerns as personal development, social cohesion, community empowerment and self-determination, ‘local’ image and identity, imagination and vision, health and well-being. (I’m much drawn towards the findings of people like Matarasso in his book about the social impact of participation in the arts.)

This leads to the second question you have identified related to forms of assessment. Should research in our areas take into account the social impact of a research activity measured in relation to its aims and those of its stakeholders? I believe it should—or rather that this might form one of the bases for the assessment of its quality.

*Submission, Prof. Victor Emeljanow, Professor of Drama, Assistant Dean (Community Relations), School of Music and Drama, University of Newcastle*

## 4.1 Origin

A great deal of HASS sector research takes place in universities, funded by the block grant. Much fundamental research is initiated in this way. However, it also needs to be recognised that a large amount of publicly funded HASS sector research is commissioned by other means—contestable grants funding from organisations such as the ARC; research commissioned by and conducted in partnership with industry, with both public and private-sector funding; policy research commissioned by government; artistic practice funded through arts grants from various levels of government; and self-funded research.

A large number of existing measures of research quality inform the process by which research is funded, commissioned and supported, even when this is by indirect means such as university block grants. These include:

- the total research output of an organisation
- research student completions
- partnership funding for research.

These familiar indicators govern the DEST block grant allocation to universities. It is worth noting that most respondents to the consultation exercises for this project regarded research students as a vital component in any overall assessment of research quality. The collection of simple quanta of research output was viewed with some suspicion but, nevertheless, was also seen to be an important piece of evidence.

Other markers of quality that are currently employed at the ‘origin’ phase of the research process include:

- peer evaluation (for contestable grant funding, such as ARC or Australia Council grants)
- industry criteria for competitive tender (often an assessment of future capability as much as past performance, with an emphasis on value for money).

Where research is commissioned for a purpose, judgments of quality are made according to specific criteria. In these instances, ‘quality’ means much the same as ‘fitness for purpose’, and the purpose needs to be known in order to judge quality. This argument was advanced in the submissions in favour of peer and expert evaluation rather than more mechanistic measurement; those submissions held that peers and experts are in a better position to judge the research in context. In the HASS sector, in particular, the context in which the research is commissioned and from which it originates is often vital in understanding its relative merits.

## 4.2 Production

In the HASS sector, there is no clear dividing line between ‘research’ as conducted within universities and other research organisations, and the broad practice of the humanities, the arts and the social sciences in society. In many cases, it may be difficult to distinguish differences in quality and impact between the research produced on issues of social relevance in universities, and that produced in government departments and private organisations. Much of the artistic and creative practice emanating from creative arts researchers in universities sits alongside, and is indistinguishable from, that of the wider creative industries. It is difficult to construct a very strict definition of ‘who’ produces HASS research, and this is a good thing. The difficulty points to the ongoing dialogue between HASS research in publicly funded research institutions, and the practice of the humanities, the arts and the social sciences more broadly—and the relevance of each sector for the other.

However, for the purpose of constructing pragmatic measures of quality and impact for research conducted in publicly funded organisations, it is possible to limit the discussion to that research conducted within universities. Unlike researchers in the STEM sector, in which a great deal of research takes place in PFRAs, HASS researchers, narrowly defined, are almost exclusively employed in universities.

Therefore, in practice, a number of measures of the quality of *researchers* are already employed on an almost day-to-day basis, principally by the universities themselves. Indeed, many of the submissions to this

project discussed some of the difficulties that arise when a single set of measures (for example, those used in assessing applications for academic promotion) is applied to researchers in fields as distinct as social work, philosophy and drama.

Nevertheless, these measures exist and are already used, and it is worth summarising them:

- *Academic promotions criteria* are those with which researchers are currently most familiar. Typically, research performance is judged alongside teaching performance and service to the university or wider communities. Most promotion processes involve specialist academic assessment of research performance (from referees who are senior figures within the discipline), which is then taken into account by a broader group—a faculty or university promotions committee—that makes an overall judgment on the merits the case. The process is tried and tested, and has the broad confidence of most academics as an assessment of quality. This is one reason why the research assessment process recommended in this report bears strong similarities to existing academic promotions processes.
- *Internal research funding.* Often, universities allocate research funding on the basis of their own internal judgments of research quality. Equally often, these criteria are nuanced by institutional priorities and the institutional mission. Given that research production takes place within, and is influenced by, this framework of institutional criteria, it is important that any national research quality evaluation process allows researchers from various institutions, groups or departments to contextualise their research when making their claims for research excellence.
- *Peer esteem.* Within the various disciplines, there are many indicators of the esteem in which a particular researcher is held by his or her peers. Academics making judgments about the quality of one another's work take into account elected membership of learned societies and academies, professional bodies and editorial and curatorial boards, invitations to speak at conferences, and many other discipline-specific markers of peer esteem. The pervasiveness and reliability of such indicators of *researcher* quality (which are indirect indicators of *research* quality) is one of the arguments that has been advanced in favour of the efficiency and reliability of peer evaluation as the basis for research assessment.

## 4.3 Outputs

Research in the HASS sector is characterised by widely diverse outputs. A great number of researchers publish in traditional scholarly forms: books, articles in scholarly and professional journals, monographs, reviews and an increasing variety of e-publications. One constant theme in submissions to this project was the need to emphasise the continuing importance of the book as a significant—perhaps the most significant—form of research output in the HASS sector as opposed to the STEM sector.

There is a great deal more. Film and television documentaries, and other multimedia forms presenting research findings, are increasingly important in many if not all HASS disciplines. Much research that has significant impact is presented in forms that are most usable by those who need to access the research—policy reports, consultancy reports, staff development for commissioning organisations, and videos, manuals and websites for highly applied disciplines such as social work and management. In the creative arts, of course, research outputs are most commonly in the form of artistic products—paintings, sculptures, musical compositions and performances, plays, novels and so on. And in the arts, perhaps more than in any other area of HASS, new technologies are enabling ever newer media and categories of research output.

One difficulty commonly reported across the sector is that this diversity of outputs is not recognised in funding and promotion systems. The May 2005 report, *Commercialisation of research activities in the humanities, arts and social sciences in Australia*, reported that the university promotion system was a major disincentive for undertaking activities outside a narrow band.

To be considered for promotion, academics need to have an acceptable publication record, and to contribute to their institution's research performance under the government's performance-based funding schemes. Many participants reported that there was limited recognition of many of their commercial activities.



This position was compounded for researchers who produced such research outputs as CDs, films and exhibitions. These activities gained little or no credit, as they did not gain performance funding ‘points’ under the system:

*Commercialisation is not seen as a benefit towards promotion, it is not valued or supported (Brisbane)*

As a result, academics reported wrestling with the definitions of their projects and consultancies in order to squeeze them into a points-gaining category:

*I have a problem about what is research and what is not. Take classroom development. You can frame it in the research agenda or you can frame it as developing new programs. It is often seen by DEST as developing new programs, which is not considered research. Whether it is research, consultancy or program development makes a difference with getting points. This makes a difference to whether we do or don't do the work. It gets very confused. It takes so much time to renegotiate boundaries all the time. I have spent so much time to renegotiate boundaries on things I thought we'd clarified. (Perth)<sup>7</sup>*

Despite this diversity, current methods of assessing output quality have features in common. In almost all disciplines, research outputs that are primarily intended for academic audiences, and many of those intended for wider dissemination (particularly in the arts), are refereed by peers as a matter of course. Whether the work is an article to be included in a scholarly journal, a book to be published by a reputable press, or a painting to be exhibited in a national exhibition, the process of selection involves robust and rigorous vetting by peers.

With research that is more applied than pure, such vetting is more commonly done by the end-users of the research. The criteria these users employ in making judgments about utility and quality are more complex, more influenced by the purposes or intended outcomes of the research, and harder to describe.

It is ... worth noting that the focus on the Arts and Humanities sector as a thing apart ignores the difficult-to-quantify benefits that derive from cross-disciplinary collaborations, especially with the sciences. As just one example, creative production techniques especially in the virtual world of digital technology open up new ways of conceptualising, presenting and delivering information. This has applications extending way beyond television and film to medical imaging and virtual laboratories. According to Gaukroger (1998) the artificial divide between science and arts is not just counterproductive but also poses potential threats to our social wellbeing:

*we cannot help but be apprehensive about developments in areas such as bio-technology, where deep and difficult ethical issues are raised about the responsibility of pursuing a whole area of research. Those working in the areas of law, the humanities, and the social sciences have as much responsibility as scientists and technologists to make sure that such issues are faced up to and dealt with in a serious way.*

In relation to the value of outputs in this area, much work has been done in the UK and the US and more recently in Australia on the notion of ‘creative industries’. Richard Florida in the US (2003) and Australia’s David Throsby (2001) are but two examples of how economists have attempted to define an economic model to track value in this sphere. This body of work shows how it is possible to capture not only the conventional ‘capital’ value of arts-based research outputs, but also the more intangible flow-on values that have hidden or longer-term impacts in society. There is no need to re-invent the wheel—the transformation of creative industries in the UK under the Blair government shows the flow-on advantages of investing in creative industries as an integrated part of an overall economic development policy. In Australia the House of Representatives Standing Committee on Communications, Information Technology and the Arts (2004, xxv) notes in its report into the film, animating and electronic games industries that

*The economic effect of the creative industries is much higher than any other sector, except education. For instance, government investment to increase demand in the creative digital industries will have a 23% higher positive impact on the economy than a similar investment in primary industry.*

*Submission, Dr Gail Phillips, Associate Professor, Chair of Media Studies, School of Media Communication and Culture, Director, Research and Development, Division of Arts, Murdoch University*

<sup>7</sup> Gascoigne and Metcalfe, op. cit., p. 27.

## References

- Florida R (2003). *The rise of the creative class*. Pluto Press, Melbourne.
- Gaukroger S (1998). The relationship of research in the humanities to science and technology. In: *Knowing ourselves and others: the humanities in Australia in the 21st century*, Volume 3 (reflective essays), Australian Academy of the Humanities, Canberra. [http://www.humanities.org.au/review/c10\\_Gaukroger.html](http://www.humanities.org.au/review/c10_Gaukroger.html)
- House of Representatives Standing Committee on Communications, Information Technology and the Arts (2004). *From reel to unreal: future opportunities for Australia's film, animation, special effects and electronic games industries*. Commonwealth of Australia, Canberra.
- Throsby D (2001). *Economics and culture*. Cambridge University Press, Cambridge.

## 4.4 Outcomes

It is in the area of the final outcomes of research, of course, that the disciplines within the HASS sector vary most widely. Obviously, a symphony will have a quite different effect from a piece of social policy research, regardless of structural analogies in the way in which they are funded and produced. However, there are broad similarities in the outcomes of the research process.

The key here is the notion of value. What do we mean by value? Valued in whose terms? Is value always quantifiable? If we look to the past one way of considering this is in terms of valuing knowledge, ideas and learning for their own sake. If we must give a value to knowledge gained from research, regardless of the outcome or how it is shared—book, article, film, media—it should be in terms firstly of the value of knowledge in and of itself, and secondly of the value of what that knowledge brings to Australian culture. We will be an intellectually famished nation indeed if we lose sight of this and attempt to attach monetary value to all forms of knowledge.

*Submission, Dr Louise Noble, Lecturer in English, School of English, Communication and Theatre, University of New England*

Much of the impact of research is on other researchers. Here, differences between disciplines are not marked. Innovation in a discipline, as reflected in published research outputs (whether academic texts, artistic products or any of the other forms of outcome), is quickly picked up 'in house' by other researchers. Extensive networks of publications, conferences and disciplinary associations facilitate the diffusion of new research, new knowledge and innovation, whether artistic or scholarly. Indeed, the perceived importance of a particular journal, organisation, venue or other locus for the dissemination of research within the academic community is one of the principal bases for the judgment of research quality and impact by academic peers.

The situation is more complex when one examines the way in which research has impacts on the broader society, and it is here that disciplinary differences are more marked. The survey of research end-users conducted for this study found that the perceived purpose of the research profoundly influenced their criteria for the judgment of its value, and such purposes could be categorised broadly according to timescale:

*Interviewees spoke of 'professional' and 'academic' research. They valued these kinds of research differently, with their preference for one or the other being closely related to their workplace time-constraints. Where lead-times and policy horizons are short (e.g. as when driven by ministerial imperatives), 'professional' research, commissioned on contract and with a precisely focused brief, was preferred. In these circumstances, higher value is placed on currency, speed of delivery, plain English write-up, and immediate usability. Where lead-times or policy horizons are longer (as in the preparation for major exhibitions or displays in, say, the Australian War Memorial, the National Gallery of Australia or the National Library of Australia), higher value is placed on methodological rigour and durability of research findings.*

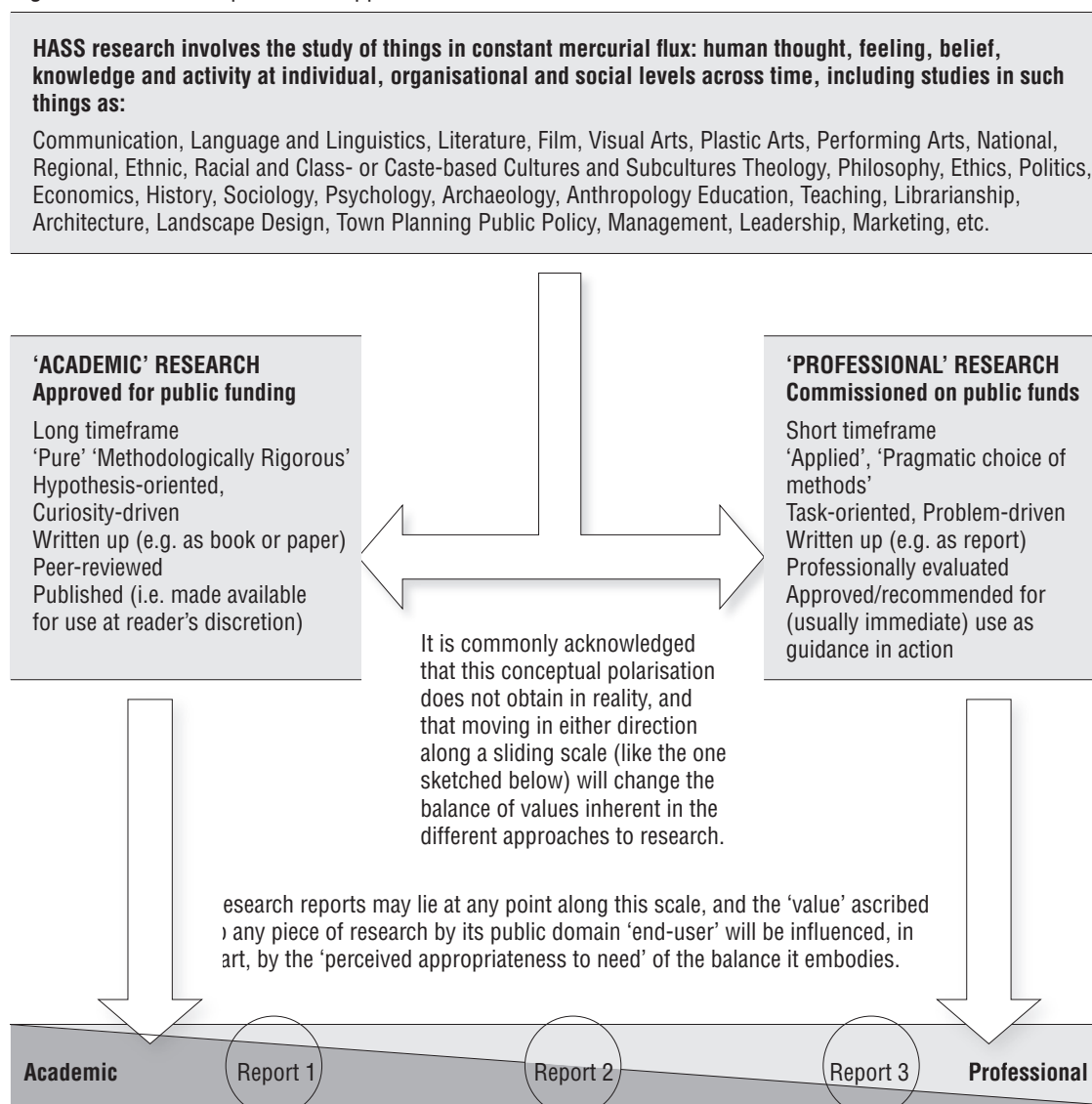
*The phrases 'immediate impact' and 'long-term effects' were raised in some interviews. It was suggested in those interviews that the word 'impact' as used in the enquiry's title 'The quality and impact of publicly funded HASS*



research' might seem like an invitation to pragmatic end-users who suffered tight deadlines to emphasise the value they have for 'professional' research and underemphasise the value they have for 'academic' research.

The survey of end-users (Appendix B) analyses this axis between 'professional' and 'academic' research, and Figure 3 maps it in more detail. Figure 4 attempts a detailed mapping of the 'impact points' of research, particularly on organisations in the wider community.

**Figure 3:** A common spectrum of approaches to research

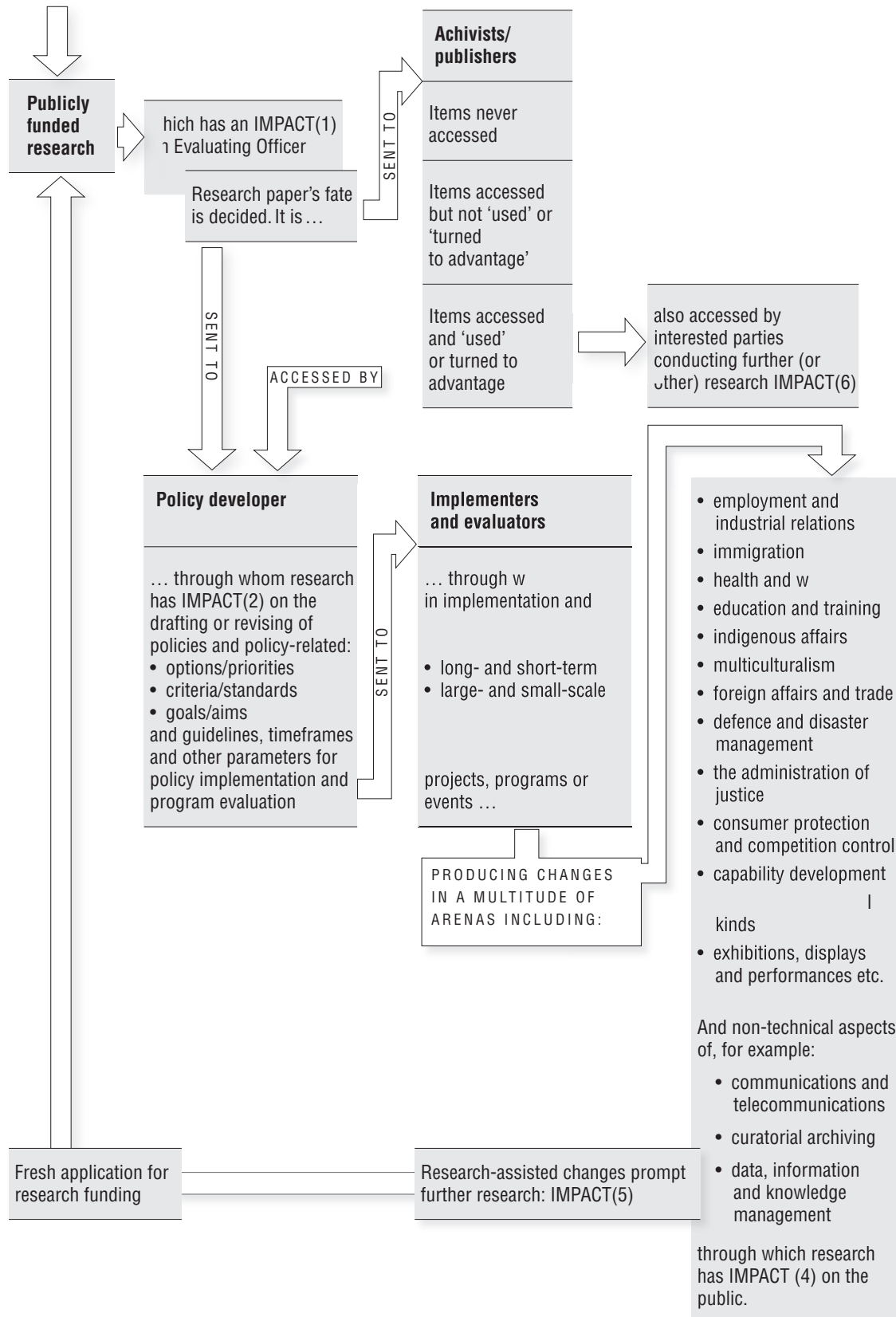


This analysis gives a good account of how research impacts, in a pragmatic and utilitarian sense, on organisations and stakeholders. However, as many submissions to this study have emphasised, it is more difficult to capture the less tangible value of research in the HASS sector. What is the value of having an informed society? What is the value of critical analysis of significant events; of public debate—informed rather than ignorant—on economic, civic, ethical and cultural values? What is the value of 'art' to society—can it not be argued that a society with art is *per se* a better society, without having to ascribe some secondary, instrumental value to art?

These questions are not easily answerable in any meaningfully quantifiable way. Nevertheless, they remain 'the most important questions of all'; they have to be measured in any research assessment process, and informed judgment has to play a prominent part.

**Figure 4: Mapping 'impact'**

White Papers, budgetary issues, policy proposals, and manifest needs or opportunities give rise to ...



The remainder of this section unfolds the conceptual ‘map’ of research presented in Figure 2 to examine some specific issues in the research process in each of the three divisions of the HASS sector. A broad map the research process in the humanities is presented in Figure 5; Figures 6 and 7 show the process in the creative arts and the social sciences, respectively. The maps display the process of research from left to right: from origin, through production, to output/outcome.

The first step might be to develop headings to capture the diversity of undertakings—a ‘Thesaurus’ (to use museum terminology). ‘Outcomes’ might be treated in the same way.

If I had to use a specific example [to illustrate the diversity of ‘impact’], it would be academic research and professional archaeological investigation into Duntroon Dairy, probably Canberra’s oldest remaining building, which resulted not only in articles, but in an exhibition and community programs which had a major influence on the subsequent restoration of the building and public awareness of its existence.

*Submission, Ms Jill Waterhouse,  
Canberra Museum and Gallery (emeritus faculty, ANU)*

## 4.5 The humanities

In some ways, the humanities display the most ‘traditional’ structure of research generation, production and use. Within each box in Figure 5, existing indicators of, or criteria for, quality and impact are identified in italics.

Three broad categories of outcome are identified: scholarly influence and dissemination; the broader community’s access to scholarly research; and education. Notably, only the last of these is not subject to some form of direct or indirect quality assessment.

It is worth reflecting for a moment on ‘education’ as a research outcome. Traditionally, universities have engaged in both research and teaching, and have argued for the inseparability of the two. Several submissions to this project argued for the importance of regarding education as a research outcome. Its nature is inextricably informed by the research conducted by the staff who teach and by the students. Perhaps the most common way in which students access the humanities while at university is through the completion of a Bachelor of Arts degree. The qualities of critical inquiry, and the ability to research, formulate and express an argument, are among the hallmarks of an Arts degree, and are also the characteristics of a research mindset. Thus, students emerge from universities equipped not only with the knowledge gained from studying with senior researchers in the field, but also in many cases with the skills and attitudes required to conduct research themselves.

In Section 5, this report argues for the inclusion of a third category—‘capability’—to sit alongside ‘quality’ and ‘impact’ in assessments of the overall value of research. A strong research culture not only produces significant research outputs, but also, in universities, produces a body of graduates with the research skills that enable them to go on being critical, inquiring knowledge-producers once they have left academia. This contribution to enhancing the abilities and capabilities of Australia is one of the important contributions of a healthy research culture, and one that must not be undervalued in overall assessments of research. Various indicators, such as research degree completions and postgraduate destinations, are useful as indicators of research capability.

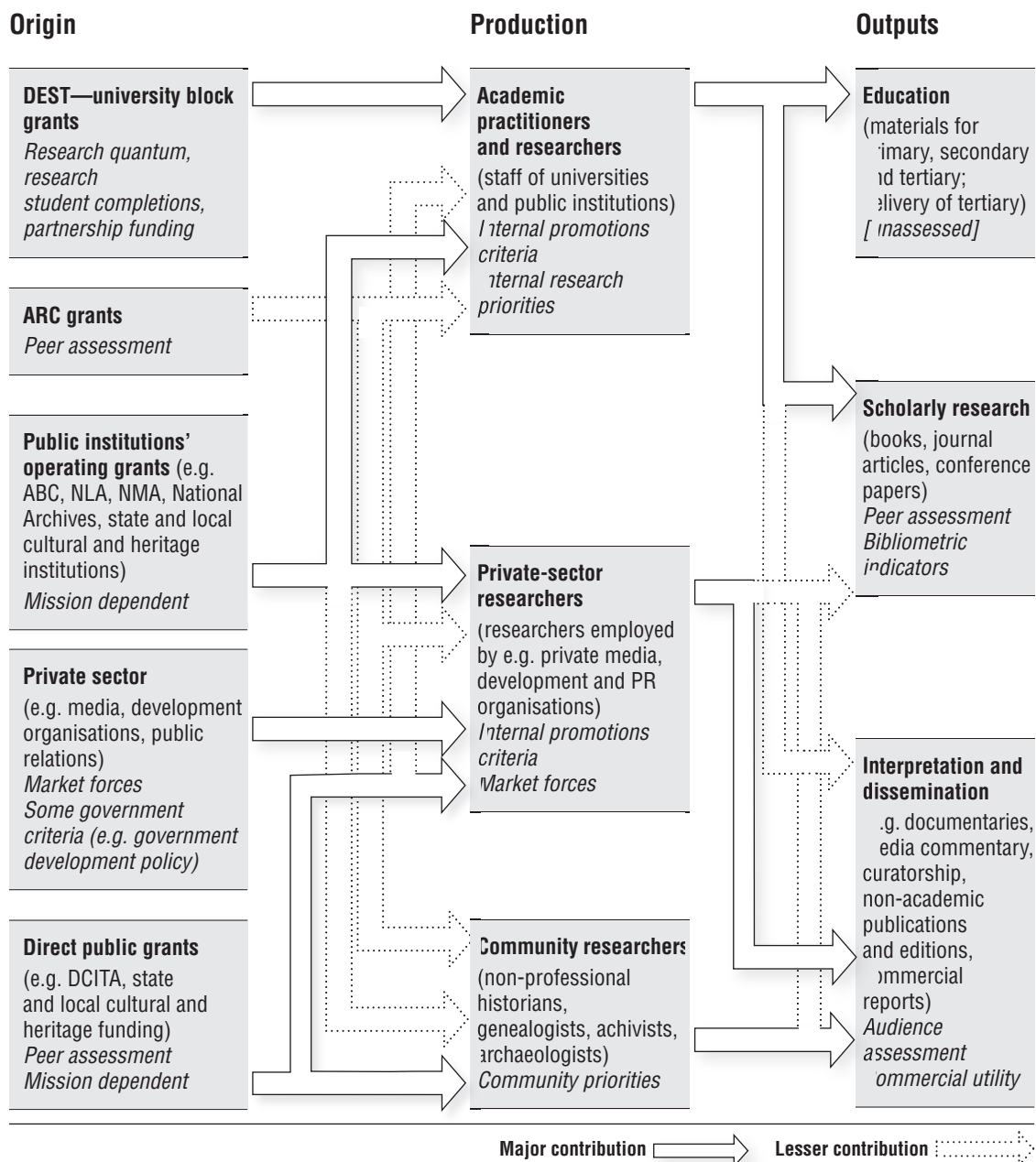
Lying behind [the CHASS survey questions] lies another question: What is the intrinsic value of the humanities? The value of the social sciences is easy to see: they help us understand the context of our lives; but what about literature, plastic arts, music? I suggest that their value lies in two outcomes. They bring before us beauty—beauty of language, of sound, of vision. Beauty is a good in itself, it is not instrumental; would anyone disagree with that? The other outcome is that they offer us insights. That’s not the same as ‘understanding’; a poem can stab us to the heart, move us to joy or tears, without

extending our ‘understanding’—indeed, it can contradict our understanding. A painting or sculpture, or indeed a beautiful person can be simply a delight to look upon, can stir pleasurable and grand emotions with no cerebration required; and similarly for music.

Now—your questions relate to research into the studies of beauty and insights: how can that be assessed and valued? I suggest it is important to keep the objects of the research, and their intrinsic value, separate from the value of the research. The value of the research is in the extent to which it helps us to appreciate the beauty and insights of the works studied. A poem of Gerard Manley Hopkins is harder to comprehend than one of Tennyson; the painting *Blue Poles*, or much of James Gleeson, is harder to get inside than a landscape by Tom Roberts; most of us need help to get the insights, see the beauty that these creations offer to those who can appreciate them.

*Submission, Emeritus Professor Richard Johnson, Australian National University*

**Figure 5:** The humanities



## 4.6 The creative arts

Traditionally, the most common difficulty in dealing with work in the creative arts disciplines as research has been in accommodating the creative outputs of artists' work within conceptual frameworks geared more to traditional scholarly publication. As long ago as 1996, the Department of Employment, Education, Training and Youth Affairs addressed this problem, commissioning the Strand Report to consider the status of creative work as research.<sup>8</sup> The reporting mechanisms for university research output were temporarily revised along the lines recommended in that report, but a simple quantitative approach to research accounting proved inadequate to handle the complexities of the issue.

This report is not the place to reopen the debate, but it is worth noting that the introduction of a new national approach to research performance in the proposed Research Quality Framework provides an opportunity to deal with the definition of 'research' in the creative arts. For the purposes of this report, the nature of the process for the generation of artistic output—its origin, production and outcomes—broadly parallels those for other types of scholarly research (Figure 6).

It is also worth noting that, in any research assessment process based on peer evaluation, there is little difficulty in accommodating artistic output alongside other research forms. This was done without contention in the United Kingdom Research Assessment Exercise.

The problem lies with both the nature of the output and the definition of research. Arts researchers get caught between two stools for on the one hand not being 'commercial' and on the other, when they are 'commercial' (as in a film or piece of software), the output not counting as research. Despite the diversity of potential outputs there are commonalities with more 'traditional' research and it really shouldn't be hard to adapt current standards to these areas. Any production which passes the editorial process in order to be showcased, either in the broadcast media, as part of an exhibition, or part of a festival, has effectively been 'peer reviewed'. This also holds for commissioned pieces in the print media which go through a rigorous editing process before getting into the public domain, and attest to the academic's standing and reputation. The idea that a scholarly paper is research whereas a consultancy report is not when both may be based on the same data is absurd. The focus should be not on the type of output but on the quality of knowledge that goes into it.

*Submission, Dr Gail Phillips, Associate Professor,  
Chair of Media Studies, School of Media Communication and Culture,  
Director, Research and Development, Division of Arts, Murdoch University*

In the course of this research project, a discussion was begun by some participants seeking to outline processes for the assessment of artistic work as research. A working document produced by members of the artistic research community recommends that three questions be posed of an artwork to determine its status as research:

*When looking at the output of arts faculties, we first have to decide whether the output should be subject to assessment as research output. If we're using an OECD-based definition of Research as 'creative work undertaken on a systematic basis in order to increase the stock of knowledge,' and to increase understanding, then we need to ask:*

- *Does the output contribute to the stock of knowledge? Does it undertake traditional scholarly research on the arts?*
- *Does the output contribute to the stock of understanding, by providing new artworks such as plays, creative writing, new compositions, new art, designs?*
- *Does the output provide a new or innovative perspective on the current stock of knowledge—for instance editions, new and novel interpretations of existing material?*<sup>9</sup>

If the answer to these questions is 'yes', the artwork constitutes research. If, on the other hand, the output represents existing material from within the stock of knowledge with limited innovation, it is more properly described as professional practice, and is most commonly funded either from the private sector or through publicly funded support for arts practice (nationally, most commonly from the Department of Communications, Information Technology and the Arts) rather than for research (most commonly from DEST).

<sup>8</sup> Strand D (1988). *Research in the creative arts*. Department of Employment, Education, Training and Youth Affairs, Canberra.

<sup>9</sup> Schippers H, Sears M and Rowley S. The payback model and research in the performing and creative arts. Unpublished discussion paper.

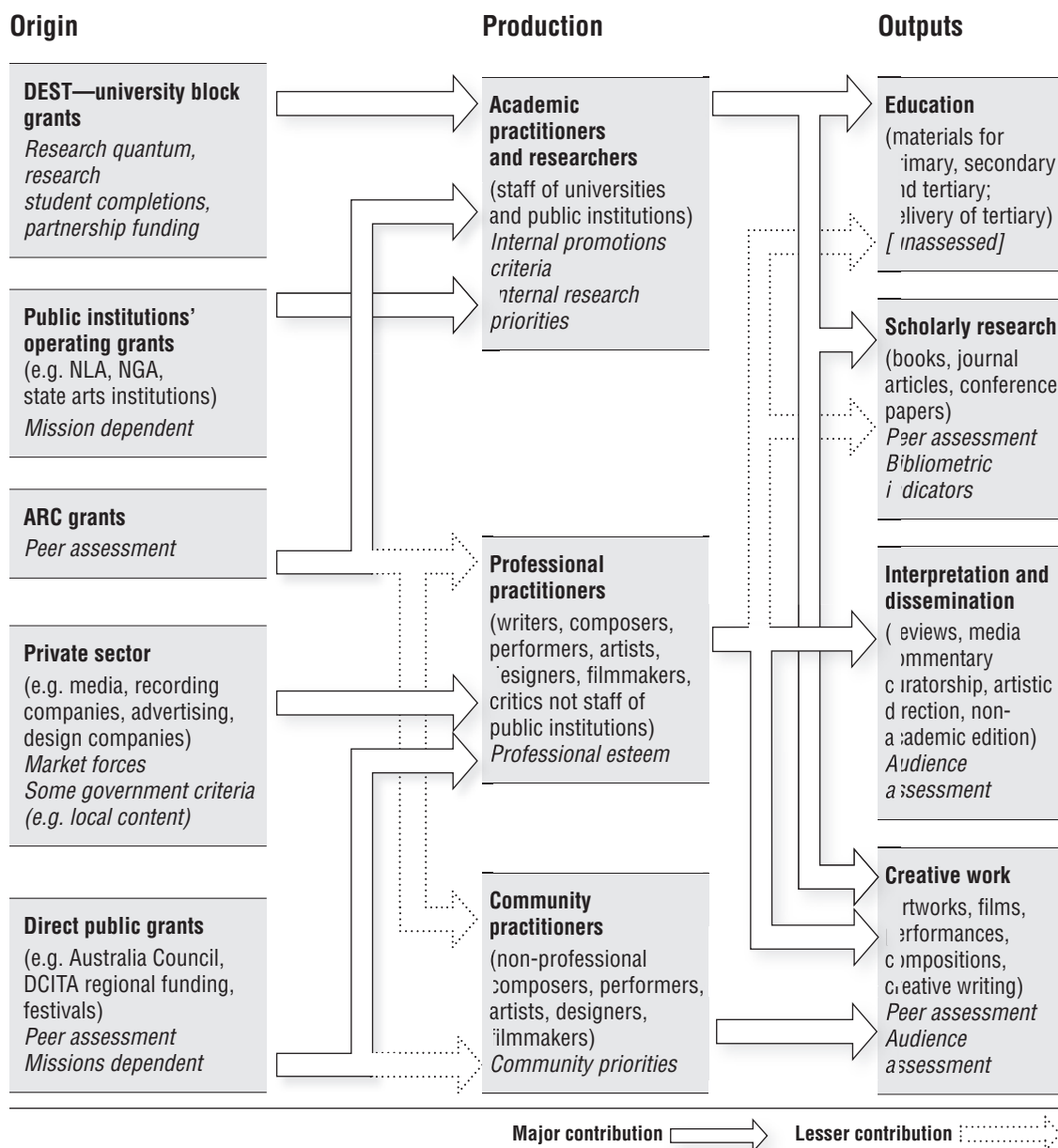
A clearly structured system of peer assessment might well answer some of the problems in this area. Much of creative practice/research has a level of performance ephemera about it and is difficult to capture accurately in an archival state. Therefore a panel of peer assessment, plus reports and submissions would help assessors to evaluate the integrity of the research and work. These reports should be part of the end of the research project thus allowing practitioners to concentrate on achieving their stated creative outcomes as the project proceeds.

The evidence of non-traditional outcomes should be archived in an accessible form so that future researchers can access and refer to these. This might involve a revisiting of copyright regulations since many creative works might involve the partial use of copyright held elsewhere (for example, an innovative production of a copyright piece, might be licensed by the artists for performance but not for recording. Such exclusion also prohibits a single recorded copy for archival purposes and fair dealing for research).

The creation and cataloguing of such a multi-modal archive of creative and performance outcomes might itself constitute a major research project since it involves the interrogation of categories of information required for retrieval of resources held.

*Submission, Faculty of Communication and Creative Industries, Edith Cowan University*

**Figure 6:** The creative arts





## 4.7 The social sciences

As with the humanities, the typical research outputs in the social sciences are most often in the form of scholarly publications. As can be seen from Figure 7, there is a significant engagement with policy formation, and the requirements of meeting the needs of organisations that directly use the research must influence criteria for the assessment of research value. In the minds of many, there is a real or perceived trade-off between ‘impact’ and ‘quality’ if the latter is measured by some set of generic, abstract attributes of all academic research.

The diversity of research outputs is not always acknowledged by the bodies that give out the ‘brownie points’ (for instance, the DEST research quantum criteria). For example, on my last and current research projects we chose to write public reports that are widely distributed and also will be accessible on websites. But my institution gets no quantum for this. In many ways in our organisational culture, there is too much emphasis on publishing in academic journals because of the assessment system and therefore publicly funded research is not always finding its way into the public realm.

Diversity of output also requires diversity of writing styles. And yet we train our new graduates to write only in an academic style which again is often very inaccessible to a wider public.

*Submission, Dr Melissa Butcher, Research Institute for Asia and the Pacific,  
University of Sydney*

The end-user survey revealed a distinction between ‘professional’ and ‘academic’ research in the minds of users. This suggests that the criteria for ‘peer assessment’ of scholarly research might be different from, or even at odds with, the criteria for the assessment of policy impact or commercial utility in the bottom two boxes in the ‘Outputs’ column of Figure 7.

The complexity of measuring impact in the social sciences is highlighted by the submission of one of the participants in the case-study exercise undertaken for this project (for a full account of this process, see Appendix C):

*The connection between research and impact may not be easy to trace. An invitation to speak, for instance, is offered because of reputations based on research, in the expectation that the talk will provide insights that may be useful.*

*By impact we refer to the form in which staff, through their research, have interacted with the community, with government or with international bodies. We contend that several indicators illustrate that research is noticed and of interest.*

- *Invitations to address government and public service departments*
- *Writing reports for Governments or NGOs*
- *Addressing public bodies*
- *Contract research*
- *Participation in documentaries*
- *Membership of Boards, Panels*
- *Continuing advice/consultancies with Government or NGOs*
- *Establishment of institutes to carry through research ideas.<sup>10</sup>*

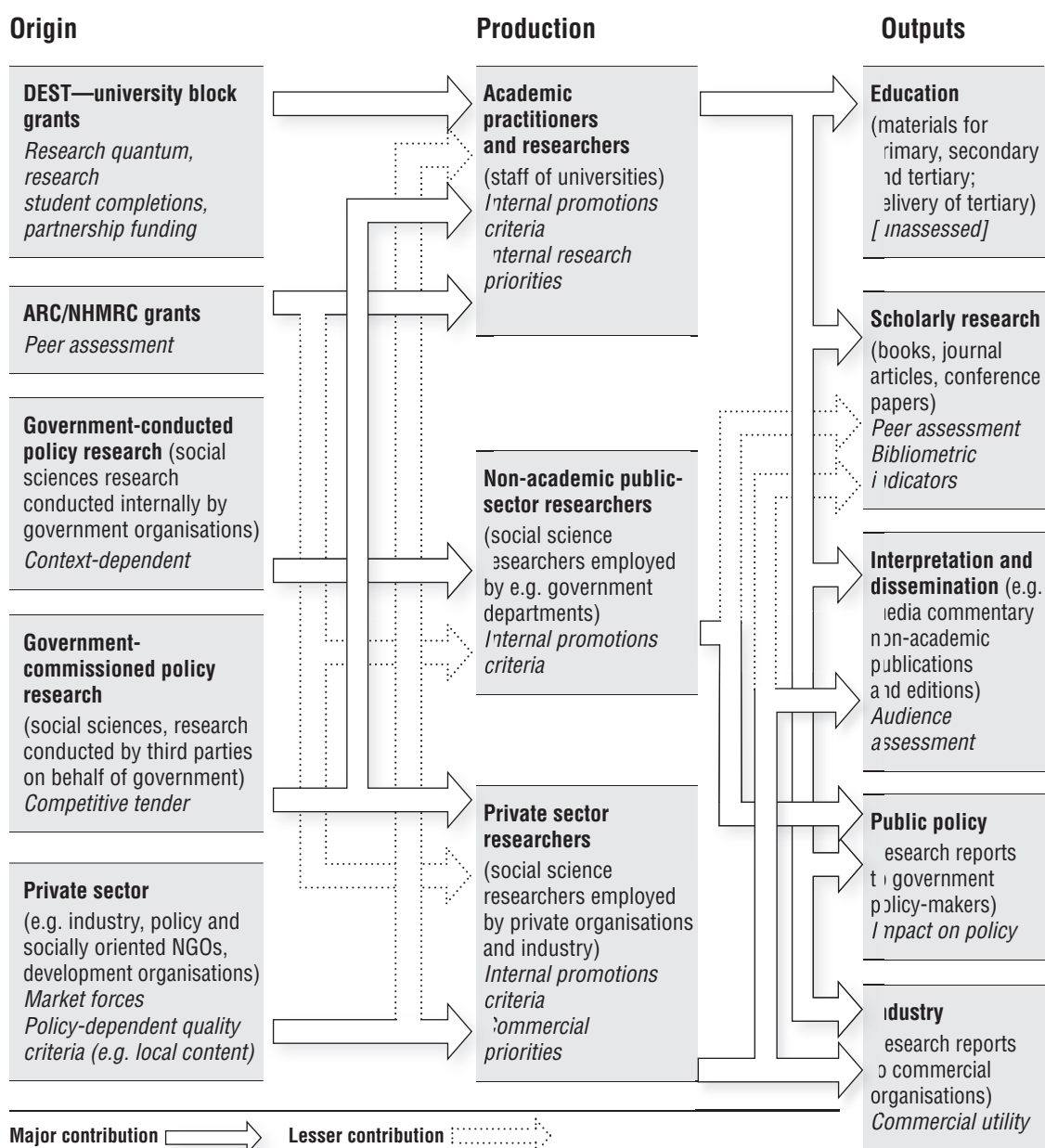
This broad definition of ‘impact’—which includes the influence of individuals, and the expertise of groups, above and beyond the influence of their publications *per se*—is characteristic of some of the issues that arose when the notion of ‘impact’ was considered in the submissions and in the case study trial. The introduction of ‘capability’ as a separate measure would capture these broad and holistic attributes of a successful research unit.

<sup>10</sup> Case for Excellence, Politics and Public Policy, Griffith University.

[In social policy research], the question of useability is critical. Much research is conducted in the name of social policy that has no bearing on political or policy decision making but is funded through selection mechanisms that entrench standard ways of analysing and viewing problems. What constitutes good social policy research products? Material that contributes to the rationing problems of government, that informs decision-makers about resource allocation emphases in trustworthy ways, and so on, would be regarded as good products within government. At the same time, research that holds a mirror up to the practices of policy (e.g. the anthropology of audit or policy) is also highly valuable, and does not have to be instrumental to be worthy. Here, critical acuity is more important ...

*Submission, Associate Professor Tess Lea, Director, School for Social and Policy Research,  
Institute of Advanced Studies, Charles Darwin University*

**Figure 7: The social sciences**





## 5 Indicators: quality, impact and capability

Conventional measures of research quality and impact have been narrowly conceived and used, and have often been too blunt in their application to HASS disciplines. This is because their design is based mainly on experiences in the sciences, and the net effect is akin to comparing apples with oranges. This is an international problem, and any solution requires a new approach. This report proposes a new and different model for assessing research quality and impact. It adds a third component into the mix: capability, which takes into account the 'people' factor.

The model has an immediate potential application: to inform any Research Quality Framework (RQF) and Research Assessment Exercise (RAE) that might be employed in Australia.

Most evidence indicates that evaluation of quality is best achieved through appropriately informed qualitative judgment. The overwhelming majority of respondents to the consultative processes undertaken for this project favoured a research assessment process based on peer assessment. Many made reference to existing practices of peer evaluation in other countries (for example, the United Kingdom RAE) and in Australia (for example, in awarding ARC grants or in making judgments about academic promotions).

No standard quantitative indicators can be relied upon in isolation, particularly in the humanities, arts and social sciences. Existing bibliometric practices in the sciences, which are arguably helpful in making judgments about quality, do not capture the variety of research outputs in the humanities and social sciences (for example, books, documentaries and policy reports), and their use would clearly be absurd in the creative arts. Indicators derived from success in research funding fail to capture the very different sources of such funding across the HASS sector, and the very different costs of conducting research in different disciplines. Obviously, however, such indicators can be useful evidence to help inform the judgments of appropriately constituted panels.

The comprehensive review of the literature and current practice in assessing the quality and impact of HASS research (Appendix A) provides evidence for this conclusion, which accords with the feedback received from the great majority of respondents to this review.<sup>11</sup>

We think that peer review by discipline experts is still the most appropriate form of assessment of quality for research in arts, humanities and social sciences. It needs to be pointed out that there are deeply divided theoretical camps in these fields, so any appointment of such experts would have to take this into account so that an assessor from one theoretical school didn't rubbish the work of someone from a different school of theory.

Attention needs to be drawn to the inadequacies of citations and indices. Citation indices just don't work in the humanities since they don't register books, and don't pick up on work in other languages than English.

A list was received a couple of years ago for comments. [It] had all the journals which were deemed by a DEST commission to be internationally refereed journals in humanities and social sciences. It was found that it was far from complete. It is highly doubtful whether it would be possible to have a complete list of refereed academic journals in the English-language and the non-English-language world.

*Submission, School of Languages and Comparative Cultural Studies,  
University of Queensland*

<sup>11</sup> See Appendix A for a discussion of the shortcomings of quantitative, citation-based approaches to research assessment for the HASS sector.

If the current measures of research quality that focus predominantly on scholarly publications, research degrees and research income are inadequate, what should a new set of measures include? As a starting point, it should incorporate competitive research grants ('grants'), scholarly publications ('publications'), use and citation of academics' published research by other academics ('citations'), qualitative assessment of academics' research by other academics ('peer review'), and research degree student training, support, supervision and completion ('research degrees').

The new set of measures should also include additional measures of particular relevance to the HASS sector. These include citations by experts other than academics, such as citations of research in the social sciences in governmental public policy and law reform documents, and the citation and use of academic research by judges in applying, interpreting and even developing the law. Nor are total outputs the only or best measure of research quality, as outputs are relative to inputs and quantitative outputs are not the same as substantive outcomes. There is also a connection between research quality and research impacts and benefits, and hence a connection between fully developed measures for assessing both.

This section discusses the three principal elements of a new system of research evaluation: quality, impact and 'capability'. It outlines the steps Australia can take to lead international best practice in the assessment of research quality and impact, and identifies more accurate and appropriate measures for assessing HASS research and better indicators for assessing and comparing the value of both HASS and non-HASS research. The section discusses indicators of research impact and benefit as well as indicators of quality, and develops processes for assessing the worth of all research and its significance for modern nations and their economies.

## 5.1 Quality

Quality can be assessed through a variety of types of evidence, including peer evaluation of the work itself; a number of indirect indicators, such as citation indices and researchers' success in obtaining grant funding; and some broader indicators of peer esteem, such as membership of learned or professional bodies. This evidence of quality in HASS research is conceptually identical to evidence of quality in other disciplines, such as science, technology, engineering and medicine.

Types of evidence that might be submitted for assessment of research quality are listed below, ranging from direct assessment of the research work, through to more indirect measures of researchers' esteem and of the measurable academic utility of the work as evidenced in citation indices.

### 1. The research

It might be assumed that the fundamental evidence to use in assessing quality lies in the work itself. This is the cornerstone of the United Kingdom RAE, and the assumption was tested in the case-study trial described in this report. Assessors were not initially provided with copies of the research—one aim of the exercise was to determine how successfully and confidently they were able to make a judgment based purely on indirect evidence.<sup>12</sup>

The feedback from assessors was that they would prefer to receive only summary documents:

*Having taken part in several research assessment exercises in the past I have never found receiving copies of actual research outputs (papers, chapters, etc.) very useful. It is mostly just time consuming. Knowing the typical journals, conferences and forums for research presentation in the respective fields is a more efficient method of assessing quality. It is admittedly slightly less accurate but the difference would be negligible.*

*Assessor, Architecture and Design*

<sup>12</sup> For full details of the case study process, see Appendix C.

## 2. Full lists of scholarly publications, or equivalent research outputs

A list of publications might simply be provided in the form of a list of scholarly publications as reported to DEST in annual university research reports.

A more sophisticated approach would be to provide a more detailed and annotated list in order to allow a weighted assessment of scholarly publications according to their focus on, and reference to, national versus international research, disciplinary versus cross-disciplinary insights, theoretical versus applied research, academic versus non-academic audiences, and so on.

The approach taken by the United Kingdom RAE is to ask researchers to identify the four best publications in the period of assessment. Feedback from assessors in the case-study trial suggested that such an approach—the selection of the *n* best publications—would be helpful by allowing assessors to focus on the crucial items. However, having full lists of research outputs available was also seen as important to contextualise the research under assessment, and to provide a comprehensive account of research activity against which the very best performances could be assessed.

Appendix A refers to the importance of non-written outputs, particularly in the creative arts:

*For example, Strand's definition of publication in the creative arts extends far beyond the written word. He noted that publication in the creative arts includes:*

- *public performance for dancers, actors, choreographers, musicians, playwrights*
- *building or manufacture for architects and designers*
- *written works for academics, researchers and creative writers*
- *exhibitions for visual artists, craftspeople*
- *computer software for designers, musicians, visual artists and so on (1998, 55)*

*Again, this is a list that can be extended and refined ...*

From the above considerations follows that a peer review approach is necessary. This is also the state of the art of science studies (including bibliometrics, which states that its outcomes must be interpreted by peers). However, some important qualifications need to be made.

- a) It is important to draw the right boundaries between fields. The social sciences, arts and humanities are much more heterogeneous than any other discipline, which means that more and smaller fields need to be constructed and to be assessed by their peers. Too large and too heterogeneous research fields to be evaluated were one of the problems reported by assessors in the ANU Quality Review (Gläser and Laudel 2005).
- b) One feature of peer review that cannot be made go away is its normalising approach. Peer review likes the very good but not the exceptional, and tends to favour the mainstream. It is therefore important not to assign too much power to the evaluation procedure, even when it is based on peer review.
- c) While we favour peer review, the one existing blueprint (UK's RAE) has some serious problems. The major problem is that with conducting and publicising financially relevant performance evaluations at the departmental level, departments become 'profit centres' that contribute to the 'income' of the university. Under conditions of an all-powerful university bureaucracy which has simultaneously succumbed to corporate business models, an unhealthy pressure is put on departments and individual scientists. This pressure is particularly strong when a department is rated so low that it 'doesn't bring in university money'.

While production-line workers and clerks in industrial enterprises might work better under such pressure (though even this has been contested), the case has yet to be made that increasing pressure, frantic hiring and firing, and ubiquitous restructuring improve the quality of research and teaching. To our knowledge, there is no conclusive evidence that there is an improvement of UK's science that can be causally linked to the RAE.

Another potentially problematic aspect of a RAE-style evaluation is its selectivity. Many of UK's academic organisations have become concerned about the selectivity of the RAE and the ensuing threats to the diversity of the UK's research landscape.

Important changes of the RAE model we regard necessary concern the time frame, which should be widened (the recently introduced seven years between RAEs in the UK are good), and the selectivity of funding, which should be reduced. The more severe (frequent and significant) the consequences of the evaluation for the 'income' of the university, the more likely are bureaucratic frenzies within universities that are detrimental for research.

- d) Ideally, a peer review would be informed by quantitative indicators. However, we have already made clear that this is a difficult business in the case of social sciences, arts and humanities. Any indicators would be partial and would be likely to create a very selective attention to these indicators. Universities would begin to learn 'how to play the game', a process which is likely to distort research by over-emphasising the aspects that are measured.
- e) Whatever evaluation procedure is going to be applied, it will have unintended consequences because universities and scientists will adapt their behavior to the conditions of funding. A careful independent analysis of these consequences is necessary.

*Submission, Dr Jochen Gläser, Dr Grit Laudel, Research Evaluation and Policy Project,  
Research School of Social Sciences, Australian National University*

### 3. Citation data

Another measure of evidence of research quality is citation data. These data include:

- academic citation of academic work, as recorded in available citation indices (predominantly for science and technology)
- academic citation of academic work, as recorded in searchable print and electronic sources in the absence of discipline-specific citation indices<sup>13</sup>
- academic citation of authoritative academic books and texts, not otherwise recorded in searchable print and electronic indices but available by other means (e.g. search for academic author names in the indexes and references of electronic copies of books held by publishers)<sup>14</sup>
- academic citation in one discipline's research and scholarship of the work of academics in other disciplines (as a means of recognising and emphasising the value of cross-disciplinary insights for other disciplines).

It is well known that citation counts are very flawed measures of quality. For example a very poor but very provocative article is likely to be cited frequently. Also it would be necessary to apply a sophisticated weighting factor to take account of citations in overseas journals as well as Australian ones. It would be better to have well respected discipline panels sample submitted research portfolios. Such a measure of quality might also serve to dampen the over-publication caused by simple counting, and also limit the proliferation of low-quality journals.

*Submission, Associate Professor David Lemmings, Research Development Manager,  
Faculty of Education and Arts, School of Liberal Arts, University of Newcastle*

<sup>13</sup> Web-based sources such as 'Google Scholar' (<http://scholar.google.com/>) and, for law, WESTLAW, LexisNexis and AUSTLII ([www.austlii.edu.au](http://www.austlii.edu.au)) can be used to track citations of an academic's work in electronic databases of journal articles.

<sup>14</sup> As most books are still not publicly available in web-based form, this might require discussion and cooperation with commercial publishers, most of whom now use electronic copies of books for proofing, printing and updating purposes.

In the trial assessment conducted for this project, citation data of some kind were presented for two of the disciplines (politics and public policy, and law). In each case, these were highly discipline-specific. It seems evident that the relevance and relative importance of citation data, and thus the weight they should be given in an overall assessment of quality, will be very specific to each discipline.

Appendix A discusses a variety of possible means for extending the notion of ‘bibliometrics’ to make it more usable for HASS research assessment. In particular, ‘sociometrics’—the embeddedness of research outcomes in social and cultural practice—has been explored in the Netherlands and may be useful in a number of HASS disciplines, particularly the social sciences.<sup>15</sup> Such discipline-specific quantitative data are potentially useful sources of evidence to support qualitative judgments about research quality.

For the purposes of this project, as an example of discipline-specific citation indicators, the possible sources of citation data, broadly interpreted, have been explored for just one discipline: law. The following potential sources were identified:

- citation of Australian HASS research (i.e. both legal and non-legal<sup>16</sup>) in Australian court and tribunal decisions, as a measure of Australian judicial assessment of research quality<sup>17</sup>
- citation of Australian HASS research in international court and tribunal decisions, as a measure of international judicial assessment of research quality<sup>18</sup>
- citation of Australian HASS research in published speeches by Australian judicial officers, as a measure of Australian judicial assessment of research quality<sup>19</sup>
- citation of Australian HASS research in published speeches by international judicial officers, as a measure of international judicial assessment of research quality
- citation of Australian legal research in written submissions and oral arguments in court proceedings<sup>20</sup>
- citation of Australian HASS research in governmental, parliamentary, judicial and other official sources.

This list indicates at least the possibility of developing sources of broad citation data for selected disciplines, which could be then presented to assessors as indicators (but not evidence) of quality.

However, there are significant difficulties in relying on bibliometrics and citation data as a useful tool for assessing the value of HASS research. These are outlined in detail in Appendix A, and include the following:

- Citations do not reveal the quality of the publication—citations may be positive or negative, and a paper may be highly cited because its findings are contested.
- A relatively small proportion of HASS literature appears in journal form (as low as 4% for law and 11% for the arts)—many of the outputs of the HASS sector are not captured by the standard bibliometrics.

Donovan quotes Royle and Over, who ‘make the strongest criticism relating to the social sciences in Australia when they argue “reaching conclusions about the relative research productivity of individuals

<sup>15</sup> For more detail, see Appendix A.

<sup>16</sup> The point here of including reference to legal and non-legal academic research is to highlight the fact that judges use, cite and quote a range of legal, economic, historical and other academic sources, rather than to open a debate about the capacity of judges to assess the quality of non-legal research. Judges are public officials within an arm of government engaged in the administration of justice, and their reference to published material counts for something as a measure of quality, just as the references by government reports and policy documents to the research of academics counts for something.

<sup>17</sup> All decisions from at least the past decade for major Australian courses and tribunals are available and searchable electronically for author citations, either via the websites of particular courts and tribunals or from a portal such as Australia’s AUSTLII.

<sup>18</sup> All recent decisions for major appellate and international courts are available and searchable electronically for author citations, either via the websites of particular courts and tribunals or from portals such as the United Kingdom’s BAILII ([www.bailii.org](http://www.bailii.org)), the United States’ LII (<http://www.law.cornell.edu/>), Canada’s CanLII ([www.canlii.org/](http://www.canlii.org/)), the Pacific islands’ PacLII ([www.pacii.org/](http://www.pacii.org/)), and WorldLII ([www.worldlii.org](http://www.worldlii.org)).

<sup>19</sup> Speeches by judges are routinely available from their court websites. For example, speeches by High Court judges are accessible under ‘Speeches’ via the High Court’s webpage ([www.hcourt.gov.au](http://www.hcourt.gov.au)).

<sup>20</sup> For example, court registries hold copies of lodged written submissions, while references to academic research in the course of oral argument before courts appear in court transcripts (e.g. High Court transcripts of proceedings accessible via AUSTLII).

or universities, or seeking to compare disciplines or nations, by reliance solely on the database used to compile the [Institute for Scientific Information] source indexes is not a worthwhile exercise”<sup>21</sup>. Donovan’s commentary in Appendix A of this report summarises the criticisms of using standard citation-based indicators to measure HASS research, although she does suggest that it should be possible to develop ‘new non-standard measures of research quality using “metrics” suitable for both HASS and STEM’.

#### 4. Evidence of academic standing and esteem

Indicators of esteem within one’s discipline or profession will vary by discipline, but might include:

- memberships of editorial boards, academic/professional bodies and other organisations based on research expertise and reputation, such as learned academies
- invitational and keynote addresses for academic, professional, business or industry conferences and seminars
- prizes, fellowships and awards.

Of course, indicators of academic standing and esteem are subjective, and their interpretation is highly dependent on the discipline. There are no meaningful ways to aggregate esteem data to provide comparable numerical measures, particularly between academic fields.

However, esteem indicators are among the most trusted by academics in current quality assessment processes, such as those for promotions and grant awards. They are succinct condensations of quality judgments already made by peers in the field. They could prove useful as part of a suite of evidence to be presented to a panel making a qualitative evaluation.

#### 5. Statistics (number and value) and qualitative accounts of successfully acquitted contestable research grants received by the researchers

The successful acquittal of research grants provides a further indicator of quality. Within the framework that this report proposes, the successful receipt of a grant that is yet to be completed is better viewed as an indicator of future capability than of past performance.

Once again, caution is needed before using such data comparatively in their raw state. The numbers and sizes of research grants available vary hugely by discipline. Often, this reflects the differences between disciplines in the cost of conducting the research. In other cases, it reflects the level of investment in particular disciplines by government or industry, made on the basis of political or economic priorities rather than quality *per se*.

However, within a discipline, a unit’s success or failure in attracting contestable funding can provide an indicator of quality.

<sup>21</sup> Royle P and Over R (1994). The use of bibliometric indicators to measure the research productivity of Australian academics. *Australian Academic and Research Libraries* 25(2):77–88

## 5.2 Impact

Assessing the ‘impact’ of research in our sector is both more difficult and, arguably, more important than for research in the scientific disciplines. Beyond its impact in the academic world, humanities, arts and social sciences research may contribute to policy development; inform the public on issues of social and cultural concern; make scholarly, artistic and commercial contributions to the nation; and, through research-based education, shape the thinking of the next generation of Australians.

*Submission, Associate Professor Walter F. Veit, Department of German Studies,  
School of Languages, Cultures and Linguistics, Monash University*

The stark reality confronting any imminent national assessment of research impact and benefit in Australia is summarised by the Allen Consulting Group as follows:<sup>22</sup>

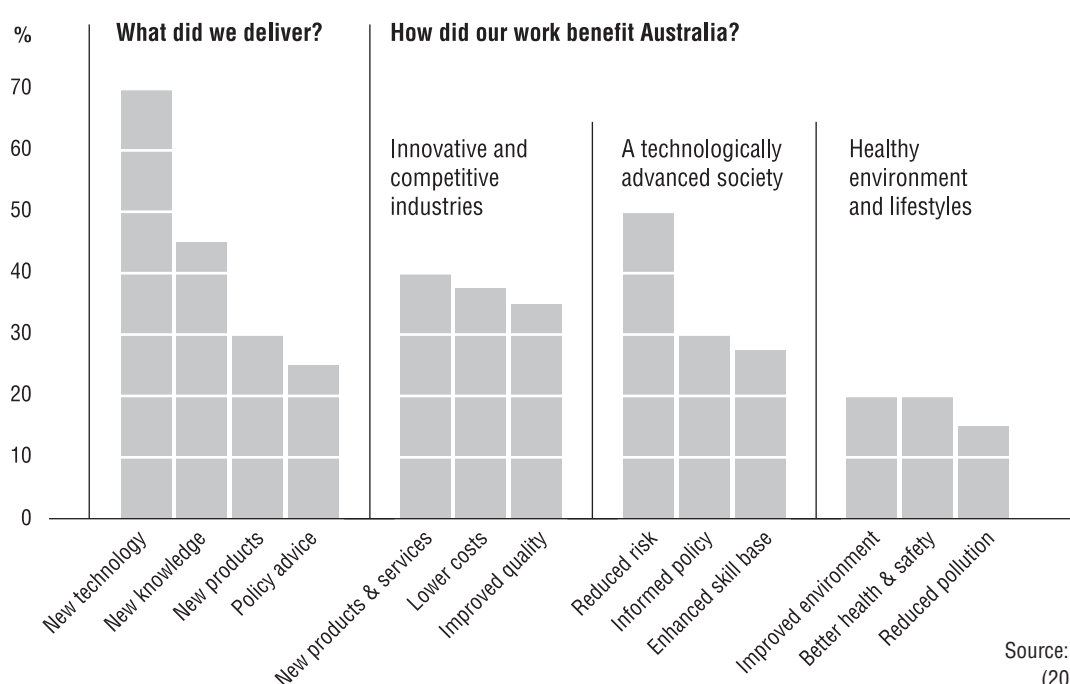
*The overall conclusion to be drawn from this review of current Australian and international practice is that attempts to collect data that capture the full range of ways that research can impact upon society are still evolving and that no comprehensive solutions to this task have yet been developed.*

Australia is not well served if the only or main audience for research is academics, and if the measures of research’s worth for funding purposes focus solely or mainly on peer review, publications and citations. None of this means, of course, that academic research and publication are not valuable, that academics are not an audience for research, or that academics are not important experts in assessing good research. But they are not its only audience, they are not its only beneficiaries, and they are not the only ones in society who assess or use research. Any system for measuring fully the quality and impact of research must include the academic arena, but also extend beyond it. Even within the academic arena, current measures of research output and outcomes—and they are not the same thing—are too blunt in design and application to capture all forms of research and its products and benefits.

A number of interesting models are available, including those of Australian science agencies. The Commonwealth Scientific and Industrial Research Organisation (CSIRO), the Australian Institute of Marine Science, and the Australian Nuclear Science and Technology Organisation use common indicators in measuring their respective impacts on commerce and government.

The CSIRO model is presented below and explained further in Appendix A.

**Figure 8:** CSIRO—Overview of selected achievements, 2003–04



Source: CSIRO  
(2004, 39)

<sup>22</sup> Allen Consulting Group (2005). *Measuring the impact of publicly funded research*, p. 52.



Although the precise methodology is not presented, we can see from Figure 8 that ‘outputs’ covers new knowledge, technology and products, and policy advice, while benefits accrue to Australian business, the technological advancement of society, and environment and lifestyle. The four main strands of CSIRO research, taken individually, produce different outcomes.<sup>23</sup>

The HASS disciplines can perhaps learn from such methodological trailblazing in the sciences.

Section 6 of this report presents a conceptual structure for grouping types of impact, from direct impact on academic peers through to the most diffuse benefits to culture and society.

The 2005 report of the Expert Advisory Group for an RQF lists ‘examples of possible research impact outputs’.<sup>24</sup> The list comprises:

- publication outputs
- editorship of international journals
- involvement in international learned societies
- invited international lectures / other international collaborations
- peer recognition: academy membership
- research income: competitive grants schemes
- research income: other grants schemes (including international)
- other research income (other contract research)
- international linkages / collaborations / memorandums of understanding
- visiting scholars and/or postgraduate students
- publication citations
- incorporation of research results into international/national policies, codes and/or practices
- research graduates employed in industry
- industry-funded research places
- academic–industry staff exchanges
- research students industry placements
- the holding of exhibitions and performances
- audience/attendances at exhibitions/performances
- media presence through articles, debates, coverage
- expert advice / submissions / panel membership at government inquiries
- patents
- commercial licences
- commercial uptake
- spin-off companies
- consultancies
- other.

Most respondents to this project enthusiastically support this expansion of research impact indicators beyond conventional (e.g. solely bibliometric and academic-based) measures. At the same time, many (if not all) of the listed examples are simply quantitative research outputs and not substantive research outcomes, and not all of them apply equally to the HASS and non-HASS sectors. The list is illustrative but not exhaustive. All these indicators are relevant, but none of them should be used in isolation or be given preference in any assessment framework in a way that unjustifiably advantages some disciplines over others.

<sup>23</sup> See Appendix A to this report.

<sup>24</sup> Research Quality Framework Issues Paper (2005), p. 24.



Most recent authoritative reports are moving towards broader and more accurate and comprehensive research measures that include, but are not limited to, admittedly important bibliometric indicators like academic citations. The Allen Consulting Group report<sup>25</sup>, for example, outlines for the HASS sector the following illustrative and hence non-comprehensive ‘collectable diffusion indicators that may relate to impacts’:

- presentations given at learned societies
- presentations given at academic conferences
- submissions made to government inquiries
- citations of research in government policy publications
- citations of research in court judgments
- articles published in the popular press
- research cited in articles in the popular press
- research from surveys of heads of policy sections in government departments as to whom they regard as ‘high-impact’ academic researchers
- the number of research students who are subsequently employed within government departments, ministerial offices and industry.

Feedback from the HASS sector supports (with several specific caveats) all these measures and more as suitable indicators. They need to be incorporated within a wider framework of indicators, such as the framework for measuring research impact outlined in this report. There are differences between the HASS and non-HASS sectors in the illustrative examples of research impact indicators outlined in the Allen Consulting Group report. These differences are manageable within a proper research management framework that genuinely compares like with like and makes justifiable adjustments for disciplinary differences in how research is used in society.

One major task for the government before the introduction of an RQF will be to develop, on a discipline-by-discipline basis, appropriate measures and evidence of research impact. The research project undertaken by CHASS to produce this report could prove a useful model for a comprehensive process.

Measuring outcomes rather than outputs is a very inexact science and often requires a resource base well beyond that available (my experience is with the human services which are taking a huge battering as they cannot often relate outcomes to specific interventions). We are often pragmatically restricted to outputs and perhaps should put our energies into broadening the types of outputs considered to be of value.

*Submission, Dr Wendy Earles, School of Social Work and Community Welfare,  
James Cook University*

The issue of impact becomes particularly significant for those research disciplines which are by nature applied. Most broad HASS disciplines have specific branches with highly applied outcomes: the results of research in such areas are intended to apply directly within society, perhaps bypassing altogether the academic context in which they are produced and by which they are theoretically informed. For example, research in social work draws on theoretical and fundamental research in the social sciences but is often intended to produce outcomes in which the direct societal benefit is greater than the contribution to an academic debate. Much consultancy work in the areas of business and commerce is valuable precisely because it speaks directly to industry. In the arts, most creative practice is intended first and foremost to enrich culture directly, and not to be mediated through academic criticism or constrained by the traditional

<sup>25</sup> Allen Consulting Group, op. cit., pp. 42–43.

format of conventional academic research. This in no way diminishes its importance or standing as research; it does, however, necessitate a different approach to the measurement of value—one that takes ‘impact’ into account as a primary rather than a secondary measure.

This [the assessment of impact] hangs around notions of value. I believe it would be possible for staff to argue the significance of their work as part of the process. That was adopted in some parts of the RAE process in the UK ... In drama and music individuals could argue the case for their works’ significance. In that way purely professional practice can be reduced as a research outcome. In addition, there are known achievements that can simply be listed (such as Australia Council grants or awards for example).

*Submission, Professor Andrew Schultz, Dean, Faculty of Creative Arts,  
University of Wollongong*

Impact is best measured and judged on the basis of highly discipline-specific evidence. No set of metrics can be common to all disciplines or research units. For example, when measuring the impact of innovative, practice-based research in the creative arts—for instance, a musical composition—metrics such as the number of performances and broadcasts are both objective and valuable. Monash University has devised an internal university points allocation system for various musical creative outputs, including published musical scores, commercially published or broadcast original compositions, recorded solo performances at commercial concerts, and national television broadcasts (ABC or SBS) of ethnographic films about music. More details are in Appendix A.

While there might be very broad conceptual categories within which, say, impact on public policy and the impact of a work of art can be compared, such categories are not practically useful in making a judgment based on concrete and specific evidence. Such a judgment also needs to be informed by the contextual knowledge of an expert in the field, who has knowledge of the broader operation of disciplinary practice in society.

This requirement for a variety of measures for assessing the impact of research contrasts with a near consensus on the most appropriate ways to assess research quality, regardless of the discipline involved.

Within disciplines, however, there are reliable, discipline-specific sets of evidence that can be used to support claims for research impact. For example, there are quantifiable mechanisms for determining the extent to which research in law has influenced judgments in Australian courts; there are bibliometric indicators which support claims that research in public policy has shaped policy formation; there are reliable measures of the dissemination of musical performance and composition, such as statistics on broadcasts and performance venues.

For this reason, this report recommends that the assessment of impact be construed differently from the assessment of quality—specifically, that assessment of impact should be conducted under an ‘audit’ model. Institutions, research groups or departments can compile evidence-based cases demonstrating the impact of their research. Appropriately qualified experts—both academic and, crucially, experts from industry, public institutions and other research end-users—should then audit some of these cases to determine whether the evidence presented supports the case made.

The use of experts from outside the academy as part of the assessment process is a significant departure from current processes. It would introduce a new element to the evaluation process: that of the consumer. This project’s consultations with the users of research suggest that many people are expertly qualified to assess the impact of research and its benefit to society, from perspectives beyond those of university academics and PFRA researchers. In the HASS disciplines, in particular, ‘impact’ is perhaps best thought of as a dialogue between researchers and the community—the many government, community, charitable, educational, media and cultural organisations and groups that are daily and vitally informed by HASS research. Both sides of this dialogue need to be represented in judgments of research impact.

The approach taken to the evaluation of impact in the case-study trial had, at its starting point, work conducted in the United Kingdom for measuring the impact of arthritis research. This recent approach to the multivalent notion of research impact was developed by the Higher Education Research Group (HERG)

at Brunel University and RAND Europe for the biomedical disciplines, and outlined by Dr Jonathan Grant of RAND Europe when he visited Australia for this project. Because the HERG–RAND approach allows for a broader conception of impact than the influence of research on purely academic work, and for a flexible weighting of the various identifiable types of impact, it seems an appropriate model for the HASS disciplines.<sup>26</sup> The categories of impact defined for the biomedical disciplines (referred to as the ‘payback’ from the research) include:

- A** Knowledge
- B** Benefits to future research and research users, including
  - i. Better targeting of future research
  - ii. Development of research skills, personnel and overall research capacity
  - iii. Critical capability to utilise appropriately existing research, including that from overseas
  - iv. Staff development and educational benefits<sup>27</sup>
- C** Political and administrative benefits, including
  - i. Improved information bases on which to take political and administrative decisions
  - ii. Other political benefits from undertaking research
- D** Health sector benefits, including
  - i. Cost-reduction in the delivery of existing services
  - ii. Qualitative improvements in the process of service delivery
  - iii. Increased effectiveness of services, e.g. increased health
  - iv. Equity, e.g. improved allocation of resources at area level, better targeting and accessibility
  - v. Revenues gained from intellectual property rights
- E** Broader economic benefits, including
  - i. Wider economic benefits from commercial exploitation of innovations arising from R&D
  - ii. Economic benefits from a healthy workforce and reduction in working days lost.

In the HERG–RAND model, these categories were used as the basis for qualitative assessment by experts. This approach was adapted for the case-study trial in a process described below. Each participating research group was given free rein to provide suitable, discipline-specific evidence for the impact of its research. The following categories—broadly, HASS sector analogies to the biomedical categories used in the HERG–RAND report—were given as guidance:

- the take-up, recognition or influence of research or creative work by users in industry, government, community and cultural organisations
- indicators of social, cultural, artistic and economic benefits to the community that flow directly or indirectly from the research or creative work, where such a link can be demonstrated conclusively
- media profile or other public exposure (e.g. online accession) of the research or creative work
- successfully acquitted collaborative projects with partners outside the academic sector
- membership of professional, industry or cultural organisations or advisory groups, where relevant to the discipline
- commercial returns on research outcomes and creative products.

<sup>26</sup> See Appendix A. The original source is Wooding S, Hanney S, Buxton M and Grant J (2004). *The returns from arthritis research*, Volume 1: Approach, analysis and recommendations, RAND Europe.  
<[http://www.rand.org/pubs/monographs/2004/RAND\\_MG251.pdf](http://www.rand.org/pubs/monographs/2004/RAND_MG251.pdf)>

<sup>27</sup> This category of ‘payback’ and its associated subcategories are defined as ‘capability’ in this report.

I believe we have to appreciate the wide range of values that are and can be attached to any project outcomes in these fields. Social, historical, aesthetic, emotional, intellectual, educational and other values may all be involved and can (and probably should) be written into the project aims. Each of these values can be measured to some degree, although the measurement tools available might be more subjective for some than for others. There is still a lot of work to be done in terms of finding effective evaluation and assessment tools, and large-scale comparative projects still seem rare ... I suspect that genuine assessment has to look at the full range of these values and how, why and whether they're being delivered (and whether they can be effectively measured yet). We also need to recognise that the research process itself can differ vastly from project to project and from that used in other fields, so a 'one size fits all' approach is probably not appropriate. Also, just as for the sciences there may be unforeseen consequences, impacts or findings as a result of the research—assessment criteria need to take this into account.

*Submission, Bobby Cerini, Science Communications Advisor, Office of the Director, Questacon—The National Science and Technology Centre*

Each participating group took a very different approach to the evidence it presented. While this may well represent disciplinary differences, assessors did see the need for some degree of standardisation in how evidence of impact should best be represented. Notwithstanding this perceived need for standardisation, the assessors—both academic peers and experts from outside the academy—found during the trial that the assessment of impact (as distinct from quality) was both achievable and significant.

### 5.3 Capability

Capability is the third strand of the research assessment process, alongside quality and impact.

In contrast to quality and impact, which are explicitly attributes of research outcomes, capability revolves around people: their research skills, their deployment, the way they are resourced and supported, and the way they collaborate. In a healthy and dynamic national research environment, people move around. Investment in the development of research capabilities pays dividends not just to the unit making the investment, but to the research capacity of the nation as a whole. A unit with a flourishing research training program that delivers high-quality researchers and research-capable professionals into the academic and non-academic workforce has made a valuable contribution to the national research capability, whether or not those researchers ultimately contribute research outcomes to the unit.

In its response to a prompt in the RQF Issues Paper, CHASS recommended strongly that any RQF should definitely include a 'forward-looking strategic element'. This element is captured within the definition of 'capability' presented here. Research training is a central component of capability. Equally important is the development of early-career researchers (and existing research assessment processes, such as the United Kingdom RAE, singularly fail to deal adequately with issues involving early-career researchers, according to most respondents to the CHASS consultative process).

Determining the impact of research is difficult because there is often a time factor involved; something might have a strong impact over a long period. There are the obvious factors in the visual arts such as art reviews, commercial success of work, articles written about artists, receiving awards, but these should be balanced against the broader impact [of individuals, not just works], which might include invitations to exhibit internationally, or to represent the nation; invitations to participate in prestigious art events, especially internationally; artists recognizing the achievements of research in their disciplines by electing artists to professional roles or awarding honours.

*Submission, Dr Bernard Hoffert, Associate Dean (External Affairs), Faculty of Art and Design, Monash University*

One problem with the allocation of research funding purely on the basis of the quality and impact of prior research is the lag before such evaluations can be made. This issue is particularly acute in the HASS disciplines. Many respondents to the consultative process stressed the length of time it can take for HASS research to have an impact on society, but the fact that some is still having an impact after a considerable time—in some cases, decades—is evidence of its quality and value. By definition, assessing ‘capability’—a unit’s likelihood of future research success—overcomes the lag in indicators of research quality and impact.

Our greatest impact is in fact in the area of capability, across various domains (most obviously and pertinently, education, professional practice, community practice). Capability incorporates relevance, and relates to the work of creating capability within the academy, and in professional and community practice.

People trained in universities do not necessarily have research skills, but do have the capacity to think, assess, evaluate and critique. How can these skills (capabilities) be measured? One approach would be to measure the numbers of PhD and MA graduates being produced by a department, their employability, and the destination of graduates. An important measure of capability is the extent to which graduates are equipped to engage in contemporary professional practice.

Appropriate assessors would vary, depending on what capabilities were being assessed and the discipline under question. If measuring the benefits accrued socially, then assessors who can see benefit being put to work—i.e., external assessors—should be used. When the benefit/capability is designed to feed back into the academic sector, then internal assessors are more appropriate. Hence we support the use of peer review, whether these are academic, artistic or professional peers. For the social sciences, the weighting would necessarily be towards external assessment because of the need to evaluate effect of research product; for the humanities it would be mainly internal because much of the work carried out here is about intellectual formation; and for the arts, a mix of internal and external assessment.

We note the demarcation between assessment of quality of teaching and assessment of quality of research, but wish to point out the nexus between effective teaching and research in any discussion of impact and capability.

*Focus group feedback, consultation with heads of research centres, 18 March 2005*

Capability can be measured by a number of pieces of data and evidence that will give a reliable indication of a unit’s likelihood of producing high-quality, high-impact research in the future, or its ability to send research-capable individuals into the academic or non-academic workforce. These cover, broadly, the strategic management and planning of the research unit; its success in finding appropriate collaborative partners inside and outside the publicly funded research sector; its management of finances and infrastructure that support the research effort; its successful recruitment and development of early-career researchers; and its success in research training and in providing education that is led by and focused on research in a way that demonstrably increases the research capability of the discipline.

For the purposes of the case-study trial, participating research groups were given the following guidelines to the types of evidence for capability that might be useful to assessors:

*The unit should provide a narrative account of its strategic research planning and management processes, drawing particular attention to how these will enhance the capabilities of the unit, its partners, and the nation. This should be supported by evidence and indicators from the period January 2000 to the present that include:*

- *Budgetary data*
- *Number of staff, and their academic position*
- *Research students’ enrolments, completions, and graduate destination*
- *Evidence of the recruitment and development of early-career researchers*
- *Details of research infrastructure and plans for its development and management*
- *Statistics and accounts of ongoing externally funded research projects*

- *Details of ongoing collaborations both with other academic units and with partners in industry, government, the community and cultural sector*
- *Alignment of research plans with National Research Priorities.*

These are among the most standard data available to institutions (for example, in annual reports) but, woven into the tapestry of an overall statement about research capability, they provide an invaluable context for assessors seeking to ground their judgments of research quality and impact. A sample of comments from assessors shows how the overview and context provided through an account of research capability was seen as an integral part of the trial process:

- *I found the prizes and industry projects sections from [the department] made a useful contribution towards my understanding of the high impact of their work. The prizes, ... perhaps because they are quantitative, were easy to grasp as evidence of impact. The industry case studies (the [deleted] project and the Linkage grant) were more difficult to quantify and I suspect that most schools of [this discipline] could come up with examples that, superficially at least, looked similar in quality and impact ... from the documents provided, I felt better able to rank capability than quality or impact.*
- *[In this case] there is the challenge of assessing a group that has only so recently come together. Understandably, we must draw heavily on the track record of the key players, particularly in assessing quality and impact. But in this situation I think we need much more on their strategic planning, their governance model, their intentions about partnerships and collaborations etc., and how the new organisation will take advantage of the elements brought together. Some sense of this is in the material, but not enough to give me confidence in rating their capability into the future.*
- *The unit's own overview, including its caveats about assessment evidence, was critical. Those without such an overview, or with a limited one, were much harder to assess, particularly by a non-expert. This also suggests you should be pretty flexible about how the detailed evidence is provided. But such evidence is important ... I am reminded of some work a few years ago on evaluating policy advice. That cautioned against emphasising too much the extent to which advice is accepted and, when it is, what happened. These things are relevant (particularly the latter obviously), but it is also important to assess the inputs (e.g. the information and the expertise) and the process (e.g. consultation, timeliness, accessibility, assessment of risks and implementation issues). In a similar way, I think impact and capability in particular need to include evidence of inputs and processes to complement the (probably limited) evidence of final societal impact or effectiveness.*

*Assessors' comments, case-study trial*

These views are further discussed in Section 6.2.

Section 6 also presents a conceptual structure for grouping types of capability, from direct capability-building through the development of a research group's skills and infrastructure, to broader and more diffuse enhancement of society's general research capability.

The only useful 'set' of measures is that which has the flexibility to accommodate such diversity. It must be broadly conceived; it must allow for emerging fields, evolving questions and paradigm upheavals and it must recognise the spectrum of knowledge types.

In addition, the view that there must be a clear link between assessment outcomes and funding needs to be debated. We need to challenge the argument that high quality research should therefore obtain relatively high funding levels, compared to emerging research developments. Rather, any research quality framework should incorporate the principle that research quality needs to be nurtured and enhanced. Appropriate mechanisms can be developed to ensure that funding is not wasted.

In the same vein, given national demographics, and the age profile of Australian researchers, succession planning to replace individuals, and to restructure research units and concentrations needs attention as part of the foundation for the continuing growth of research excellence.

*Submission, Associate Professor Suzanne Franzway, Dean Research,  
Division of Education, Arts and Social Sciences, University of South Australia*



## 5.4 Results from the case-study trial

In the course of the trial assessment process conducted for this research project, seven research groups from a variety of HASS disciplines each presented a ‘case for excellence’ for its research quality, research impact and research capability. Table 3 summarises the evidence presented by each unit and provides a convenient overview of the range of indicators that might be used in any RAE—perhaps as a ‘menu’ of possible indicators from which research groups can select the most relevant.

One of the issues that emerged from the trial—and one of the lessons learned—was that it was not always possible or desirable to distinguish or prescribe which indicators applied primarily to ‘quality’, which to ‘impact’ and, to a lesser extent, which to ‘capability’. This was often context-dependent. In general, assessors found it relatively easy to discriminate and to synthesise their judgments in each of the three categories on the basis of the whole body of evidence presented. Table 3 gives the categories ‘Q’, ‘I’ and ‘C’ in order of priority; for example, ‘IQ’ suggests that a particular indicator might be useful mainly in judging impact, but also in assessing quality. The entries are suggestions and should not be taken as prescriptions.

**Table 3:** Summary of evidence of research quality, impact and capability presented during the case-study trial

	Architecture and design	Multidisciplinary cultural studies	History	Politics and public policy	Marketing	Law	Music
<b>Overview data</b>							
Narrative describing the research and research-training activities of the group, research management, plans and organisation	QIC	QIC	QIC	QIC	QIC	QIC	QIC
Quantitative data: numbers of staff, students, research projects, and details of research infrastructure and equipment	QIC	QIC		QIC		QIC	
Individual researchers’ statement of research impact				I		I	I
Financial management data		C					
<b>Research production and dissemination</b>							
Traditional (e.g. current DEST-eligible) academic publications, and academic conference papers	QI	QI	QI	QI	QI	QI	QI
Professional publications (usually not currently DEST-eligible), and professional conference papers	IQ					QI	QI
Non-written research outcomes (e.g. creative works, designs, documentaries, multimedia websites)	QI	QI					QI
Policy and other reports				IQ		IQ	
Self-assessment of publications and scholarly/artistic outputs (e.g. ‘best five’ publications by researcher, journal rank, etc.)				QI	Q	Q	
Media appearances, articles and reviews	I	I	I			I	QI
Public awareness programs		IQ					
<b>Bibliometric data</b>							
Citation indices				Q		I	
<b>Research funding</b>							
ARC grants awarded (Discovery grants, fellowships)	QI	QI	Q	Q	Q	Q	
ARC Linkage grants awarded	IQ	IQ		QI		QI	
Other government (local, state, federal non-ARC) funding for research	IQ	IQ	QI	QI		QI	
Industry research funding	I			QI		IQ	I
Competitive research funding internal to the institution	QC	QC	Q		Q		QI
Industry funding for education-linked research	CI	CI					
Competitive funding internal to the institution for education-linked research	C						
Data on successful <i>and</i> unsuccessful grant applications		QC					
Assessors’ reports on grant applications				Q			Q

	Architecture and design	Multidisciplinary cultural studies	History	Politics and public policy	Marketing	Law	Music
<b>Awards, prizes, fellowships and markers of esteem</b>							
Fellowships of academies and learned societies			Q	Q	Q	Q	Q
International lectures, keynote addresses, invitations to address government and non-government organisations	QI	QI	Q	QI		QI	QI
Academic prizes and awards		QI	Q	Q	Q	Q	
Industry/professional prizes and awards	IQ			Q	QI	QI	
Editorships and memberships of editorial boards	QI	QI	QI	Q		Q	Q
Convenorships of academic conferences	QI	QI	QI		QI		QI
Convenorships of professional conferences, exhibitions and/or artistic festivals.	IQ						
Offices held in academic societies and organisations			Q	Q		Q	
Offices held in professional or industry organisations or government consultative bodies	IQ	IQ	IQ	Q		IQ	
Academic standing of scholars-in-residence and adjunct researchers	QIC	QIC	QIC	QI		Q	
Professional standing of advanced students and scholars-in-residence from industry	CIQ						
<b>Partnerships</b>							
Collaborations with other academic institutions and networks (national and international)	QI	QI		IC		QI	QI
Collaborations with government	IQ	IQ	IQ	I		IQ	
Collaborations with industry	I		IQ	I	I	IQ	QI
Collaborations with community organisations	I	I	I	I	I	IQ	
<b>Research training, education, and early career researcher development</b>							
New researcher recruitment and advancement strategies		CQ	CQ				
Research degree enrolments and completions	C	C		C	C	C	C
Destinations of research students and postdoctoral fellows			CQ	CQ			
Research-based education initiatives	C			CQ			CQ
Research-student publications	CQ	CQ			CQ		CQ
Prizes and grants for teaching and learning in research-based fields of education	CQ						
<b>Equipment and infrastructure</b>							
Self-assessment of research equipment and infrastructure	C		CQ	C		C	CQ



## 6 The proposed model

*How is this diversity best captured in any set of measures of quality and impact?*

Forensically. If you can make a good case, then make it. The problem will be ‘how are these self-reports to be judged?’ And by whom? Here is where a small annual expert panel from both within and without (for benchmarking purposes) [individual universities] would be called in.

*Submission, Dr Ben Bradley, Research Management Committee,  
Charles Sturt University*

The assessment framework will focus on the three primary elements of research quality, research impact and research capability. The model has these features:

- 1** Research groups or departments are the most suitable unit of assessment for HASS sector research, enabling the best compromise between specificity, accountability and the accommodation of multidisciplinary research, which is increasingly a feature of research in HASS.
- 2** Assessment is by the presentation of an evidence-based case for excellence. This should use appropriate evidence and indicators (discussed in Section 5 of this report) to support the group’s claims for the excellence of its research in each of three categories: quality, impact and capability.
- 3** This case is assessed in a two-stage process, akin to the standard processes for the adjudication of applications for academic promotion, and similar to existing processes for the assessment of ARC grant applications:
  - i) In the first stage, academic peers with specific disciplinary expertise comment on the academic quality of the work, much like referees.
  - ii) Then, a conglomerate panel of experts, including academic peers, experts from outside the academy (such as from industry) and experienced research managers makes an assessment of capability and impact, including by selectively auditing a sample of impact claims for veracity and reliability. Taking into account all the evidence, including assessments from (i) above, they then make an overall assessment of research excellence.

### 6.1 Criteria

Broad criteria for the assessment of quality and capability, and in particular for impact, will have to be developed to take into account disciplinary differences. Ideally, it should be possible to develop a menu of indicators from which individual research groups would be able to select the most appropriate. The case-study trial conducted for this exercise provides a starting point—one that has been tested on seven academic disciplines in the HASS sector (see Table 3). We must be cautious before adopting existing international practice and criteria without further reflection. Even the limited trial conducted for this project has demonstrated that some criteria suitable for the United Kingdom RAE, for example, are not as appropriate in an Australian context.

For the purposes of the case-study trial, criteria for ‘quality’ were adapted from those used generically in early rounds of the United Kingdom RAE. These refer to research of ‘international significance’, ‘national significance’ and ‘lesser significance’. An adaptation of this wording was used for the criteria for ‘impact’. Criteria for ‘capability’ referred to a scale of national best practice. Full criteria used can be found in Appendix C.

Broadly, such terms are metaphors for standing within the research community and, as such, seem to have functioned adequately in practice in the United Kingdom. However, the particular needs of a smaller nation such as Australia point to a responsibility to conduct research that is relevant to the Australian environment, region, culture and society and to Australia's indigenous traditions and population. Indeed, this is reinforced by the nature and jurisdictional focus of Australia's official National Research Priorities. At times, the national importance of such research is at odds with its 'international significance'. As one assessor in the trial commented:

*I don't like the International/National/less-than-National wording at all, and the explanation was inconsistent. The first sentence is fine: 'the intrinsic academic or artistic excellence'. Why is there any need to have anything other than an International Best Practice standard? Why is there an assumption that 'the best work in the field in Australia' is necessarily lowering the bar vis-à-vis 'the best work in the field [internationally]?*

...

*It should be obvious that something unequivocally local—e.g. a splendid history of North Queensland—can be absolutely of the highest international standards within its genre of local/regional history, yet have pretty minimal 'international significance'. It is the intrinsic quality that ought to matter under this criterion. So why not just use one measure—something like 'International Best Practice'?*

Some framework is required in order to group the types of evidence presented for quality, impact and capability coherently, so that appropriate criteria can be applied. Such a framework is articulated below, drawing on the model for grouping types of impact posited initially in the 2005 Howard Partners report, *The emerging business of knowledge transfer*.<sup>28</sup> The categories proposed are simply convenient ways of conceptually grouping the different domains within which research quality, impact and capability can be identified. As such, they may be helpful markers for institutions and groups preparing cases for excellence, and convenient markers for assessors to use when discussing and evaluating these cases. There is, of course, no rigid methodological reason to use the taxonomy proposed below over any other (compare, for example, the similar taxonomy of impact used by RAND Europe in assessing arthritis research, discussed in Section 5.2). However, this set of conceptual categories has the advantage of being simple but broad enough to encapsulate all the types and measures of quality, impact and capability discussed in this report.

Within each of the categories, specific criteria can be developed for making comparative judgments between research groups. While the twelve categories are robust enough to be common among all disciplines, the criteria for making judgments on a group's success within each category will, to a greater or lesser extent, be subject-specific.

The four fundamental categories for assessing **research quality** should be:

- **Knowledge Generation:** new discoveries, inventions, insights and so on
- **Knowledge Manifestation:** research-based publications, performances, exhibitions, and so on
- **Knowledge Acceptance:** academic peer acceptance of the worth or quality of a researcher's research and scholarship, as demonstrated by academic citation of it by other academics, prescription of academic texts by other academics in their courses, academic peer and user views, and so on
- **Knowledge Expertise:** a researcher's or research group's reputation, as demonstrated through their track record, including successful research grants, consultancies, sponsorships, internal university funding support for research projects, and so on.

The four fundamental criteria for assessing **research impact** should be the four categories suggested in the Howard Partners report<sup>29</sup>:

<sup>28</sup> The taxonomy developed extends that introduced in Howard Partners (2005), *The emerging business of knowledge transfer*, DEST publications. <<http://www.howardpartners.com.au/publications/knowledge-transfer.pdf>>

<sup>29</sup> *ibid.*, p. x.

- **Knowledge Diffusion:** the broad, industry-wide adoption of research findings through education, capacity building and communication
- **Knowledge Production:** the explicit commercialisation and commodification of research
- **Knowledge Relationships:** research collaboration and partnerships to exploit intrinsic or ‘tacit’ knowledge
- **Knowledge Engagement:** the impact of research that stems from shared interests and concerns that transcend the boundaries of universities *per se*.

These four terms provide a conceptual framework that, interpreted in ways specific to individual disciplines, defines ways of demonstrating research impact in the wider community.

The four fundamental categories for assessing **research capability** should be:

- **Knowledge Infrastructure:** the research infrastructure and support provided internally and externally for researchers, whether they are organised into teams, groups, centres, laboratories, schools, departments, faculties, divisions or some other chosen assessment unit
- **Knowledge Culture:** the human and sometimes less tangible elements surrounding Knowledge Infrastructure that support an active research environment, including a critical mass of researchers and research degree students, high-quality visiting academics, active grant-based and non-grant-based research projects, and so on
- **Knowledge Training:** training and support for research degree students and their supervisors
- **Knowledge Strategy:** the forward-looking aspects of developing research capability now and for the future, specifically addressing and covering the need, highlighted in the RQF Issues Paper, to assess both research track records and future research strategy and operational plans.<sup>30</sup>

Having twelve fundamental criteria—four criteria each for the three basic elements of the assessment framework—should be easily understood, managed and accepted by various stakeholders. The criteria encapsulate many (if not all) of the features that academics and their employing institutions recognise and use in recruitments, appointments and promotions.

What is needed is a range of measures of quality and impact. Academic quality and academic impact are best assessed by international peer review (perhaps in concert with quantitative information about grants, HDR completions, and the like). But impact outside the academy requires quite different kinds of measures. Perhaps we should be thinking about some kind of national or community peer review, where the reviewers are not themselves members of the academy.

*Submission, Prof. Graham Oppy, Associate Dean Research, Faculty of Arts,  
Monash University*

There should be an illustrative set menu of particular indicators or metrics relating to each of the three basic elements of the framework (quality, impact and capability) and the four fundamental criteria that relate to each of the three elements. Many, if not all, of those indicators are outlined in this report, as well as in the RQF Issues Paper, the Allen Consulting Group report, and the Howard Partners report. Accordingly, this report builds upon the expertise and suggestions for the RQF and RAE in those earlier documents.

A basic menu of indicators would promote certainty, consistency and coherence. Research units could also be allowed some flexibility in choosing and justifying additional metrics or indicators that they claim are better suited to their particular circumstances. Units could be asked to use a limited or otherwise capped number of ‘best’ indicators of their fulfilment of each criterion.

<sup>30</sup> Research Quality Framework Issues Paper, March 2005, DEST.

## 6.2 Assessment panels

In the feedback received, most project respondents thought it important that the broader value of research in the HASS sector be identified, and that using appropriately qualified experts from the professions, industry, government, and cultural and artistic organisations would be a valuable way to do this. However, there was some contention about the composition of assessment panels, with some academics expressing a degree of anxiety at the prospect of non-academics assessing the impact of their work.

We do a great deal of research in PNG for bodies such as AusAID and the World Bank on things like rural development, food production, agricultural systems, poverty assessment, roads and HIV/AIDS. We can write reports that are critical for policy and aid giving, but we receive almost no recognition from academia ...

However, as soon as we publish a paper in a refereed academic journal we get instant academic recognition and sometimes very high praise. We have decided we will just have to go the extra mile and publish as often as we can. It makes for more work, but there seems to be no way around it.

I don't know what this means for assessing the value of research. Over the last 10 years or so belated recognition by the universities and the ARC and DEST of the amount of money brought into the university from outside has changed the position we used to find ourselves in ... We are now very highly valued because of our DEST earnings. However I suppose we still yearn for recognition within the academy for the value of our intellectual achievements.

Peer review is the way to go, but who are the peers? They will have to include both academic peers and non-academic peers. They should also include, in the case of our sort of work, senior researchers and public servants from the countries in which we work and because our work takes place within a country or a region, academics from outside our discipline. We do work that falls across geography, environmental studies, development studies, economics, anthropology, agricultural science and so on.

*Submission, Dr Bryant Allen, Land Management Group, Department of Human Geography, Research School of Pacific and Asian Studies, Australian National University*

The national–international balance of assessment panels requires careful attention on a discipline-by-discipline basis. Clearly, the international standing of research is an essential measure of research quality, but much research, particularly in the HASS sector, is directly engaged with issues specifically relevant to Australia. The international impact of research in, for example, Asia–Pacific development economics, Australian history or Indigenous Australian art is perhaps less important than its vital significance in building a prosperous and self-aware Australia.

This issue was discussed in Section 6.1, with reference to the formulation of appropriate criteria. It seems logical that the composition of assessment panels should be determined by those best qualified to assess against those particular criteria. If the criteria refer to the international standing of research, representatives from the international academic community are necessary; if the criteria refer to the impact of research outside the academic community, experts from outside that community need to be part of the process.

The two-stage assessment process outlined at the beginning of this section and in Figure 1 seems an appropriately articulated model that would enable various types of assessor to make contributions at the points at which their expertise is most relevant. Given the highly discipline-specific nature of research impact, it seems vital that both academic peers and qualified experts from non-academic institutions be involved in the process of 'auditing' claims for impact.

From the consultation and trialling CHASS has conducted, the best way to capture both quality and impact in an environment of disciplinary and institutional diversity, and increasingly multidisciplinary and collaborative research, is to allow a heterogeneous body of assessors to make inputs at various stages.

## 6.3 Flexibility of weighting

Different disciplines and types of research group vary widely in the goals of their research and the ways in which it meets the needs of society. This report recommends that a nuanced research assessment process should accommodate those differences. Following the example of the proposed modifications to the United Kingdom RAE for 2008, different disciplines and, arguably, different research groups or departments should be able to stipulate different weightings for quality, impact and capability to reflect their differing research foci.

A weighting of 40–70% for quality, 20–50% for impact and 10–30% for capability was suggested to several of the consultative and reference forums for this project and received cautious endorsement. Allowing ranges, rather than specifying precise proportions, would enable the group being assessed to ‘weight’ the assessment process according to the group’s strengths, once it had met the minimum requirements of 40% for quality, 20% for impact and 10% for capability.

## 6.4 Unit of assessment and reporting

What is the appropriate unit or level at which the assessment of research should be conducted: by individual researcher, by research department or group, or by discipline within a faculty or a university as a whole?

This apparently simple question is in fact extremely vexed. In many ways, all these levels of aggregation are useful and informative, and it is possible to come to conclusions about research quality and impact even for individual items of research.

There is perhaps no simple solution. Different stakeholders have differing perspectives on Australia’s overall research performance. For example, those responsible for managing the research of entire institutions may prefer assessments to be conducted, reported and funded at the institutional level. This would enable them to evaluate their own performance in developing institutional research strength, and give them most discretion in the internal allocation of research funding and in aligning research in the institution with institutional priorities.

However, a perspective more concerned with the overall shape of Australia’s research effort would focus more on research assessment conducted and reported by discipline or by focused research group or department. This would give greater transparency to the research assessment process, and would enable a greater alignment between research planning and performance and national priorities.

First, it is important that all academics participate in the assessment exercise. This is because extremely large departments will benefit from a measure that samples only 80% of academics. The assessment should be evaluating the standing of the University and departments within it on the basis of level of academic engagement, that is, how much on average each academic is contributing. The assessment should also be weighted for academic appointment and length of time as an academic so that Level As 2 years out of their PhD are not being viewed as needing to contribute at the same level as a Level E ... It is also important to evaluate Universities at the Faculty level rather than at the University level to allow for the specialisation some universities have in some areas but not others.

Second, individuals should put together a portfolio much like they do for promotion/study leave. It contains copies of work completed and scholarly activities during the assessment period, proof such work was disseminated as claimed. This would be assessed by peer reviewers of international and national standing in the area. While time consuming it would also be helpful in managing and lifting the overall quality of academic performance in this country.

*Submission, Professor Charmine Hartel,  
Director of the Centre for Business Research and Associate Dean (Research),  
Faculty of Business and Law, Deakin University*

The strong message received in the course of this project is that the HASS sector as a whole would prefer the second scenario. The context within which research success should be judged is national or international, rather than institutional. Research is increasingly collaborative, but this collaboration takes place between cognate research groups in different institutions, and between such groups and industry, rather than between different units of the same university or organisation. The Cooperative Research Centre initiative has been widely cited as the best example of this successful shift in research organisation.

Any new system of assessment will cause behaviour change in researchers as they learn the new rules and adjust their behaviour to maximise the chance of success. The research group or department is close enough to individual researchers to allow them to appreciate the implications of a new system, whereas an assessment at the institutional level is more distant from the researcher's immediate concerns and will be seen as less likely to affect them.

In the model proposed, there are strong grounds for assessing quality at the discipline level, primarily because academic peers who are familiar with existing disciplinary criteria and processes are best placed to make such an assessment.

Impact, on the other hand, is best assessed at the level of the research project, which even in the HASS sector is nowadays, more often than not, conducted by a team. For example, assessing the impact of a piece of social policy research involving contributions from demographers, sociologists, economists and political scientists is best done for the project overall, rather than through some next-to-meaningless attempt to disaggregate the relative 'impact' of each discipline. Indeed, the CHASS consultation with research end-users suggests that focusing academic research to an intradisciplinary academic audience, rather than using a multidisciplinary approach to answer pressing questions for the wider society, is a significant reason why much academic research has less impact than it might.

It is important to distinguish between the level of aggregation at which judgments are made and the level of aggregation at which they are reported. For example, implicitly in the United Kingdom RAE, judgments are made about individual items of research; if enough are deemed to be of sufficient standing, a judgment is reported at an aggregated level.

One principal benefit of any research assessment process is the shaping of positive behaviours among researchers and research managers, who are skilled in shaping their own behaviour to optimise the results from any assessment system. This should be the primary consideration in determining the level of reporting.

Reporting at too low a level—for example, that of the individual researcher—will lead to divisive practices and cause morale problems within the sector. In submissions to the project, much reference was made to the 'transfer market' mentality that some see as pervading the United Kingdom RAE. In that system, great importance is placed on an individual's past research record, to the detriment of enhancing the unit's potential to succeed in the future and the individual's long-term development as a researcher.

Reporting at too high a level—for example, at the faculty level—will lead to a lack of incentive for projects and groups to demonstrate specific success and be rewarded for it.

Whichever level is adopted, some way of disaggregating findings about research quality by discipline will be necessary. A great deal of research is multidisciplinary or interdisciplinary, and dealing with this adequately is a major issue for the RQF. A multilayered assessment process will allow impact and capability to be assessed by very broad panels of conglomerate disciplines (containing experts both from inside and from outside the academy). These panels' judgments should be informed by specialist, discipline-based, peer advice as to research quality.

Some form of disciplinary hierarchy will be needed to organise this process; the RFCO ('research fields, courses and disciplines') classification used by the Australian Bureau of Statistics provides such a hierarchy, although it needs to be updated to reflect recent research trends, particularly in the organisation of the human society and culture areas of scholarship.



## 6.5 Broader applicability of the model

Both ‘quality’ and ‘impact’ are important in research and in some ways are inseparable. Any evaluation of the significance of a piece of research should take into account both quality and impact; however the measures to evaluate the significance of research should be flexible as ‘impact’ will have varying boundaries. More and more ‘impact’ appears to be more aligned with very broad and extensive boundaries. However, it is possible to have a high quality piece of research that has apparently less impact because it might be undertaken within a specific context (or limited) context. For example research undertaken within a specific Indigenous or other community may ultimately have what appears to be a small impact which is limited to the community that it was conducted for and within, however the actual impact on that specific community may be enormous. Any evaluation of research should have enough flexibility to take these situations into account. Not all research will be high quality, high impact (as in international impact) research. It is important that overall social/societal wellbeing and agendas are included and seen as valid in any evaluation of the significance of a piece of research.

*Submission, Ms Darlene Oxenham, Director, Curtin Indigenous Research Centre,  
Centre for Aboriginal Studies, Curtin University*

Using a narrow definition of ‘research’, most HASS research in Australia takes place in universities. This is not the case with STEM research, in which a large amount of research takes place in publicly funded research agencies (PFRAs).

If one considers only research outputs, there are many similarities between the research conducted by universities and that of the PFRAs. But the universities have an additional role—developing the research capability of the nation into the future. They fulfil this through training the next generation of researchers, nurturing researchers early in their careers, and providing to the workforce a steady stream of graduates whose capabilities—and therefore the nation’s—are enhanced by the universities’ research and research skills.

For the university sector, then, the notion of ‘capability’ as the third key element of an overall assessment of research is significant. It is less significant for PFRAs because they do not have a certified educational role, but is still an important consideration because of their engagement in building infrastructure, culture and strategies for the nation’s research capability.

## 7 Conclusion

After extensive consultation with the HASS sector, an examination of existing research evaluation literature and practice, and a trial assessment exercise, this project has developed a model for assessing the value of research in HASS. The process combines quantitative, evidence-based performance data with qualitative peer and expert assessment, so that expert judgments are informed by objective evidence and not simply by subjective opinion.

In this process, a high-level group, consisting of a broad mix of specialists from cognate disciplines (broadly analogous to an ARC panel or a university promotions committee) makes an overall assessment, taking into account all the available evidence. This includes both evidence of quality, impact and capability provided by the research group and very specific evaluations of the research made by academic peers expert in the discipline. The case for impact made by the research group is selectively audited for veracity and reliability by appropriately qualified experts.

It is worth highlighting the similarities between this process and the established method of assessing promotions cases. The familiarity of the process increases the level of acceptability and comprehension among the academic community. Indeed, ‘acceptability’ is one of the main criteria for any proposed RAE model for Australia.

A strong element of the model is the flexibility it provides. The research group being assessed is able to choose from a menu the most appropriate evidence and indicators to make its case. It is also able to select relative weightings (within a range) for the three primary elements of quality, impact and capability, according to a balance of the three that is appropriate to the field of research and to the nature and aims of the research group.

The conclusion of this report is that the methodology detailed above has sufficient flexibility to be applicable to the assessment of all research in Australia, not just to that conducted under the banner of the humanities, the arts and the social sciences. In researching and developing this model, it was extremely important to find a process that was flexible enough to accommodate all the diversity between disciplines, goals, orientations and audiences that characterises HASS. It seems likely that the accommodation of such diversity extends further, to the diversity of research in science, engineering, technology and medicine.

The consultative process revealed a general acknowledgment—perhaps clearer now than in the past—of the surprising degree of accord between the HASS and STEM sectors. As both HASS and STEM move beyond the cloistered thinking of ‘academic’ quality and purely ‘academic’ impact, it becomes clear that there is a need to move past the tendency to hide behind synthetic measures or indicators that merely demonstrate academics’ engagement with other academics. Such measures and indicators are only meaningful when contextualised and part of a coherent strategy to maximise research quality, impact and capability.

In short, Australia’s new method for assessing research must be based on a case, supported by evidence, about which judgments are made by experts.

By following the well-tested model suggested in this report, the Australian RQF and RAE can make a groundbreaking and innovative contribution to worldwide best practice in recognising and measuring all facets of research performance.



# Appendix A

---

## Setting the scene: a review of current Australian and international practice in measuring the quality and impact of publicly funded HASS research

---

### A1 Introduction

The purpose of this review is to describe current practice in measuring the quality and impact of publicly funded humanities, arts and social sciences (HASS) research in Australia and internationally, and then to consider what we may glean from Australian and international practice in assessing both HASS and science, technology, engineering and medicine (STEM) research, together with proposed innovations to help us push forward the policy agenda and develop new and alternative thinking on appropriately capturing and assessing the quality and impact of research. While the terms ‘quality’ and ‘impact’ are contested, this review follows policy language and defines ‘research quality’ as research outcomes which benefit and influence academia, and ‘research impact’ as research outcomes which benefit ‘users’ and the Australian public.

### A2 Current measures of the quality and impact of publicly funded HASS research in Australia

This section describes current Australian practice in measuring the quality and impact of HASS research. Section A2.1 reveals that there is no existing direct measure of research quality or research impact tied to government block funding of Australian universities. Section A2.3 examines measures of quality and impact used to assess STEM and HASS research supported by publicly funded research funding agencies. We find that, despite advances made in the STEM sector, the notion of research impact is not operationalised to take into account specific HASS benefits, and that publication-based and citation-based indicators are the dominant form of assessing HASS research quality. The following descriptions of current Australian practice are interwoven with cautionary analyses of the application of these STEM-derived publication- and citation-based quality measures to HASS streams.

#### A2.1 University block funding—the IGS and RTS formulas

The Institutional Grants Scheme (IGS) and Research Training Scheme (RTS) were introduced as ways to link government block funding of universities to weighted performance measures, while maintaining the principle that universities set their own research priorities and manage their own research activities. According to 2004–05 estimates, the IGS and RTS formulas which constitute ‘performance based block

funding' guide the distribution of the greatest single financial outlay within the Australian Government's science and innovation budget, accounting for \$1,203.5 million or 22.5% of spending.<sup>31</sup> However, these measures do not directly gauge research quality or research impact.

IGS formula weightings	RTS formula weightings
60% = Research income	50% = All higher research degree completions
30% = Commonwealth-funded research students	40% = Total research income
10% = Research publications	10% = Research publications

The formulas contain two input measures: student load, which has no bearing upon research outcomes; and research income, which may be indirectly interpreted as a quality measure where grants are won on the basis of competition and peer review, although winning a grant does not guarantee the quality or impact of research outcomes. There are two output measures: higher degree completions, which is related to research training; and research publications, which is the only measure linked to research outcomes and accounts for 10% of the weighting in each formula. It is therefore the use of this 'measure' or 'indicator' that we shall investigate.

### A2.1.1 The research publications component

Currently, a university's share of research publications determines the proportion of funding it receives under the IGS and RTS formulas. 'Publications' currently correspond to four Department of Education, Science Training (DEST) categories:

- A1 a scholarly book produced by a commercial publisher (worth five 'DEST points')<sup>32</sup>
- B1 a chapter in a scholarly book produced by an international publisher (one DEST point)
- C1 an article in a scholarly refereed journal (one DEST point)
- E1 a peer-reviewed paper presented at a conference of national or international significance and published in its proceedings (one DEST point).

However, this measure only detects the volume of work produced, which is not the same as assessing research quality or research impact.

In Australia, the effect of distributing funds on this basis has run contrary to government intentions because increased research production, rather than improvements in research quality, has been rewarded. For example, Butler (2003a; 2004b) found a relationship between the introduction of the research publications component in Australian performance-based block funding and a sharp rise in journal publications, but in lower impact journals, a trend that holds for STEM and HASS fields alike.<sup>33</sup>

The 2004 *Knowledge and Innovation Reforms Consultation Report* discussed the publications component of the IGS and RTS formulas, and identified three weaknesses (DEST 2004, 10). The first, as identified above, was that the publications count rewarded quantity rather than quality. The second was that the publications component was highly correlated with the other elements of the IGS and RTS formulas, particularly research income, and thereby provided limited extra value. The third weakness was that the count is open to criticism of the types of publications that are and are not included in the DEST categories, an issue that is of particular interest for the HASS sector and is now discussed further.

<sup>31</sup> <[http://www.dest.gov.au/science/analysis/pdf/Science%20and%20Innovation%20Budget%20Tables%201%20to%205\\_2004-05.pdf](http://www.dest.gov.au/science/analysis/pdf/Science%20and%20Innovation%20Budget%20Tables%201%20to%205_2004-05.pdf)>

<sup>32</sup> Funds tied to publication ÷ all publications (weighted by type) = value per 'publication unit' or 'DEST point'.

<sup>33</sup> Universities began supplying publication details in 1993, and the publications measure was first applied in 1995 as part of the Research Quantum, then superseded by the IGS and RTS schemes in 2001–02.

The key issue is that publication counts tied to the current DEST categories favour communication practices appropriate to the experimental sciences and the more quantitative social sciences (i.e. those disciplines in which journal papers and conference presentations are the main modes of disseminating research). While books and book chapters are included, and are important publication sources for many fields of social science and humanities research, the creative arts are not well served by these arrangements.<sup>34</sup> As the coverage of DEST returns changes, so do the potential rewards for various HASS sectors and fields. For example, in 1994 there were over 20 DEST categories, including audiovisual recordings, computer software, technical drawing, designs and models, and other creative works, including recorded works and individual exhibition or representation of original art. The four current DEST categories came into effect in 1995, although in 2002 a reduced list of creative arts and design categories was reintroduced based on 2001 returns, but was dropped by 2003. The inclusion or exclusion of these additional categories has clear implications for assessing and rewarding research production in the creative arts where standard academic publication is not the norm. This also sends out strong signals about what is or is not considered legitimate research.

These considerations raise the question of whether there are any past practices related to publication counts that benefit HASS and can be resurrected and modified? The answer is that any revamped publication measure would remain count-based and would not assess research quality or research impact.

### **A2.1.2 Conclusion**

In Australia, there has never been a direct measure of research quality or research impact linked to university block funding. The only research-related output measure has been the number of publications produced, which is not something to be retained for any field of STEM or HASS research if the aim of government policy is to measure research quality and research impact.

## **A2.2 Measuring research production: a closer look at publication indicators and HASS**

Section A2.1 explains that research productivity does not relate directly to research quality or research impact. The aim here is to draw upon the bibliometrics literature to examine this issue in more detail and to focus on specific issues involving publication-based measures and HASS research.<sup>35</sup>

‘Bibliometric indicators’ are quantitative measures of published academic literature (or research ‘output’) and the citations these works make or receive. The simplest form of bibliometric indicator is the publication count, and this indicator is the most frequently used in research evaluations (Martin 1996; Toutkoushain et al. 2003), due in part to the ease of basing assessments purely upon counts, the potential variety of publications that may be included (or excluded), and the fact that there is no necessary reliance on specialised databases for access to information, the starting point being academics’ publication lists.

But why is it that publication output does not constitute research quality? Surely, for example, the current DEST categories of journal articles, books, book chapters and conference proceedings point to research quality, as they all require peer review prior to publication? In the bibliometrics literature, the concept of research ‘quality’ is taken to be quantitatively inaccessible and is therefore eschewed, although the importance of research can be measured by calculating the relative number of citations a publication receives. In other words, it is not possible to divine the value of a publication until we can assess its influence on subsequent literature—the more it is cited, the more influence it has. In this light, publication-based indicators are redundant, and citation-based indicators become the focus of bibliometricians’ activities (see Section A2.4 for an overview of citation indicators and HASS).

<sup>34</sup> Beyond this there are further concerns with the current DEST categories when, taking social science as an example, papers presented at arguably the most internationally significant political science and sociology conferences (those of the American Political Science Association and the American Sociological Association) are ineligible for inclusion in DEST returns because they do not follow the required refereeing or publication protocols. Section 2.2 takes this discussion further and draws out particular considerations for publication measures applied to HASS generally, by stream and by field.

<sup>35</sup> For an extensive review of the literature on quantitative indicators, see *Quantitative indicators for research assessment—a literature review* (REPP 2005). <<http://repp.anu.edu.au/Literature%20Review3.pdf>>

### A2.2.1 The social sciences and humanities

When discussing publication-based measures, we must note that, while the bibliometrics literature is largely devoted to the natural and physical sciences, there is a smaller, discrete area of study dedicated to the social sciences and humanities. This separation is premised on the view that research methods and orientations in the social sciences and humanities are distinct from those of the experimental sciences, that their communication practices or literatures are consequently structured differently, and that this has bibliometric consequences (Nederhof et al. 1989, 427; Glänzel and Shoepflin 1999, 31; Luwel et al. 1999, 13; Hicks 2004, 473). Publication-based measures were originally developed for STEM, using journal articles and conference proceedings as the basis for scholarly communication. When applied to HASS, such a model excludes core categories of publication in humanities and social science fields, such as books and book chapters, and ignores non-scholarly ‘enlightenment’ literature aimed at practitioners and the public.<sup>36</sup> In fact, the idealised model of scientific communication does not equally serve all STEM fields: it has long been recognised that in engineering there is a greater focus on producing technical reports, which are excluded by journal- and proceedings-based counts (Bourke et al. 1996).

The current DEST publications count does serve many social science and humanities fields relatively well because it includes books and book chapters—hence the call in the 2004 evaluation of the knowledge and innovation reforms to retain this element of the RTS and IGS formulas (DEST 2004, 43).

### A2.2.2 Creative arts and design

While there is a separate focus in the bibliometrics literature on the social sciences and humanities, the creative arts are, in contrast, a very neglected area. There are, however, many possibilities for evaluating creative arts and design which radically push the boundaries of what we understand to be ‘publication’. Strand discussed at length how we may think about defining publication in the case of the block funding of Australian creative arts, and presented the case that ‘publication may take many forms, of which the written word is merely one’ (Strand 1998, 55). He argued that to publish something means to make it publicly known, which is to present ‘for public viewing, criticism, analysis and evaluation, an author’s ideas, knowledge, stories or information’ and in this way ‘[t]he author’s ideas become available for public comments and testing, made available for public use, building on to previous knowledge’ (Strand 1998, 53–4). This process can involve a variety of formats, such as ‘the production and distribution of books, reports, papers and such that consist of words, drawings and photographs. In the broader sense, however, publication of a work may entail hearing it, viewing it, reading it or experiencing it in other ways’, and it follows that a publication may ‘be presented in ephemeral and less tangible ways, such as through performance on a stage or in some other public forum’ (Strand 1998, 54). In this light, we could say that the creative performance is the publication.

Such a wide variety of potential publication formats can be illustrated in several ways. For example, in the case of music, a text is something that can be read, ‘whether it is in visual, audio or written forms’ (Strand 1998, 54), and an opera ‘could be performed on stage, recorded on compact disc and filmed all at the same time’ (Strand 1998, 55). Strand borrowed an example from Webber (1994) and asserted that in architecture the publication is the actual building and not photographs of it or written descriptions of it in journal papers. Strand emphasised that a narrow view of evaluating publication in the creative arts in written terms alone creates anomalies whereby a painter’s paper about their own exhibited painting counts as a publication but the painting does not, and a critical paper on a musical composition counts as publication while a performance of the composition, and even the composition itself, do not count as publication.

The view that only scholarly texts count as publications clearly excludes the central activities of creative arts and design. For example, a draft review of Australian Research Council (ARC) humanities and creative arts research draws attention to the importance of output, identified in final ARC reports, which falls outside of the DEST categories (ARC 2004, 11). Wissler (2004) noted that, while there were various operational difficulties when the DEST publications list included measures of creative activity (see Section A2.1 above),

<sup>36</sup> Luwel et al. (1999, 30–31, 156) use law and linguistics to represent the social sciences and humanities, respectively, and present an extensive range of scholarly and non-scholarly publications which academics routinely produce but which will be excluded from standard publication counts.

Australia's current counting system when applied to arts, media and design is inadequate and lags far behind international practices.

### **A2.2.3 Conclusion**

To sum up, if an indicator is designed to measure STEM publications, it will detect the most STEM-like publications when applied to HASS, and will privilege highly quantitative social science fields such as economics and psychology. Publication counts which include books and book chapters will benefit the social sciences and humanities more generally, although non-scholarly literature aimed at the public or practitioners will be overlooked. Creative arts and design are neglected by the bibliometrics literature, and are practically invisible when using conventional publication measures.

While this discussion has illuminated a variety of issues relating to publication indicators and HASS, and is important because publication counts are the only 'output' measure of research currently tied to Australian university block-funding, we must not be too distracted by a debate that is ultimately divorced from assessing research quality and research impact. Despite the fact that the consultation exercise for the knowledge and innovation reforms recommended from a social science and humanities perspective that research production should be retained as an indicator in the IGS and RTS formulas, it can be expected to disappear.

The core concern for the HASS disciplines is that citation indicators are likely to be strongly advocated as a replacement for the current publication count, and these measures will be based purely upon the STEM-friendly model of journal publishing. HASS research will not be favoured by such arrangements, a subject to which we shall turn in Section A2.4. For HASS, this would not be so much a case of throwing the baby out with the bathwater, but keeping the bathwater and throwing out the baby.

## **A2.3 Publicly funded research agencies and research funding agencies—current measures of quality and impact**

Public money is used in Australia to fund three STEM-dedicated research agencies: the Australian Institute of Marine Science (AIMS), the Australian Nuclear Science and Technology Organisation (ANSTO) and the Commonwealth Science and Industrial Research Organisation (CSIRO).<sup>37</sup> A brief overview of their current measures of research quality and research impact is included here because, as demonstrated above, it is the norm for STEM-based measures to be applied wholesale to HASS, and it is therefore important to consider the potential consequences of this. There follows an assessment of current measures of quality and impact employed by Australia's publicly supported research funding agencies: the ARC, which provides dedicated funding for the humanities and creative arts and the social, behavioural and economic sciences; and the National Health and Medical Research Council (NHMRC), which funds some HASS research within its medical and health remit. The collective practice in these agencies poses some concerns for HASS if research quality is measured using journal-based citation indicators, but there are promising STEM-led initiatives which could benefit HASS in measuring research quality using non-bibliometric indicators, and in adapting innovative models for assessing the benefits and outcomes of publicly funded STEM research to measuring research impact in economic, policy, social, intellectual, artistic, cultural, educational and community terms.

Existing measures and data that can aid this project's central aim of developing indicators of HASS research quality and research impact are identified.

### **A2.3.1 Publicly funded research agencies**

It may seem strange to discuss AIMS, ANSTO and CSIRO in a HASS-centred report, as these are STEM-focused institutions, but there is an important reason for briefly doing so: STEM practice in evaluating research has traditionally provided the template for national research policy. The Research Quality Framework (RQF) consultation process considers measures of quality and impact which are to

<sup>37</sup> The Defence Science and Technology Organisation is excluded here because it falls outside the remit of DEST, being under the leadership of the Chief Scientist as part of the Department of Defence.

be applied to university-based research and to the research produced by these block-funded scientific organisations. It is therefore vital to consider the strengths and weaknesses of the approaches these institutions collectively adopt to evaluate scientific research, in case these practices become influential models in the RQF deliberations, and to anticipate how these practices may impact on the HASS sector.

As part of their triennial funding agreements AIMS, ANSTO and CSIRO are undertaking internal research assessment specifically to feed into the RQF consultation exercise. It follows that, at the time of writing, these institutions are considering measures of quality and impact for STEM research, so it is prudent to ascertain whether they will advocate traditional STEM-friendly measures or alternative indicators which may be more complementary to HASS—developments which shall be discussed in Sections A4.1 and A4.5.<sup>38</sup>

### Measures of research quality

All these publicly funded research agencies use ‘number of publications’ as a performance indicator, although we have established that measuring research output does not unearth any truths about research quality. AIMS does not specify what kinds of publications it counts, while ANSTO and CSIRO both include journal articles, conference papers and abstracts, books, book chapters, monographs, and commercial and technical reports, and CSIRO also includes the number of citations to patents.<sup>39</sup> In addition, CSIRO adopts intellectual property as an output measure, in the form of numbers of inventions, patents and applications for patents; Australian and foreign trade marks; and Australian and foreign registered designs.

AIMS and CSIRO both employ citation analyses of publications as a quality measure, an approach which is based on the indexed journal literature alone. ANSTO alludes to assessing the quality of science and technical publications together with conference proceedings, but provides no details of how this quality is gauged (ANSTO 2004, 18).<sup>40</sup>

All three organisations list the awards and prizes received by their staff to denote research quality, and ANSTO presents the number of internationally recognised visitors it receives in a year as a quality measure. AIMS mentions external peer assessments and reviews of its research.

There is a great deal of common ground in the use of quality measures, and fairly standard indicators are adopted. There are no separate evaluation procedures for any HASS research funded by the agencies. Prizes, honours and peer assessment are measures of quality which can in principle be applied equitably to STEM and HASS alike. However, while citation indicators drawn from the indexed journal literature are well suited to evaluating the experimental sciences, citation-based measures are extremely problematic and disadvantageous for Australian HASS research, an issue elaborated in depth in Section A2.4.

### Measures of research impact

There is less cross-organisational consensus to be found in approaches to identifying research impact, although this is presented in the shared instrumental contexts of commercialisation, relationships with business, and advice given to government. AIMS and ANSTO measure research commercialisation and

<sup>38</sup> To further bolster the case for considering current (and potential future) measures of research quality and impact used by these STEM bodies, it is worth remembering that CSIRO received an estimated \$576.5 million, or 10.8% of the science and innovation budget for 2004–05, and ‘Other R&D agencies’ (which include AIMS and ANSTO) received \$398.6 million or 7.5%, compared with \$1203.5 million or 22.5% allocated for university block funding under the IGS and RTS formulas. STEM-dedicated cooperative research centres attract \$193 million (3.6% of Commonwealth funds), and in 2003–04 the CSIRO alone participated in 50 of the 71 centres. In other words, at 21.9%, funding for these three areas of dedicated STEM expenditure is almost equivalent to that for the whole university sector through the IGS and RTS. See <[http://www.dest.gov.au/science/analysis/pdf/Science%20and%20Innovation%20Budget%20Tables%20120to%205\\_2004-05.pdf](http://www.dest.gov.au/science/analysis/pdf/Science%20and%20Innovation%20Budget%20Tables%20120to%205_2004-05.pdf)> for Australian Government science and innovation expenditure figures.

<sup>39</sup> See AIMS (2003, 15), ANSTO (2004, 18–21) and CSIRO (2004, 23ff).

<sup>40</sup> AIMS and ANSTO commission ANU’s Research Evaluation Policy Project (REPP) to provide *ad hoc* citation analyses of their publications, and REPP supplies CSIRO with regular citation studies.



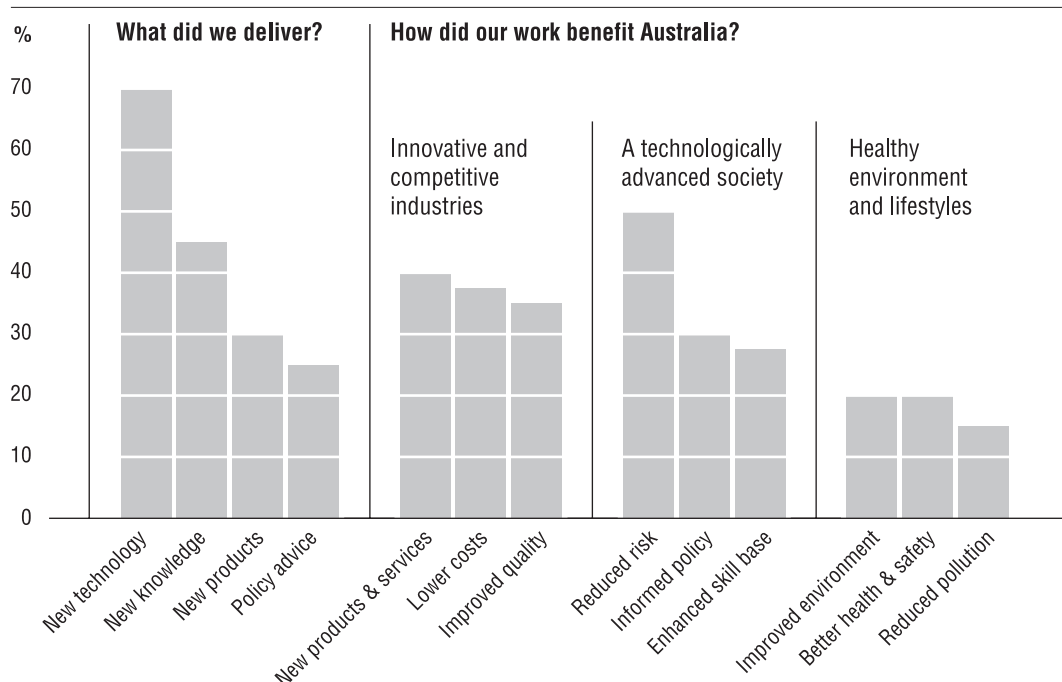
technology transfer variously in terms of the revenue generated, the number of commercial disclosures and arrangements, and the number of start-up companies created. CSIRO judges the impact of research on large and small business using ‘customer value’ ratings, while AIMS also calculates levels of an undefined ‘customer satisfaction’.

All three institutions gauge research impact through the advice they give to government: ANSTO in terms of its level of involvement in international policy developments by number of projects, person years and money expended; CSIRO through submissions, verbal reports and presentations, advice given for policy development, memberships of committees and invitations to briefings; and AIMS through a count of the number of advisory submissions.

ANSTO uniquely thinks about its impact in terms of the number of visitors who use its facilities, such as students, postdoctoral researchers and members of collaborative research projects.

Thus far, the thinking behind these impact indicators displays a shared concern with benefits to commerce and government. CSIRO, however, has broken new ground in ways of imagining and formulating the impact of its research. For example, in 2002 it created the Social and Economic Integration Emerging Science Initiative, which integrates social and economic considerations into the design, conduct and delivery of research with the aim of increasing the positive impact of R&D on people’s lives. This type of thinking has been operationalised in how CSIRO accounts for ‘delivering impact from our science’, where an ‘Outcome–Outputs Framework’ is presented and provides indicators for the economic, social and environmental benefits of CSIRO’s four main arms of research, in sum (presented in Figure A1) and individually (CSIRO 2004, 38–41).<sup>41</sup>

**Figure A1:** CSIRO—Overview of selected achievements, 2003–04



<sup>41</sup> CSIRO’s four research groups are Information Technology, Manufacturing and Services; Sustainable Minerals and Energy; Environment and Natural Resources; and Agribusiness and Health.

Although the precise methodology is not presented in Figure A1, we can see that impact ‘outputs’ cover new knowledge, technology and products, and policy advice, while benefits accrue to Australian business, the technological development of society, and environment and lifestyle. The four main strands of CSIRO research, taken individually, produce different outcomes.

To sum up, we again find that there are no specific arrangements for evaluating the impact of HASS research supported by the STEM agencies. Measuring HASS research impact is less-travelled ground in Australia, and internationally, than measuring research quality. However, it is apparent that some areas of HASS can follow STEM’s lead and demonstrate the impact of research on commerce and government policy, although it would be desirable to produce measures which also reflect the broader intellectual, social, community, artistic, educational and cultural impacts of HASS.

In this light, CSIRO’s outcome–output framework provides a potential starting point for HASS, and this type of approach has an inherent flexibility which may be adapted to aggregated HASS stream (or specific discipline or research group) impact ‘outputs’ and a wide variety of resulting benefits—a theme taken up in Sections A4.1 and A4.5.

### **A2.3.2 Publicly funded research funding agencies**

Turning to the research funding agencies, we encounter directly sponsored HASS research in the form of the ARC’s Humanities and Creative Arts Program and Social, Behavioural and Economic Sciences Program, and in the form of medical and health-related HASS research funded by the NHMRC. These research councils are considered below, in terms of current measures employed to ascertain the quality and impact of the general research they support and whether they use distinct arrangements for evaluating HASS research. We find that the ARC is quite naturally our most important consideration, and that its current evaluation practices of quality focus on publication or output (particularly on citations gained); as explained in Section A2.4, this will not favour HASS. Research impact is considered sporadically in the form of various commissioned reports, and only routinely evaluated in terms of ‘outcomes’ provided by end-of-grant reports, although a forthcoming ARC review of the humanities and creative arts reveals a likely combination of data sources and community benefits for future application.

While the NHMRC applies similar citation measures, it also employs and is currently developing a broad range of quality and impact measures which, while directed towards health and medicine research, hold promise for future models of HASS assessment.

#### **The Australian Research Council**

In the 2003–04 financial year, the ARC received \$482.4 million, or 9.0% of the Australian Government’s science and innovation budget. Total ARC research funding for 2002–06 gives over a fifth of funds to HASS research, representing the largest single Australian source of competitive HASS funding.<sup>42</sup>

#### *Measures of research quality*

The ARC adopts several approaches to aid the task of assessing the quality of its whole research remit which make no fundamental distinction between how STEM research and HASS research are evaluated. For example, information is collected from end-of-grant reports, including the number of project ‘outputs’ such as publications encompassing and extending beyond the core DEST categories (to include, for example, software and creative works) and measures of commercialisation activity (patents, licenses and start-up companies founded), although the latter is clearly more STEM-relevant. However, as end-of-grant reports may be submitted not later than six months after the research is completed, several outputs are undoubtedly missed. A further ARC measure of quality is provided by gauging the extent of international collaboration in Linkage grants, in the form of the number of projects, incidences of collaboration and number of countries involved.

<sup>42</sup> The HASS share of expenditure on Discovery Grants is 23.3%, Research Fellowships 24.2%, Federation Fellows 20% and Linkage Projects 22.2% (2002 grant statistics in ARC 2004, 153–157).



The ARC also commissions bibliometric studies of journal publications attributable to its funding (Butler 2004a)<sup>43</sup>; measures used cover the number of publications, citations per publication, actual and expected citation rates, and most highly cited publications. There is also analysis of international collaboration linked to level of research (e.g. fellowship schemes, various grant schemes, research institutes and other sectors). A distinction is made for HASS streams, in that they are evaluated separately using the relevant fields and subfields of the Australian Bureau of Statistics' 'research fields, courses and disciplines' (RFCD) classification scheme (economics; studies in human society; behavioural and cognitive sciences; social sciences; and arts and humanities). We are alerted to the fact that the size of publication sets determines the reliability of findings, with behavioural and cognitive sciences best served and social sciences, arts and humanities worst served, and that the coverage of the indexed database employed<sup>44</sup> clearly favours STEM over HASS, a point elaborated in Section A2.4.

### *Humanities and creative arts*

Quality aspects of humanities and creative arts (HCA) research are considered separately in a draft ARC review of this discipline grouping (ARC 2004).<sup>45</sup> While the review mainly discusses research 'impact', several areas fall into the remit of research 'quality' as defined in Section A1 of this paper. The ARC report does not formally present measures or indicators of HCA research; however, it examines major areas of concern for the HCA sector, identifies potential data sources (ARC applications, grants, end-of-grant reports) and, although it does not use these terms, reveals likely novel indicators which may be supported through mining ARC databases. While the focus is on HCA and its distinctive features, these indicators can readily be applied to the social sciences and to STEM.

A core concern is that many HCA researchers believe their research performance is currently assessed using inappropriate publication-based and citation-based indicators, which 'has the effect of hindering recognition of strength and opportunities in Australian HCA research,' and that productivity and performance measures must be developed 'in relation to what are actual and appropriate outputs in HCA fields of research' (ARC 2004, 50). For example, in terms of publication-based measures, we are reacquainted with the familiar and pressing point that HCA research includes significant creative works, designs, exhibitions and audiovisual recordings which are simply not counted because they fall outside the core DEST block-funding publication categories (ARC 2004, 11). As for citation-based indicators which depend on Institute for Scientific Information (ISI) journal coverage, we again meet the perennial concern that HCA research is more likely to appear in book or chapter form and will have less ISI coverage than the sciences or social sciences. The report draws upon the work of Butler (2004a) to demonstrate that a disproportionate level of high-impact international HCA journals are not indexed by ISI (ARC 2004, 21).

However, several potential indicators of HCA research quality, supported by existing ARC data, are revealed in a survey of 1996 Large Research Grants and Research Fellowships (ARC 2004, 64–66). Non-standard publication-related indicators include published reviews of research, republication (including translations and pirated web copies), invitations to contribute to other publications and to write prefaces, invitations to review or referee research, and invitations to join editorial boards or edit journals. Other non-bibliometric indicators include invitations to present conference papers and lectures (especially keynote ones); visiting professorships, fellowships or other appointments overseas; conferences dedicated to specific research; special journal editions; and invitations to launch books. Contributions to teaching which develop from research are identified in the form of invitations to write textbooks, to supervise or examine postgraduate research theses, and to teach in postgraduate programs. National and international prizes, honours and awards also recognise the quality of published writing or an individual's research standing.

Measures of research quality in the social, behavioural and economic sciences are not similarly reviewed by the ARC.

<sup>43</sup> Again, commissioned from REPP at ANU.

<sup>44</sup> Comprising the ISI's Science Citation Index, Social Science Citation Index and Arts and Humanities Citation Index.

<sup>45</sup> A draft version of the ARC's *Review of the Humanities and Creative Arts Discipline Grouping* is cited in this report with the kind permission of Mandy Thomas.

### *Measures of research impact*

Indicators of impact have been largely untapped in the ARC's assessment of the research it funds. This is quantified when figures are presented for the number of partnerships and linkages with industry by sector in terms of funds contributed, co-authorship and co-patenting. A promising source of data is identified but not described further: researchers and industry partners provide descriptions of research outcomes and benefits for final grant reports. This is an interesting data source that may be tapped for qualitative analysis of the impact of HASS and STEM research.

The ARC commissioned a report, *A wealth of knowledge: the return on investment from ARC-funded research* (Allen Consulting Group 2003a), which used an economic approach. While the economic return on investment is important, there are missed opportunities for HASS: when the report addresses the 'social rate of return' of ARC research, we find that this is defined as 'the permanent increase in GDP as a percentage of the dollar cost of the investment that led to this increase'—in reality, an economic calculation. The report goes on to say that the 'estimate of returns relating to ARC-funded R&D does not fully capture all possible sources of benefit. Health, environmental, social and cultural benefits for instance have not been [sic] quantified in this study' (Allen Consulting Group 2003a, 1), 'but which are nevertheless valued by the community' (Allen Consulting Group 2003a, 10). However, this report does identify a range of interesting research benefits: 'benefits from building the basic knowledge stock', 'direct benefits from improving the skills base', 'benefits from better-informed policy-making' and 'health, environmental and cultural benefits' (Allen Consulting Group 2003a, 6). Due to the limitations of the study, these benefits are quantified in financial terms, but no indicators are produced. However, we can envisage these categories forming the basis for providing qualitative indicators of research impact (see Section A4).

### *Humanities and creative arts*

The ARC's draft review of HCA (ARC 2004) addresses research impact but does not discuss any measures. However, the range of the discussion, combined with the descriptions of data held by the ARC, points the way towards the future development of qualitative indicators of research impact. As outlined above, the report identifies international collaboration and linkages with industry, government and community, but more interesting for our purposes is the detail given on the contribution of ARC-funded research to the Australian community, partly derived from a survey of leading researchers about the impact of their research (ARC 2004, 15, 145–152). For example, the following areas emerge:

- 'Understanding ourselves and our community': enabling Australians to better understand themselves and others; addressing major contemporary social and cultural challenges including the impact of new technologies on society; playing a role in building Australia as a cohesive, sustainable society that values tolerance and respect for others.
- 'Understanding other societies and cultures': enabling Australians to better understand themselves and others; addressing major contemporary social and cultural challenges including changing world security and globalisation.
- 'Informed public debate, improved policy-making and economic impact': fostering the capacity for informed public debate; improving policy particularly through research contributions in law, architecture, cultural studies, media studies, anthropology and Asian studies.
- Knowledge/education: providing an enriched 'ideas base' and improved education curricula which bring community benefits.
- Cultural: enriching public life through the research development of exhibitions, performance and film; providing expertise in supporting national collections in museums, galleries, libraries and archives; contributing to an innovation economy through the creative and cultural industries, particularly in the areas of design, digital arts, computer games, publishing, film, television and radio, and music.

These categories, combined with researcher and partner descriptions of benefits contained in end-of-grant reports, point the way towards the development of qualitative models of research impact informed by academic and 'user' perspectives, a point discussed in further detail in Section A4.

Measures of research impact in the social, behavioural and economic sciences are not similarly reviewed by the ARC.

## The National Health and Medical Research Council

In the 2003–04 financial year, 'NHMRC and Other Health' received \$428.3 million, or 8.0% of the Australian Government's science and innovation budget. The NHMRC is dedicated to research, ethics and advice in health and medicine. Some HASS research falls within the NHMRC's remit, although as we shall see it is not separately distinguished in processes of identifying research quality and impact. In September 2004, the NHMRC adopted a performance management framework that uses the Department of Finance and Administration's preferred 'outcome–output' model for performance measurement.<sup>46</sup> The outcome–output framework demonstrates what the NHMRC produces ('outputs') and the consequences for the community ('outcomes'). Various headings include 'Creating new knowledge', 'Utilising knowledge', 'Enhancing capacity to innovate', 'Ensuring high ethical standards' and 'Strengthening communications and collaborations'. These are supported by a range of indicators, discussed below, for measuring research quality and research impact.

### *Measures of research quality*

Several indicators of research quality appear in the NHMRC's performance management framework, although data sources are not revealed and technical details are not discussed. In the category of 'Creating new knowledge', the subcategory 'Growth in internationally competitive knowledge' employs measures of the number of publications and citations per publication, presumably based on journal-based output, which allows for comparison with national and Organisation for Economic Co-operation and Development (OECD) averages. A further category, 'Enhancing capacity to innovate', includes 'Increased international recognition of Australian researchers', which employs new measures; data to be supplied and benchmarked in 2005 will cover non-bibliometric indicators (awards, prizes and other honours, keynote addresses, presentations and speaking engagements).

The NHMRC also commissions bibliometric studies of journal publications attributable to its funding (Butler 2003b); the measures used include number of publications, citations per publication, actual and expected citation rates, and most highly cited publications. There is also analysis of national and international collaboration and level of research (e.g. various grant schemes, fellowship schemes, research institutes and other sectors).

The measurement of NHMRC research quality does not separately analyse any HASS research funded, although the introduction of non-bibliometric indicators would likely be favourable to HASS.

### *Measures of research impact*

A particularly interesting feature of the NHMRC's performance management framework is that it incorporates a variety of measures of research impact. Many of these indicators are new, and data will be supplied and benchmarked in 2005. Again, data sources and technical details are not revealed, but are on the cusp of new directions in measuring research performance in STEM which embrace research impact beyond the academic realm. Under the category of 'Utilising knowledge' there are indicators of 'Increased uptake of NHMRC health advice' including (non-academic) citation rates of NHMRC advice and stakeholder awareness and satisfaction with NHMRC advice and information (to be identified by stakeholder groups through a benchmarking exercise). For 'Improved transfer of knowledge into health policy and practice', indicators include the trend in NHMRC research rolled over into further research, research leading to changes in public health policies and practices, and the uptake of new or updated advisory products. For 'Increased commercial activity', a new measure is the growth in research leading to commercial outcomes. A further general category of 'Strengthening communications and collaborations' includes indicators for 'Growth in collaborations and partnerships', including the level of external funding from collaborators and the number of national and international collaborations and partnerships. Finally, 'Increased engagement with the community' incorporates the trend in consultations with the public and stakeholders (which appears to be measured in terms of positive feedback gained); growth of links with other organisations; demand for NHMRC products (sales and download levels); stakeholder awareness of NHMRC support; advice and satisfaction levels; community awareness of issues addressed by the NHMRC; and changes in community attitudes on issues addressed by the NHMRC.

<sup>46</sup> See <<http://www.nhmrc.gov.au/aboutus/pmf.htm>>

It is not clear whether the aim is to quantify all data, or whether qualitative interpretations will be incorporated. It is, however, interesting to note that within the category of ‘Creating new knowledge’ there is a heading of ‘Research results with outstanding scientific impact or potential to impact highly on public health policy and practice’. This is a new indicator to be benchmarked through a peer-review process in 2005, the outcomes of which could be presented in either quantitative or qualitative terms. The NHMRC is certainly not averse to the idea of qualitative benchmarking, currently being considered by its Research Committee on Measures of Research Impact and Achievement, in its plans for allocating research funds through its Programs Grants Scheme on the basis of previous research impact, especially to humanity, through gains in knowledge, health and commercial development. The idea is to use peer review, combining quality and impact data relating to each of these areas on scales spanning from ‘activity’, ‘recognition’ and ‘acclaim’ to ‘global impact’. However, while this approach is currently only being developed to assess research proposals, it merits further discussion as a potential approach to measuring research outcomes (see Section A4).

Again, we find that any HASS research is not separately identified or evaluated.

### **A2.3.3 Conclusion**

In terms of measuring research quality, we find that all agencies collect data on the number of publications produced, that records often extend beyond the core DEST categories, and that ARC records include creative outputs. We have previously noted that the number of publications or ‘outputs’ produced does not necessarily correlate with research quality. A common thread is that all agencies use standard journal-based citation indicators to denote research quality. While this practice is well suited to the natural and experimental sciences, it is problematic and disadvantageous when applied to HASS, as elaborated in the ARC review of HCA (ARC 2004) and in Section A2.4 below. This is a vital consideration if citation indicators are proposed to replace the current publications element in future funding arrangements.

AIMS, ANSTO, CSIRO and the NHMRC all gauge research quality using non-bibliometric indicators such as awards, prizes and honours, and the ARC holds similar data. These non-bibliometric indicators are more promising measures of research quality for HASS and can in principle be applied equitably between the STEM and HASS sectors.

For measures of research impact, all agencies produce indicators relating to research commercialisation or technology transfer, and for advice given to government, while the NHMRC has also introduced an indicator of its actual policy influence. The ARC and the NHMRC measure levels of international collaboration and community linkages. The NHMRC is unique in gauging non-academic citation of the research it funds, and in measuring community awareness of its research and any changes its research produces in community attitudes.

The most exciting area in development is the lead that CSIRO and the NHMRC have taken with their outcome–output models, which potentially combine quantitative and qualitative data with peer assessment, and present HASS with the opportunity to design similar models for research impact assessment. Such models can highlight what HASS has in common with STEM and also what makes HASS research distinctive and worthy of public financial support. The ARC’s report on research in HCA lists in detail the impacts of its research upon the community, culture, public policy, debate, the economy, knowledge and education, but the ARC has not yet operationalised this thinking in the style of CSIRO or the NHMRC.

It is curious that it is STEM which is most actively promoting its wider social and community benefits, when HASS and its spectrum of activities have the most to gain by adopting this approach. The humanities and creative arts are a step closer to this, but the social sciences remain silent to date. The social sciences are the HASS stream with the expertise to be at the forefront of developing such indicators and experimenting with the types of quantitative and qualitative representations of research impact discussed in Sections A4.1, A4.3 and A4.5 below.

## A2.4 Measuring research quality: a closer look at citation indicators and HASS

We have seen that AIMS, ANSTO, the ARC, CSIRO and the NHMRC use journal-based citation indicators derived from the ISI's Science Citation Index, Social Science Citation Index, and Arts and Humanities Citation Index as proxies for measuring the quality of publicly funded research.<sup>47</sup> These indicators can be benchmarked against the performance of fields or subfields by unit, department and institution, both nationally and internationally. The most common citation indicators used are citations per publication or set of publications (often compared with expected citation levels), most highly cited publications (i.e. the top 1% or 5% of most frequently cited publications), and journal impact factor (the average number of citations to articles in a journal in a particular period). This review has observed that practice in evaluating STEM research tends to be applied to HASS wholesale, and that any policy move to amplify the importance of citation-based indicators will place HASS research at a disadvantage, a concern now elaborated using insights gained from the specialist bibliometrics literature on citation indicators.

It is important to note first that the bibliometrics literature insists that citation-based indicators do not reveal the inherent quality of publications: citations may be positive or negative, and a paper may be highly cited because its findings are contested. Bibliometricians therefore prefer to talk of citations as indicators of research impact, or of the influence written work has on subsequent literature. However, this definition of influence is purely in internal academic terms (i.e. academic papers that are cited by other academic papers) and is not the broader definition of 'impact' adopted by the RQF consultation process and hence used in this review.

### A2.4.1 Citation indicators, science policy and HASS

The literature speculates about why bibliometrics has come to be applied to the social sciences and humanities, and there is consensus that the driving force has been the desire of funding agencies or policy makers to evaluate the whole research sector. There has been 'considerable debate' between academics and research evaluators about the 'usefulness of bibliometric indicators as an evaluative tool for the social sciences' (Katz 1999, 1) motivated by a need for 'humanities and social sciences to develop methodological tools to assist evaluation agencies or policy-makers in carrying out their tasks, in the same way as current SCI-based methodologies provide supplementary research assessment tools in the natural and life sciences' (Luwel et al. 1999, 13). However, while Nederhof et al. (1989, 425) earlier acknowledged that 'bibliometric tools have proved their usefulness as monitors of developments in the natural and life sciences', they cautioned that 'evidence on this point is almost completely lacking for the humanities and many of the social science disciplines'; these concerns persist today.

### A2.4.2 Coverage of HASS publications

The most pressing issue concerning the coverage of Australian HASS literature is the relatively small proportion of this that appears in journal form. Butler (2004a, 153–4) calculated that in 1995 just under 50% of all Australian publications in the social sciences and humanities were in article form, and that in 1991 the proportion of all Australian journal articles covered by ISI's databases varied widely: law (4%), arts (11%), political sciences and public policy (23%), historical studies (40%), population studies (44%), economics (55%), sociology (61%) and psychology (62%). HASS compares badly with Australian STEM journal article coverage: pharmacology (81%), clinical sciences (81%), astronomical sciences (82%), botany (83%), chemical sciences (83%), neurosciences (85%), immunology (87%) and biochemistry (93%).

ISI-based citation studies are confined to studying journal literature and the citations made by the journal literature, and the figures above demonstrate that, while STEM coverage is generally high,<sup>48</sup> we must question whether bibliometric approaches suited to measuring research output or research impact in the

<sup>47</sup> For a comprehensive review of quantitative bibliometric and non-bibliometric indicators, including a detailed analysis of all forms of citation indicators, see *Quantitative indicators for research assessment—a literature review* (REPP 2005), from which Sections 2.4 and 4.4 of this report are derived. <<http://repp.anu.edu.au/Literature%20Review3.pdf>>

<sup>48</sup> It is interesting to note that engineering is less well served by journal-based indicators, as the discipline relies heavily on technical papers which are not indexed by ISI.



natural and experimental sciences can legitimately be applied to the social sciences and humanities. In other words, 'standard' science (ISI journal-based) approaches do not capture the essence of those disciplines in which 'non-standard' publications (books, monographs, chapters and other non-ISI literature<sup>49</sup>) are important or even dominant, and creative works are not served by this approach at all. Nederhof and Noyons (1992, 255) caution that 'focus on article publications alone could thus lead to seriously flawed results', and Hicks maintains that by replicating practices suited to science publication 'bibliometric evaluation produces a distorted picture of social science fields' (Hicks 2004, 473).<sup>50</sup>

A further issue is that indicators designed to measure research in the natural and experimental sciences will detect research most like that in the natural and experimental sciences, illustrated by the greater coverage of highly quantitative social science research, particularly in economics and psychology, which more closely mirror scientific publishing practice through a more journal-based and international orientation (van Raan 1998, 3; Katz 1999, i). There is an inherent danger that undetected social science literature may be viewed as 'soft' or less mature, or that invisible research will simply remain unacknowledged.

A further issue for Australian HASS is the relatively small number of Australian-oriented journals covered by ISI's databases. Such journals lose visibility alongside predominantly American and European publications. In a study of Australian university bibliographies, Bourke et al. found that, by excluding non-source items (and due to the under-representation of the 'periphery' of Australian and regional journals), journal-based indicators in the social sciences and humanities used as a 'surrogate for total publications citation rates will be more misleading than in the sciences' (Bourke et al. 1996, 54). Royle and Over make the strongest criticism relating to the social sciences in Australia when they argue that 'reaching conclusions about the relative research productivity of individuals or universities, or seeking to compare disciplines or nations, by reliance solely on the database used to compile the ISI source indexes is not a worthwhile exercise' (Royle and Over 1994, 86).

### A2.4.3 Conclusion

It is clear that citation-based indicators will disadvantage HASS when database coverage is limited to journal literature: books, book chapters, 'enlightenment' literature and creative works are excluded, several regional journals are not covered by the ISI, and regional issues are less likely to be cited by the American and European journal papers that dominate the database. There is a case for applying standard citation-based indicators to some social science fields. However, there is a stronger case for developing new non-standard measures of research quality using 'metrics' suitable for both HASS and STEM.

For example, if database coverage changes to include a wider range of publications, HASS may be better served by citation-based indicators. In this respect, there is some interesting work in progress (Butler and Visser, in press) which applies citation analysis to non-source data currently provided to DEST by Australian universities, including books, book chapters, journals not indexed by ISI, conference papers and technical reports. This research is discussed further in Section A4.3. There is also much scope for developing non-bibliometric quality indicators (also discussed further in Section A4.3), which are currently adopted or collated to some extent by AIMS, ANSTO, the ARC, CSIRO and the NHMRC. Bibliometricians would advocate that citation and other 'quality' measures should not be used in isolation, but rather selectively combined and applied to appropriate subfields of research within an overall context of expert peer judgment.

<sup>49</sup> Other non-ISI literature includes 'grey' or 'enlightenment' literature such as internal reports, reviews, notes, and literature aimed at practitioners and the general public in the form of non-scholarly journals, media contributions and other specialist publications.

<sup>50</sup> For further concerns about Arts and Humanities Citation Index and Social Science Citation Index coverage, see Nederhof and Zwaan (1991, 323–333), Nederhof and Noyons (1992, 255), Burnill and Tubby-Hille (1994, 148), Glänzel (1996, 292), Cronin et al. (1997, 264; 270) and Hicks (1999, 197, 210–212).

## A3 Current international measures of quality and impact of publicly funded HASS research

The aim of this section of the review is to examine the measures of quality and impact of publicly funded HASS research currently adopted by other countries, and to note the use of any novel indicators.<sup>51</sup> It is not the intention to describe the research evaluation frameworks adopted by various countries, such as formula-based block funding or peer review-based research assessment exercises, but rather to describe the indicators that can in principle feed into any overarching RQE. This overview of the international scene takes as its starting point a comprehensive study of research assessment practices (von Tunzelmann and Kraemer Mbula 2003), which does not discuss in detail the types of indicators used but does identify any HASS-specific policy initiatives. Web searches were conducted on a country-by-country basis to determine the kinds of measures used and, where the information was accessible, the specific indicators employed.

Many countries have not developed a national system of research evaluation tied to measures of research quality or impact, but the trend is very much in this direction. For example, France, Germany and the United States are currently experimenting on an institutional or regional basis. We find that the Netherlands and New Zealand are the countries with the most innovative approaches to measuring both research quality and research impact, while in the United Kingdom several initiatives are in development, ranging from standard and non-standard bibliometric measures through to more abstract conceptions of HASS impact. In this light, we are on the cusp of major international developments and changes in the application of 'metrics' to research evaluation.

### A3.1 Measures of research quality

In the review of current practice in 18 countries, we find that the Netherlands, New Zealand and the United Kingdom have the most to offer in measures of research quality, either currently being used or in development.

In the Netherlands, national assessment of research is carried out internally by universities. Currently, assessment is comparable but not rigorously comparative between departments. Self-assessment is carried out every three years, and external evaluation every six years. This exercise currently has no connection with the level of funding received, but is in force to improve the public accountability of research activity. Several measures of research quality are employed (see VSNU 2003) and fall under various meta-labels. For example, 'Productivity' includes bibliometrics (number of publications and citations), technometrics (number of patents and citations of patents), and sociometrics (quantifiable socioeconomic performance of 'embeddedness' of research). Publications are taken from the previous six years and include papers in refereed and non-refereed journals, books and monographs, book chapters, PhD theses, and professional publications and products. Research 'Relevance' is sought through developments in the international scientific community (user relevance is also gauged, and is discussed below in terms of impact). 'Vitality and feasibility' of research group performance is measured, including planned research directions and the quality of project management. Other data are sought relating to esteem, including academic awards, prizes, editorships of academic journals and memberships of scientific boards. Efforts have been made to standardise the evaluation methodology across all disciplines, an approach mainly directed at evaluating science research. It has, however, been acknowledged that from a HASS perspective the extended, uniform use of quantitative publication indicators is problematic.

In New Zealand, the Foundation for Research, Science and Technology introduced 'outcome indicators' in 2001, and collects data for six main performance indicators (FRST 2004). Metrics and peer review are combined for evaluation and funding purposes. Indicators of research quality are employed under the performance indicator of 'Intellectual property' (IP) in terms of the number of IP items which have been commercialised or for which protection is being sought. The performance indicator 'Publications, awards and/or keynote presentations' uses measures in the form of the number of published scientific publications

<sup>51</sup> 18 countries are investigated: Austria, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, the Netherlands, New Zealand, Norway, Sweden, Switzerland, Taiwan, the United Kingdom and the United States.

in peer-reviewed journals or books (including extended abstracts in journals, conference proceedings and Masters and PhD theses). Non-bibliometric indicators include awards, defined as ‘any recognised science achievement for research’, and keynote presentations, defined as ‘any plenary/invited written or oral presentation delivered at a recognised science forum’.

In the United Kingdom, it has been recommended that quantitative performance indicators should play a significant role in the 2008 Research Assessment Exercise (Roberts 2003). However, while there is a list of metrics in the guidance given to panels for the 2008 RAE, it does not mention bibliometrics or any other indicator of quality or impact (HEFCE 2005). To date, as esteem measures, departmental ‘units of assessment’ have been able to provide written evidence of peer esteem and evidence ‘of significance’ which will be regarded highly by a panel of reviewers. However, publication indicators remain undefined. Indeed, neither the guidelines nor the Roberts review (2003) specify what ‘metrics’ are, but say that these are still in development. Some light is shed in a discussion on performance indicators: ‘other measures, such as esteem indicators or bibliometric measures, might also be incorporated where they were more closely approximate to quality’. (Roberts 2003, 210). It is recommended that the United Kingdom’s funding councils and research councils work to develop discipline-specific performance indicators, and that ‘the weight placed upon these indicators as well as their nature, should be allowed to vary between panels’ (Roberts 2003, 50). So while quality metrics are a work in progress there is, nonetheless, sensitivity shown to HASS research with the acknowledgment that ‘reliable quantitative performance indicators may prove hard to develop in some disciplines, especially the arts and humanities’ and that the process should be ‘sufficiently flexible to ensure that, where the subject community has less confidence in performance indicators, the indicators would not be over-influential in determining the outcome of the assessment’ (Roberts 2003, 212). The weight of such measures is expected to vary between subject areas, and ‘such variation a healthy reflection of the genuine differences between disciplines’ (Roberts 2003, 213).

The United Kingdom Arts and Humanities Research Council accordingly has major policy initiatives underway, including developing metrics for arts and humanities research intended to ‘capture the depth and diversity of knowledge generation and exploitation in the sector’. However, in the face of an outcry from its constituency, the council has abandoned an attempt to list the top ten ‘most significant and important’ journals in ten subject areas. Objections included the claim that such an exercise would kill off emerging disciplines and publications, and could ruin developing research careers if subsequent competition meant that only the established elite appeared in these ‘top’ publications (THES 2005). The United Kingdom Economic and Social Research Council is developing social science bibliometrics to estimate the citation level of non-ISI journals, establish a hierarchy of book publishers, identify the top journals in each discipline, and investigate quantitative measures of policy-related work, such as citations to government reports.

In other countries, we find that Flanders (Belgium) has produced the most concerted effort to introduce bibliometric methods for evaluating university research performance. Block funding is weighted, with 50% allocated on the basis of student numbers, etc., and 50% using standard journal-based bibliometrics derived from the ISI’s indices and patent and scientific innovation data. However, ISI data is not being used for HASS due to recognised limitations of ISI coverage for HASS in general and in small, non-English speaking countries in particular. In a similar vein, Norway is seeking a way of employing measures to evaluate quality and efficiency in research performance. However, it is not including scientific publishing in a funding formula because of the lack of reliable data, but is producing a database for number and type of publication. Similarly, Finland does not use citation indicators and counts publication numbers, although no detail is apparent on the types of publications covered. Finland does collect ‘quality’ data in the form of the number of foreign visitors a department receives. Taiwan is considering employing bibliometric profiling, although it also recognises problems with ISI coverage of HASS and the non-English language barrier.

### **A3.2 Measures of research impact**

Again it is New Zealand, the Netherlands and the United Kingdom which have taken the lead in developing measures of research impact. For example, for the performance indicator ‘New or Improved Products, Processes, or Services Developed for Users’, New Zealand seeks information on any policy advice given. An interesting impact indicator is ‘Reports, Presentations, and Publications for Users’, which seeks data



on published outputs targeted directly at users; reports commissioned by users under contract or formal arrangement; presentations at which users are in attendance or which involve contact with users (papers, oral presentations, static displays and web presentations); publications which are not peer reviewed but intended for users (papers not in peer-reviewed journals, series or books; conference papers or abstracts, excluding keynote presentations; research books or monographs; popular books/articles; web-based publications, including websites and downloadable files; and trade journals/magazines). Finally, for the performance indicator 'Partnerships, Contracts and Linkages', these relationships are counted and a brief description of each is sought. Here, 'partnership' is defined as a formal relationship with another party (an organisation or individual) with whom there is some form of alignment of objectives or strategy, or some close association; a 'contract' as a legal agreement entered into for a defined piece of know-how arising from research, usually in exchange for money; and a 'linkage' is a less formal relationship formed on the basis of sharing information and/or resources.

For 'Relevance', the Netherlands measures research impact in the form of its influence upon developments or questions in society at large. The United Kingdom is constructing possible impact metrics, and notes that it is desirable for users and commissioners of research to actually be involved in the assessment process to give their perspective on the range of research under question (Roberts 2003). 'Units of assessment' are free to provide information on relationships with research users, including industry and commerce. The United Kingdom Arts and Humanities Research Council is breaking promising ground by developing 'a radical new approach to impact assessment based on users of research knowledge rather than producers'. Finally, in the United States, Hicks et al. (2004) note that, while peer review has been adopted rather than quantitative indicators for research evaluation, there is growing interest in impact indicators in policy circles, spurred particularly by the desire to evaluate the economic and social outcomes of research. However, these observations are related to STEM-based research only.

### A3.3 Conclusion

For research quality measurement, we find novel ideas currently in development for the somewhat controversial notion of creating 'top ten' lists of journals in various disciplines, and other non-standard bibliometric indicators such as the citation level of non-ISI journals and a hierarchy of book publishers. A common message is that HASS is not well served by standard bibliometric indicators, particularly in small countries, and even more so in non-English speaking countries. There is also a tendency for evaluation systems to be science-based, particularly with respect to publishing-related components, and there is a danger for HASS in standardising the evaluation methodology across all disciplines.

Interesting areas for development in assessing research impact include quantifying the array of publications, presentations and displays aimed directly at users, and involving users and commissioners of research in the assessment process. We find that New Zealand and the Netherlands evaluate research in categories which may be adapted to impact modelling, highlighting the user and societal benefits of research, and user and public interaction with research products.

The idea of self-assessment on the basis of indicators of quality and impact is also novel.

## A4 Potential measures of quality and impact of publicly funded HASS

This section draws together several conclusions concerning the range of current measures available to evaluate the quality and impact of publicly funded HASS research in Australia. Sections A4.1 and A4.2 present lessons which we may glean from Australian and international practice. Section A4.3 introduces promising new ways of measuring research quality and impact; these are being postulated but are not currently adopted in standard international or Australian practice. Section A4.4 takes a closer look at measuring research quality using non-bibliometric indicators, and Section A4.5 outlines cutting-edge peer-review and user-based qualitative models which may be adapted to measuring the impact of HASS research.

We have seen that ‘research quality’ is taken to refer to the value of research ‘outputs’ as judged by other academics, and that this is measured using standard journal-based quantitative publication indicators. Quality may also be accessed through evidence of peer esteem, such as honours, awards and prizes received. Measuring research quality is highly amenable to quantitative methods, which should always be used alongside expert discipline-specific opinion. On the other hand, ‘research impact’ denotes the utilisation of knowledge beyond academia in the form of economic, policy, social, community, cultural and artistic benefits. Impact is more open to qualitative assessment incorporating user/beneficiary opinion.

The general message is that quantitative indicators, particularly standard journal-based citation studies, can be very disadvantageous for HASS, although non-standard bibliometric analyses and esteem measures will be more favourable. Impact measures may hold the greatest potential for HASS streams to demonstrate their relevance and the benefits they bring to richly varied aspects of the nation’s life, although STEM has currently stolen the march in this area.

For measuring HASS research quality, the most promising areas to develop appear to be non-standard bibliometric analyses (i.e. of publications other than ISI-indexed journal papers) and non-bibliometric indicators of esteem. Likely measures of research impact to pursue relate to public and policy uptake of HASS research. The idea of developing qualitative models of research impact by HASS stream breaks new ground: by following the lead of STEM, this is an area in which HASS can effectively demonstrate its importance and relevance in its own compelling terms. Exploring such novel indicators of quality and impact will allow the arts, humanities and social sciences to move beyond the disadvantages inherent in the traditional performance measures to which, unfortunately, the sector is still wedded.

#### **A4.1 Lessons from Australian practice**

To date, no measure of research quality or impact has been linked to the public funding of Australian research, but what may HASS and policymakers learn from current institutional practice in this respect?

We find that the publicly funded research agencies (AIMS, ANSTO and CSIRO) and the research councils (the ARC and the NHMRC) all employ standard journal-based citation indicators to gauge research quality, a practice which is extremely disadvantageous for HASS. The humanities and social sciences would benefit from the introduction of non-standard bibliometric measures of citation counts which are currently in development and are drawn from existing extended university DEST returns (see Section A4.3). As for the arts, some of these returns and ARC records give the number creative works produced, and while research production does not predict research quality, the potential exists to include creative performances and products aimed at the public or users under the rubric of ‘impact’ measures. The above agencies all employ non-bibliometric indicators or hold data on research quality, such as data for honours, awards and prizes. A range of quality measures therefore already exists, and the gaps identified may be addressed by developing novel indicators and tapping into existing Australian initiatives.

In policy terms, we may be heartened to note that the RQF Issues Paper observes that any measures of quality and impact should be ‘sufficiently flexible to accommodate the breadth of different fields of study’, that ‘it is important that excellence in humanities and social sciences can be compared with that of the physical sciences on a level playing field’ and that a “one size fits all” approach is likely to be deemed inappropriate’ (DEST 2005, 16). While the arts are overlooked in this instance, this is likely to have been an oversight, and these sentiments are extended in the observation that ‘in the arts and humanities, it may be more relevant to include performances, exhibitions, and other esteem indicators, whereas in the sciences, more focus might be given to citation indices’ (DEST 2005, 20).

The above bodies employ indicators of research impact relating to commercialisation and technology transfer, and advice given to government. Various other measures are distributed between them, such as actual policy influence, international collaboration and community linkages, non-academic citation of research, and community awareness and uptake of research findings. CSIRO and the NHMRC have taken the lead with their respective outcome–output models of research impact and in combining quantitative and qualitative data with peer assessment. A valuable lesson is that the opportunity exists to apply similar models of research impact to HASS; this could highlight what the HASS sector has in common with STEM

and also what makes HASS research distinctive and worthy of public financial support. This would be a step beyond approaches which tend to quantify HASS qualities and reduce impact to financial indicators (e.g. Allen Consulting Group 2003a), and this alternative strategy presents the opportunity to explore qualitative analyses and impact visualisation models already embraced by the Australian STEM sector. A more poignant lesson is that the HASS sector has been unnecessarily conservative in its thinking about promoting the value of its research through restricting its thinking to presenting ‘impact’ in purely economic or commercial terms.

## **A4.2 Lessons from international practice**

In OECD countries, we find that science policy founded on science principles provides the framework for regulating both STEM and HASS research, and this policy resonates with the terms ‘utility’, ‘commercialisation’ and ‘wealth creation’. The idea is that governments believe that scientific discovery creates social and economic progress, and wish to harness research in pursuit of the dual goals of national technological advance and enhanced international competitiveness. In pursuing these goals, governments wish to derive maximum utility out of finite public funds—hence the need for research evaluation to ensure efficiency and value for money (Donovan 2005). The use of indicators is desirable because independent measures are perceived to be more neutral than peer-review processes in which academics self-regulate their activities. A consequence of such wealth-maximising, STEM-dominated and indicator-led thinking is that the distinctive qualities and benefits of HASS research are likely to be lost within standardised evaluation systems. This has indeed been the case in several countries where HASS is unreflexively judged by STEM-derived quality measures, most notably standard journal-based citation analyses.

However, several countries are sensitive to the plight of HASS, and there are numerous initiatives underway to develop more equitable or separable metrics, most notably in the direction of creating non-standard bibliometric indicators which capture non-ISI journals, books and chapters. New Zealand and the Netherlands also provide the opportunity for the impact of HASS research to be identified in terms of public and user benefits and public and user interaction with research products, including publications, presentations and displays. There is a trend to involve the users and commissioners of research in the actual assessment process, and also much concern to establish international benchmarking in research evaluation. There is much work underway in the development of novel measures of quality and impact of publicly funded HASS research, particularly in the United Kingdom, where it is hoped that these metrics will feed into the 2008 Research Assessment Exercise.

Australia is in a position to cherry-pick the most desirable aspects from international and local institutional practice, and to eschew various undesirable elements. The RQF can adopt quantitative and qualitative measures of quality and impact which are sensitive to a variety of disciplines and fields, and which are tailored to identifying and stimulating the best research outcomes Australia can offer in STEM and HASS alike.

## **A4.3 New ways of measuring quality and impact currently being postulated**

### **A4.3.1 Measures of research quality**

We find that there are a couple of new suggestions for measuring the quality of HASS research. The first relates to music (Monash University School of Music 2004, 4–6) and suggests an internal university points allocation system for various musical creative outputs, including published musical scores, commercially published or broadcast original compositions, and recorded solo performances at commercial concerts in well-known venues (with a deposited CD and a five-page or longer written account of the performer’s research into the work and its performance). The proposal also includes points allocations for national television broadcasts (ABC or SBS) of ethnographic films about music, and commercially released ethnographic field recordings of 40 minutes or more with an accompanying substantial booklet of 30–40 pages containing an essay and research-based commentary.

Work currently in progress (mentioned briefly in Section A2.4) is investigating non-standard citation analyses that promise to be more HASS-friendly than standard ISI journal-based citation studies (Butler and Visser, in press). This research uses existing information provided by universities in their DEST returns,

which often extends beyond the four main DEST categories, and mines the Web of Science<sup>52</sup> for citations to books, book chapters, non-ISI journals and conference proceedings. This approach has been applied to some STEM areas and has so far been extended to economics, history and philosophy, although more testing is required.

#### A4.3.2 Measures of research impact

While discussions on research measurement tend to be quality focused, we have noted an increasing interest in appraising research impact. This is an endeavour which is in genesis, and the potential exists for HASS to actively engage in this process rather than have STEM-led impact measures imposed upon it. A recent discussion on the impact of publicly funded research provides an interesting analysis of how the benefits of this research can be measured (Allen Consulting Group 2005). While that discussion is largely STEM-focused, HASS is mentioned, particularly the social sciences. Impacts are divided into four main societal values: material, human (largely related to health and wellbeing), environmental and social. The social is defined as ‘encompassing social attachments, freedom from crime, the level and security of political rights and the extent to which the population engages in political processes’; indicators include ‘marriage and divorce rates, persons living alone, participation in voluntary work’ (Allen Consulting Group 2005, vii). There is also a focus on economic versus non-economic benefit approaches, although the non-economic is confined to health-related issues. We find that social science is understood in technocratic, Fabian terms and as an activity that can provide a ‘policy fix’. This may be a methodological consequence of deriving benefits from categories covered by Australian Bureau of Statistics social statistics rather than from the viewpoint of researchers or beneficiaries, and the resulting picture of social science we are presented with is one that makes sense to natural scientists or to economists.

In the same study, we find that the humanities and creative arts are considered in interesting if occasionally hedonistic ways. For example, research may improve the quality of people’s lives through ‘enhancing people’s understanding of the world and broadening possibilities for imagination and thought’, and ‘research may increase the possibilities for pleasurable experience’ by ‘contributing to the cultural, literary and entertainment products that people can experience and enjoy’ (e.g. computer gaming and history of theatre) (Allen Consulting Group 2005, 22). There is consideration of the contribution of HASS research to the previously defined social dimension, including intervention to reduce social exclusion; explaining Australian history and current society to contribute to common identity and community; providing an independent voice in policy debate and increasing political accountability; and research in history, political science and philosophy highlighting dangers associated with particular political and legal structures, and thereby raising community awareness of the importance of protecting certain legal and political rights (Allen Consulting Group 2005, 23). However, no HASS impacts are operationalised. We are presented with familiar indicators which we would mostly classify as ‘quality’ measures in the context of our study, and a couple of novel impact ideas are introduced which may be applied to the HASS sector: surveys of heads of policy sections in government departments about whom they regard as ‘high impact’ academic researchers, and the number of research students employed in government departments, ministerial offices and industry (Allen Consulting Group 2005, 43).

The International Federation of Arts Councils and Culture Agencies presents a detailed discussion of types of cultural indicators and what these indicators should ideally capture, but none are operationalised (IFACCA 2004). However, the discussion does not have a higher education or research policy focus, and the impact indicators are somewhat tangential to creative arts research, the only detailed example discussed being a general indicator for employment in cultural occupations.

One area of especial interest is an impact modelling exercise conducted jointly by Brunel University’s Higher Education Research Group (HERG) and RAND Europe for the Arthritis Research Campaign in order to evaluate the outcomes and outputs (or ‘payback’) resulting from research funding (Wooding et al. 2004). While this approach has been utilised within biomedical research, as in the CSIRO outcome–output model, HASS may plunder the approach and adapt it to its purposes.

The attraction of this approach is that measures of research quality and impact are combined with peer and user evaluation to produce a logic model of the research process (see Figure A2). Category A (‘Knowledge’)

<sup>52</sup> <<http://www.isinet.com/products/citation/wos>>

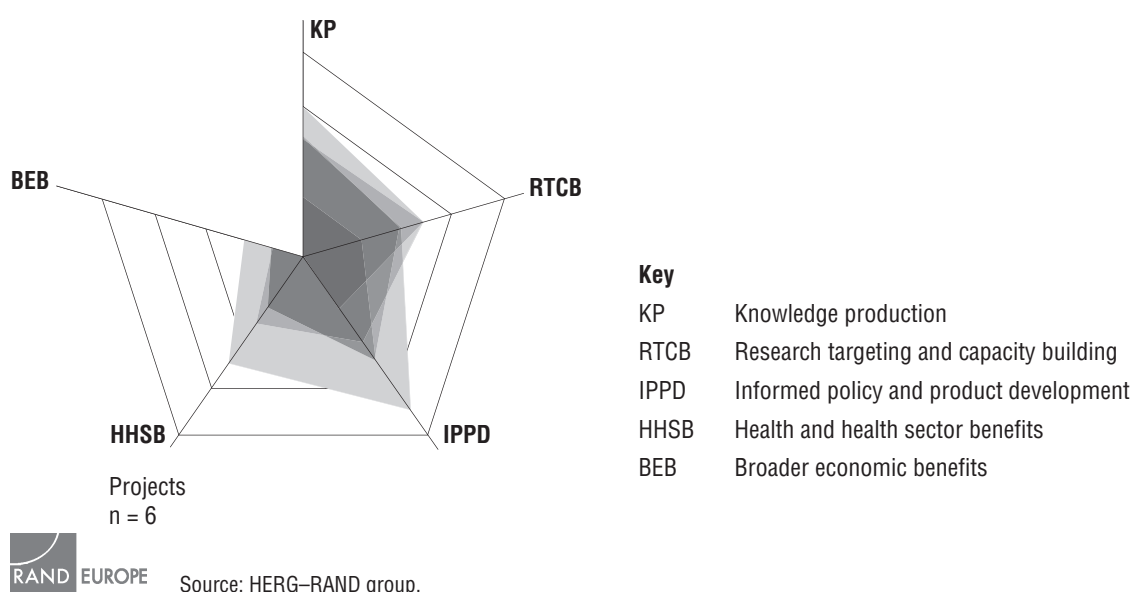
in Table A1 includes bibliometric analyses of publications resulting from research funding, and categories A to E are considered by panels of expert peers and research users who rate each category on a scale of 1 to 9. This is then mapped, resulting in various profiles like that illustrated in Figure A2. The other categories are 'Benefits to future research and research users' (which is largely forward looking and capacity building), 'Political and administrative benefits', 'Health sector benefits' (including cost reduction and qualitative improvements in service delivery, improved equity and IP revenues), and 'Broader economic benefits' (including wider economic benefits from commercialisation and a healthy workforce).<sup>53</sup>

**Table A1:** Payback categories for arthritis research

<b>A</b>	Knowledge
<b>B</b>	Benefits to future research and research users, including <ul style="list-style-type: none"> <li>i. Better targeting of future research</li> <li>ii. Development of research skills, personnel and overall research capacity</li> <li>iii. Critical capability to utilise appropriately existing research, including that from overseas</li> <li>iv. Staff development and educational benefits</li> </ul>
<b>C.</b>	Political and administrative benefits, including <ul style="list-style-type: none"> <li>i. Improved information bases on which to take political and administrative decisions</li> <li>ii. Other political benefits from undertaking research</li> </ul>
<b>D.</b>	Health sector benefits, including <ul style="list-style-type: none"> <li>i. Cost-reduction in the delivery of existing services</li> <li>ii. Qualitative improvements in the process of service delivery</li> <li>iii. Increased effectiveness of services, e.g. increased health</li> <li>iv. Equity, e.g. improved allocation of resources at area level, better targeting and accessibility</li> <li>v. Revenues gained from intellectual property rights</li> </ul>
<b>E.</b>	Broader economic benefits, including <ul style="list-style-type: none"> <li>i. Wider economic benefits from commercial exploitation of innovations arising from R&amp;D</li> <li>ii. Economic benefits from a healthy workforce and reduction in working days lost</li> </ul>

Source: HERG–RAND group.

**Figure A2:** Example of 'payback profile' impact assessment



<sup>53</sup> A report by the Royal Academy of Engineering (2000, 17ff) attempts to measure excellence in engineering research and similarly uses logic or 'footprint' modelling to visualise categories such as 'Strategy', 'Vitality and sustainability' and 'Scholarship' which are generated through mostly quantitative and some qualitative judgments about research group activities.



As with the CSIRO outcome–output approach discussed in Section A2.3.1, we can imagine similar logic models being developed for HASS research, and a particular attraction is the emphasis on research impacts, which can be tailored to HASS streams and disciplines. This approach would also appeal to the desire, in the RQE, for peer- and user-informed research assessment.<sup>54</sup> The potential of this approach for evaluating HASS research is discussed further in Section A4.5.

#### A4.4 Measuring research quality: a closer look at non-bibliometric indicators

Non-bibliometric indicators of research quality are measures that are not derived from published research ‘outputs’; prime examples are indicators of esteem, such as awards, honours and prizes. Such qualitative indicators of esteem are recognised as particularly important for those areas of HASS in which bibliometric indicators are difficult to apply (van Raan 1998, 5). Little research has been conducted to operationalise the use of non-bibliometric indicators, a rare example being an exploratory but comprehensive study of law and linguistics in four Flemish universities by Luwel et al., who demonstrated that ‘bibliometrics is much more than conducting citation analyses based on the ISI indexes, as citation data do not play any role in this study’ (Luwel et al. 1999, 13).

Examples of non-bibliometric indicators of research quality include honours and awards, election to and roles within learned societies, journal editing, editorial board membership, editing special issues of journals, special journal editions dedicated to one’s research, invited lectures at conferences (particularly keynote addresses), organising conferences or workshops, activities in providing academic advice (e.g. assessing research applications, manuscript refereeing, supervision and examination of PhD theses), contributions to dissemination/popularisation of research in the media, policy preparation research, membership of government advisory bodies, membership of a jury for a research award, visiting professorships or fellowships, and conferences dedicated to specific research. This list is by no means exhaustive but is meant to give a taste of the range of possibilities available, and each case can be qualified further.<sup>55</sup>

The mainstream literature tends to discuss non-bibliometric indicators in terms of esteem indicators; the social sciences and humanities, and non-written outputs and the creative arts, have generally been overlooked. However, because of the importance of non-written work, this is an area ripe for novel investigation. For example, Strand’s definition of publication in the creative arts extends far beyond the written word. He notes (Strand 1998, 55) that publication in the creative arts includes:

- *public performance* for dancers, actors, choreographers, musicians and playwrights
- *building or manufacture* for architects and designers
- *written works* for academics, researchers and creative writers
- *exhibitions* for visual artists and craftspeople
- *computer software* for designers, musicians, visual artists and so on.

Again, this is a list that can be extended and refined, but it gives a flavour of the potential for using non-bibliometric indicators.

A novel take on non-bibliometric indicators is suggested by a Royal Academy of Engineering report which, under the heading of ‘Scholarship’, employs what it terms ‘independence indicators’ that denote research quality to the extent that ‘it is free from financial and political control’ (Royal Academy of Engineering 2002, 13). Indicators include ‘Prizes for non-basic, non-private funded research’, ‘Proportion of research not funded by industry’ and ‘Proportion of research not included in Foresight recommendations’ (Royal Academy of Engineering 2000, 28). Non-bibliometric independence indicators can be applied equally to STEM and HASS research.

<sup>54</sup> For a more detailed account of data collection, methods and the evaluation framework, see Wooding et al. (2004, 9–29). <[http://www.rand.org/pubs/monographs/2004/RAND\\_MG251.pdf](http://www.rand.org/pubs/monographs/2004/RAND_MG251.pdf)>

<sup>55</sup> See *Quantitative indicators for research assessment—a literature review* (REPP 2005) for a more technical discussion of non-bibliometric indicators. <<http://repp.anu.edu.au/Literature%20Review3.pdf>>

Non-bibliometric indicators may also be applied to investigating policy impact. For example, Cave et al. produced a report for the United Kingdom Economic and Social Research Council which discussed potential performance indicators for research funded by the council. They examined the possibility of evaluating the relevance of social science research by using ‘policy usefulness’ as a novel non-bibliometric performance indicator, but did not put this method into practice. They believed that ‘it should be possible to scrutinise key policy developments and identify any contribution of research and other disciplined enquiry to, for example, patterns of care and prevention in drug abuse or child abuse’ (Cave et al. 1988, Section 9.1.5), and that ‘[t]his could be done by a process of identifying the relevant group of policy makers or practitioners and interrogating them about their knowledge of research results obtained and the benefits derived from it’ (Cave et al. 1988, iii). However, we may also note that publication-based analyses of policy documents, Hansard and judges’ decisions, for example, may be a more direct route to accessing and assessing research impact.

## A4.5 Measuring research impact: a closer look at qualitative modelling

Sections A2.3.1 and A4.3 of this review describe models used by STEM and the biomedical sciences to assess research impact (namely, the CSIRO’s outcome–output approach and the HERG–RAND ‘payback’ logic outcomes model), and we have speculated on the potential to apply these approaches to HASS. We have also noted that impact measurement lends itself to largely qualitative forms of analysis, and that this process can incorporate peer and user review. We have also seen that publicly funded research organisations such as the NHMRC and AIMS employ broad performance indicators for areas such as ‘New knowledge and collaborative R&D’, ‘Research services, advice and specialised consulting’, ‘Licensing, patenting and spin-offs’, ‘Enhancing capacity to innovate’ and ‘Increased engagement with the community’, which can potentially lead to similar qualitative modelling. It has also been argued that it would be timely and extremely beneficial for HASS to adopt similar approaches, given the disadvantages inherent in standard citation-based research measurement.

It appears that such a move would not involve a great change in thinking, only in the way indicators are represented and used to the HASS sector’s advantage. We find that several possible ‘performance indicators’ or categories have already been proposed or ingeniously elaborated.

For example, the Allen Consulting Group (2003a, 6) identified a range of interesting research benefits, but because of the constraints of the study quantified these only in financial terms. However, we can see that these could be applied more broadly to HASS. These categories are ‘Benefits from building the basic knowledge stock’, ‘Direct benefits from improving the skills base’, ‘Benefits from better-informed policy-making’ and ‘Health, environmental and cultural benefits’. We may already note the emergence of common threads, such as knowledge creation, educational benefits and contributions to policy making.

The ARC Humanities and Creative Arts Review discusses the contribution of ARC-funded research to the Australian community, partly derived from a survey of leading researchers about the impact of their research (ARC 2004 15, 145–152). The following areas emerge,<sup>56</sup> and we can see that these would be appropriate categories for qualitative modelling if appropriate measures of quality and impact are developed to support this approach:

- ‘Understanding ourselves and our community’
- ‘Understanding other societies and cultures’
- ‘Informed public debate, improved policy-making and economic impact’
- Knowledge/education
- Cultural/economic.

<sup>56</sup> These areas are defined in detail in Section 2.3.2 above.



Similarly, the British Academy (2004, vii–ix) examines in detail five key functions of the HASS disciplines:

- contributing to cultural and intellectual enrichment
- contributing to economic prosperity and wellbeing
- contributing new knowledge and understanding of major challenges facing both the United Kingdom and other countries
- contributing to public policy and debate
- providing a rigorous, beneficial and fulfilling education.

We can see striking similarities between these HASS examples, and there is also considerable overlap with STEM and biomedical objectives. It follows that there is much scope to produce qualitative impact modelling based on a composite of various ideas which elaborate what STEM and HASS share, and which are also suited to HASS and its streams and disciplines. For research evaluation purposes, we may for once adapt STEM to the use of HASS, rather than HASS to the use of STEM.

## References

- AIMS (Australian Institute of Marine Science) (2003). *Research Plan 2003–2006*.
- Allen Consulting Group (2003a). *A wealth of knowledge: the return on investment from ARC-funded research*. Report to the Australian Research Council, 4 September 2003.
- Allen Consulting Group (2003b). *The ARC's implementation of government decisions from Knowledge and Innovation and Backing Australia's Ability*. Final report to the Australian Research Council, October 2003.
- Allen Consulting Group (2005). *Measuring the impact of publicly funded research*. Report to the Australian Government Department of Education, Science and Training.
- ANSTO (Australian Nuclear Science and Technology Organisation) (2004). *Annual Report 2003–04*.
- ANU (Australian National University) (2004a). *ANU capabilities and performance statement, report for the Committee of Review of ANU Quality*, July 2004.
- ANU (Australian National University) (2004b). *ANU: University with a difference*. Report of the committee established by the Council of the Australian National University to evaluate the quality of ANU performance, September 2004.
- ARC (Australian Research Council) (2004). *Annual Report 2003–04*.
- ARC (Australian Research Council) (2004). Draft of *Review of the Humanities and Creative Arts Discipline Grouping 2004*, dated 14 October 2004.
- Bourke P, Butler L and Biglia B (1996). *Monitoring research in the periphery: Australia and the ISI indices*. REPP Monograph Series No. 3, Research School of Social Sciences, ANU.
- British Academy (2004). *That full complement of riches: the contributions of the arts, humanities and social sciences to the nation's wealth*. The British Academy: London.
- Burnhill P and Tubby-Hille ME (1994). On measuring the relation between social science research activity and research publication. *Research Evaluation* 4(3):130–152.
- Butler L (2003a). Modifying publication practices in response to funding formulas. *Research Evaluation* 12(1):39–46.
- Butler L (2003b). *NHMRC-supported research: the impact of journal publication output 1996–2000*. National Health and Medical Research Council, Canberra.
- Butler L (2004a). *ARC-supported research: the impact of journal publication output 1996–2000*. Australian Research Council, Canberra.
- Butler L (2004b). What happens when funding is linked to publication counts?. In: *Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S&T systems*, Moed HF, Glänzel W and Schmoch U (eds), Kluwer Academic Publishers, Dordrecht/Boston/London.

- Butler L and Visser M (in press). Extending citation analysis to non-source items. *Scientometrics*.
- Cave M, Kogan M, Hanney S and Burningham D (1988). *Performance indicators for the Economic and Social Research Council: report on preliminary project*, prepared for United Kingdom Economic and Social Research Council.
- Cronin B, Snyder H and Atkins H (1997). Comparative citation rankings of authors in monographic and journal literature: a study of sociology. *Journal of Documentation* 53(3):263–273.
- CSIRO (Commonwealth Scientific and Industrial Research Organisation) (2004). *Annual Report 2003–04*.
- DEST (Department of Education, Science and Training) (2004). *Evaluation of knowledge and innovation reforms consultation report*. DEST, Canberra.
- DEST (Department of Education, Science and Training) (2005). *Research quality framework: assessing the quality and impact of research in Australia—issues paper, March 2005*.
- DEST (Department of Education, Science and Training). Science and Innovation Budget Tables 2004–05.
- Donovan C (2005). The governance of social science and everyday epistemology. *Public Administration* 83(3):597–615.
- FRST (Foundation for Research, Science and Technology) (2004). *Outcome indicators 2003/04: guidelines for data collation by research providers*. [http://www.frst.govt.nz/Evaluation/downloads/provider/2004\\_Guidelines.doc](http://www.frst.govt.nz/Evaluation/downloads/provider/2004_Guidelines.doc)
- Glänzel W (1996). A bibliometric approach to social sciences: national research performance in 6 selected social science areas, 1990–1992. *Scientometrics* 35(3):291–307.
- Glänzel W and Schoepflin U (1999). A bibliometric study of reference literature in the sciences and social sciences. *Information Processing and Management* 35:31–44.
- HEFCE (Higher Education Funding Council for England) (2005). *RAE 2008: Guidance to Panels, January 2005*.
- Hicks D (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics* 44(2):193–215.
- Hicks D (2004). The four literatures of social science. In: Handbook of quantitative social science and technology research, Moed M, Glänzel W and Schmoch U (eds), Kluwer Academic Publishers, Dordrecht/Boston/London.
- Hicks D, Tomizawa H, Saitoh Y and Kobayashi S (2004). Bibliometric techniques in the evaluation of federally funded research in the United States. *Research Evaluation* 13(2):78–86.
- IFACCA (International Federation of Arts Councils and Culture Agencies) (2004). *Statistical indicators for arts policy: discussion paper, July 2004*.
- Katz JS (1999). *Bibliometric indicators and the social sciences*. Prepared for United Kingdom Economic and Social Research Council.
- Luwel M, Moed HF, Nederhof AJ, De Samblanx V, Verbrugghen K and van der Wurff LJ (1999). *Towards indicators of research performance in the social sciences and humanities*. CWTS, Leiden.
- Martin BR (1996). The use of multiple indicators in the assessment of basic research. *Scientometrics* 36:343–362.
- Monash University School of Music—Conservatorium (2004). *Proposal for recognition within Monash University of music performance, music composition and commercial compact disc publication as research equivalent activity, 10 August 2004*.
- Nederhof AJ and Noyons ECM (1992). International comparison of departments' research performance in the humanities. *Journal of the American Society for Information Science* 43(3):249–256.
- Nederhof AJ and Zwaan RA (1991). Quality judgements of journals as indicators of research performance in the humanities and the social and behavioural sciences. *Journal of the American Society for Information Science* 42(5):332–340.
- Nederhof AJ, Zwaan RA, De Bruin RE and Dekker PJ (1989). Assessing the usefulness of bibliometric indicators for the humanities and the social and behavioural sciences: a comparative study. *Scientometrics* 15(5–6):423–435.
- NHMRC (National Health and Medical Research Council) (2004). *The NHMRC's performance management framework 2003–2006*. <http://www.nhmrc.gov.au/aboutus/pmf.htm>

- REPP (Research Evaluation and Policy Project) (2005). *Quantitative indicators for research assessment—a literature review*. REPP Working Paper, Research School of Social Sciences, The Australian National University, Canberra. <http://repp.anu.edu.au/Literature%20Review3.pdf>
- Roberts G (2003). *Review of research assessment: report by Sir Gareth Roberts to the UK Funding Bodies, Higher Education Funding Council for England (HEFCE), Scottish Higher Education Funding Council (SHEFC), Higher Education Funding Council for Wales (HEFCW)*.
- Royal Academy of Engineering (2000). *Measuring excellence in engineering research*.
- Royle P and Over R (1994). The use of bibliometric indicators to measure the research productivity of Australian academics. *Australian Academic and Research Libraries* 25(2):77–88.
- Strand D (1998). *Research in the creative arts*. Department of Employment, Education Training and Youth Affairs, Canberra.
- THES (Times Higher Education Supplement) (2005). Journals ‘top ten’ sparks a rebellion. Phil Baty, 28 January 2005.
- Toutkoushian RK, Porter SR, Danielson C and Hollis PR (2003). Using publications counts to measure an institution’s research productivity. *Research in Higher Education* 44:121–148.
- van Raan AFJ (1998). Assessment of social sciences: the use of advanced bibliometric methods as a necessary complement of peer review. *Research Evaluation* 7(1):1–6.
- Verbeek A, Debackere K, Luwel M and Zimmermann E (2002). Measuring progress and evolution in science and technology—I: the multiple uses of bibliometric indicators. *International Journal of Management Reviews* 4(2):179–211.
- Von Tunzelmann N and Kraemer Mbula E (2003). *Changes in research assessment practices in other countries since 1999*. Final report to the Higher Education Funding Council for England, 28 February 2003.
- VSNU (Association of the Universities in the Netherlands), Nederlandse Organisatie voor Wetenschappelijk Onderzoek and Koninklijke Nederlandse Akademie van Wetenschappen (KNAW) (2003). *Standard Evaluation Protocol 2003–2009 for Public Research Organisations*.
- Webber P (1994). Creative arts as research. In: *Creative arts and research*, Cusack C (ed), Centre of Creative Arts, University of Sydney.
- Wissler R (2004). Research outputs—up for auction. In: *Innovation in Australian arts, media and design: fresh challenges for the tertiary sector*, Wissler R, Haseman B, Wallace S and Keane M (eds), Post Pressed, Flaxton.
- Wooding S, Hanney S, Buxton M and Grant J (2004). *The returns from arthritis research*, Volume 1: Approach, analysis and recommendations, RAND Europe. [http://www.rand.org/pubs/monographs/2004/RAND\\_MG251.pdf](http://www.rand.org/pubs/monographs/2004/RAND_MG251.pdf)

# Appendix B

---

## The value to the end-user of publicly funded humanities, arts and social science research

---

### B1 Executive summary

This report summarises information provided by 28 interviewees and 12 questionnaire respondents representing four major national archives, three government departments, one international charitable organisation, two university-located research centres, one national research institute, two professional consultancy groups and one newspaper.

The strong consensus derived from the interviews and questionnaire responses is that humanities, arts and social science (HASS) research is inseparable from, and vital to, the shaping, implementation, evaluation and review of public policy and to the strengthening of individual and communal capabilities of people participating at every level in programs designed to increase the quality and effectiveness of, for example:

- health and welfare interventions
- education and training
- social and cross-cultural interaction and communication
- public dissemination of valid and reliable information via radio and TV
- public service delivery
- data, information and knowledge management system and process development
- curatorial archiving
- exhibitions, displays and performances
- and many other socially or culturally significant projects.

Two kinds of value attach to publicly funded HASS research: ‘inherent value’ and ‘value in use’. These values are multiplied by the fact that there are at least six stages in the life-cycle of a piece of research at which fresh valuations of either or both kinds will be made by different ‘end-users’. The same values are multiplied again by the fact that the research accumulates in national and departmental archives and libraries which can and do serve researchers whose primary involvement may or may not be in short-term policy development. Accumulated holdings of publicly funded HASS research support longer term research leading to major cultural events, including exhibitions, displays and presentations, such as those for which the Australian War Memorial, the National Gallery of Australia, the National Museum of Australia and the National Library of Australia and nationally significant opera, ballet, orchestral and theatre companies are known. The holdings also serve the needs of academic, professional and non-professional researchers from all sectors of society.

Measures of this value will always be partial and imperfect, but may be made by capturing the inherent value and the value-in-use of HASS research as perceived by organisational research committees and panels, policy developers, policy implementers, evaluators and reviewers (each of whose roles is itself the

subject of HASS research), and by beneficiaries of visitor/audience/reader research relating to the quality of cultural events associated with museums, galleries, libraries, archives, and opera, ballet, orchestral and theatre companies across the nation and the dissemination by print and electronic media of valid, reliable information.

## **B2 Summary of main responses to interview and questionnaire questions**

### **Q1 How would you define/describe the kind of work you do?**

A summary of interviewees' and questionnaire respondents' positions and roles is to be found in Appendix B.

### **Q2 What do you understand 'research' to be?**

Interviewees and questionnaire respondents were clear about their views on research. When giving their answers, they understood that we were talking about HASS research. There were five categories of response:

1. Research means discovery for improvement:
  - 'Finding out how things work, how to make things work.'
  - 'Providing clearer, deeper understandings for people with problems to solve.'
2. Research means discovery, creativity and communication for improvement
  - 'Creating and contributing new knowledge'
  - 'Looking into things more carefully, more thoughtfully than usual. Analysing and interpreting what you see, using all available knowledge to test the reasonableness of your analysis and your interpretation. Publishing whatever passes the test so that it can be put to use.'
3. Research means discovery, creativity, communication for improvement and being methodical
  - 'It's either the action research model: you know—the plan, act, evaluate, reflect spiral—plan, act, evaluate, reflect, plan, act, evaluate ever onwards and upwards ... or it's the new factory model where you've got an input–throughput–output model with a self-regulating feedback loop or two and your job is to understand how it works, watch it, test it, know its strengths and weaknesses, anticipate the need to fix it or improve it or retire it, and know how to advise the people who'll have to wear it if the necessary changes aren't made in the best possible way at the best possible time.'
  - 'For us? Observation. Unobtrusive participation if you can manage it, negotiate it, whatever. And reflection. And interviews, focus group sessions, questionnaire surveys, data gathering, data analysis, finding out what hasn't been found out by the buggers who got there before you. Adding new perspectives anyway. And reporting on it, writing and presenting papers, getting peer reviews, publication. The lot. The whole routine.'
4. Research means discovery, creativity, communication for improvement and having a methodology
  - 'It means figuring out whether you're, say, an ethnographer, an ethnologist or a cultural anthropologist, or a phenomenographer, a phenomenologist or a hermeneutic semasiologist, or just somebody who does interviews and surveys or whatever, and acting accordingly. And writing it up. Always writing it up so it can be turned to advantage.'
  - 'Research means six things: methodology, ontology, epistemology, a field to work in, enough money to keep you alive while you get on with it, and an audience at the end—peers, colleagues, someone you can talk to. It also means having something like the NHMRC's levels of evidence and performance measurement framework to work with—the first five of their performance measures at any rate: creating new knowledge / enhancing people's capacity to innovate / putting knowledge to use / ensuring high ethical standards / improving communications and strengthening collaborations.'

5. Research means finding your own comfort zone on the spectrum that runs between ‘pure’ and ‘applied’, or ‘academic’ and ‘professional’ research

(See Figure B1: A common polarisation of approaches to research).

### **Q3 What kind of research, or research product, do you wish you had access to, but cannot get?**

Two generalised responses emerged in interviews and questionnaire responses. Various phrased, needs were noted for:

1. Longitudinal data to use when turning a policy proposal into a policy
  - ‘The kind [of research] that generates longitudinal data continuously. Research that gives us more than annual stats. Research that’s longitudinally funded and keeps on providing fresh, non-trivial, non-superficial updates and insights we can use to underpin our work in the world of policies, plans and possible futures.’
2. Research into sensitive and effective ways of operationalising policy
  - ‘Policies reach us cold. Some of them read like they were created in a research vacuum. And we’re expected to operationalise them as if we know what they’re all about. I’d like to know the research that shaped the policy and I’d like some research that would help us see how best to implement it—especially when it’s sensitive.’

Two specific needs, variously phrased, were voiced on behalf of research libraries, archives, galleries and museums:

1. more research, or more funds to conduct research, or the right to initiate ARC-fundable research, into ‘human information-seeking behaviours’ and the assistance that can be provided by digital and other technologies
2. greater understanding among funding agents
  - that increased use of digital technology does not necessarily result in greater economies of floor-space, staff salaries or overhead costs
  - that digitisation may augment the power of libraries, archives, galleries and museums to serve the needs of researchers and others, especially those who are remote and under-resourced, but that it does not necessarily replace or remove the need for the preservation of material objects, or for personal or manual services already in place.

One specific and highly prioritised need for more longitudinal research into issues relating to Indigenous health was voiced on behalf of workers in Indigenous affairs.

One specific and highly prioritised need for more longitudinal research into issues relating to citizenship, social justice and poverty alleviation was voiced on behalf of charity organisations.

Each voiced need focused on the need for ongoing research, with constant publication or communication of substantial, valid and reliable findings on which more effective programs of action could be based.

Discussion on this topic reinforced the perception that, while the interviewees and questionnaire respondents were more accustomed to dealing with short-term research projects aimed at meeting specific policy needs or solving specific problems, they had a felt need for the kinds of research products that were more likely to derive from longer-term programs of research.

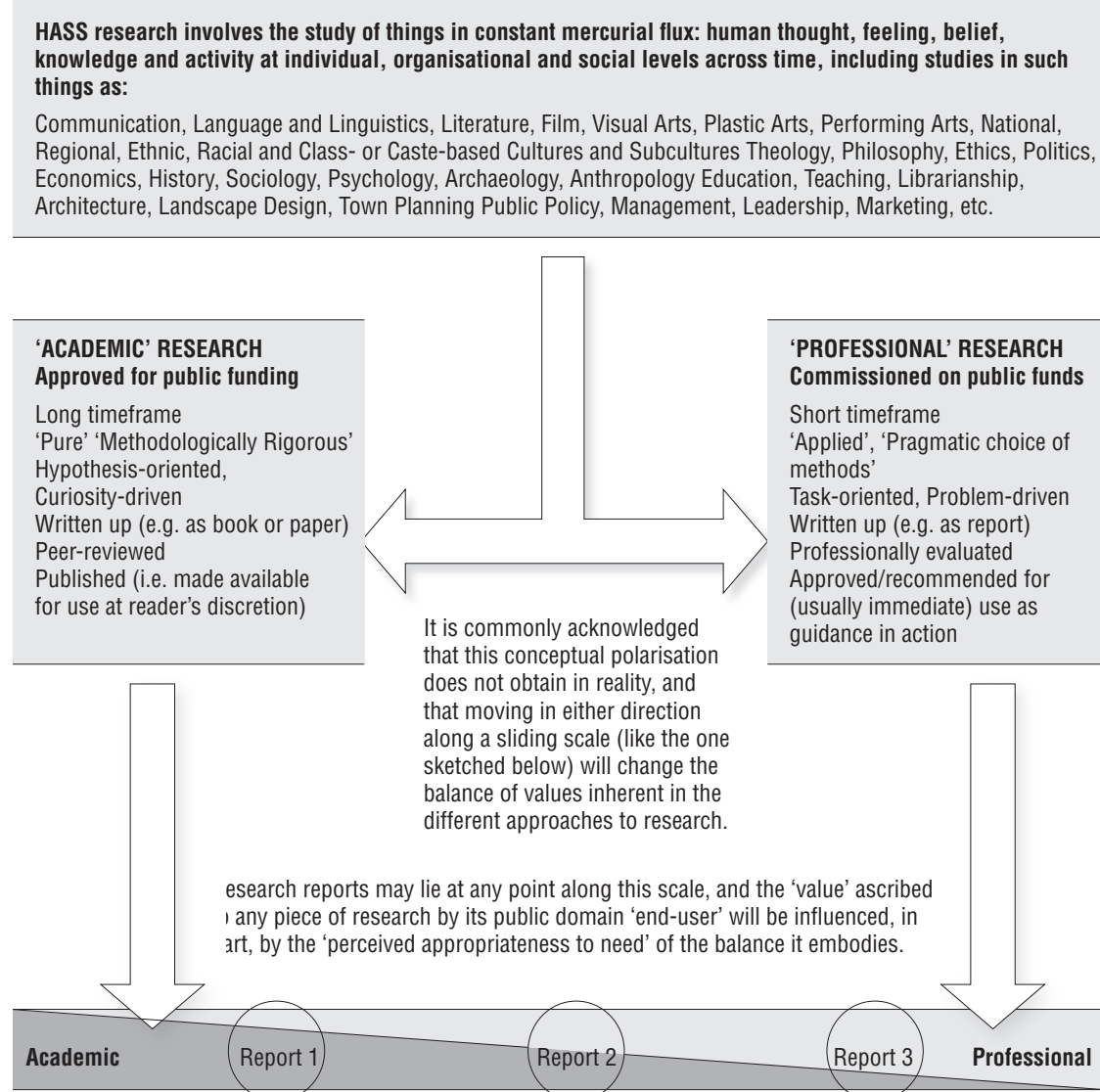
### Preferences for kinds of research are dictated by time pressures

Interviewees spoke of ‘professional’ and ‘academic’ research. They valued these kinds of research differently, with their preference for one or the other being closely related to their workplace time-constraints. Where lead-times and policy horizons are short (e.g. as when driven by ministerial imperatives), ‘professional’ research, commissioned on contract and with a precisely focused brief, is preferred. In these circumstances, higher value is placed on currency, speed of delivery, plain English write-up and immediate usability. Where lead-times or policy horizons are longer (as in the preparation for major exhibitions or displays in, say, the Australian War Memorial, the National Gallery of Australia or the National Library of Australia), higher value is placed on methodological rigour and the durability of research findings.

The phrases ‘immediate impact’ and ‘long-term effects’ were raised in some interviews. It was suggested in those interviews that the word ‘impact’ as used in the enquiry’s title (‘quality and impact of publicly funded HASS research’) might seem like an invitation to pragmatic end-users suffering tight deadlines to emphasise the value they have for ‘professional’ research and underemphasise the value they have for ‘academic’ research.

See Figure B1: A common polarisation of approaches to research.

**Figure B1 A common polarisation of approaches to research**





**Mercurial flux, the polarisation of research types, and where ‘value’ lies**

It was observed that HASS research focuses on things that are in mercurial flux. It was also observed that ‘academic’ HASS research typically takes longer than ‘professional’ HASS research to ‘capture and present usable truths about what’s continually changing’.

Where policy development is driven by political or social urgency, high value is therefore often placed on ‘professional’ research, the achieved understandings (or usable truths) of which are more quickly delivered. However, high value is *reserved* for research that achieves its findings over a longer time; where the research is continuous; where currency is achieved by regular updates on, and/or confirmations of, earlier findings; and where the most current updates/confirmations are readily and constantly available to those responsible for drafting policy proposals and those responsible for drafting, operationalising and revising policy.

These observations usually indicated a desire for a shift from the short-term to the long-term view of policy development, policy implementation and program maintenance.

Note: It was acknowledged that, while the conceptual polarisation of ‘academic’ and ‘professional’ research served a useful purpose when it helped make a point in casual conversation, the polarisation itself is never absolute in reality.

By contrast, however, the prevalence of the short-term view of policy was asserted as a reality. While the short-term view prevails, the value placed on any piece of research by its end-user will depend, in part at least, on the appropriateness of the compromise reached by the researcher between methodological rigour and speed of delivery.

**Q4 What kinds of research do you have access to, and use?**

Interviewees and questionnaire respondents collectively had access to all the kinds of research referred to in answers to Questions 2 and 3 and Figure B1.

**Q5 Where do you go for research expertise (e.g. do you do it yourself in-house, employ consultants, go to professional research institutes, go to universities and other academic sources)?**

*In-house:* It was claimed that much research deemed necessary by government departments is done in-house by their own staff. Nowadays, it was said, government departments frequently have staff who can tap professional research institutes and centres and national and international bibliographical databases for necessary information with all the panache of a practised consultant or professional researcher.

*Universities and consultants:* Universities are frequently named as the source of researchers, but there was general agreement that university-based researchers were sought out less because of their university affiliation than because of their affiliation with specific research centres or institutes. It was also observed that much of the work many university-based researchers were contracted to do was ‘professional’ rather than ‘academic’. University researchers typically function as consultants.

*The Australian Bureau of Statistics (ABS):* Statistical data relating to such things as the prevalence and incidence of phenomena being addressed in a policy, or potentially affected through policy implementation, were sought from the ABS. The bureau was identified as providing global, longitudinal and current, statistical data, and a data processing service that made it possible to generate a tailored statistical context within which to frame policy development.

*Institutional ‘meta-view’ holders:* The ABS provides a statistical version of the global overview, or meta-view. Equally comprehensive but less exclusively statistical meta-views are provided by such organisations as the

Australian Institute of Health and Welfare, the Australian Public Policy Research Network, the Australasian Cochrane Centre, the Cochrane Collaboration, the Australian Collaboration, the Australian Council of Social Services, and the National Health and Medical Research Council. These providers are frequently referred to as a first port of call when scoping a response to a policy proposal. They are spoken of as holding comprehensive and up-to-date understandings of their fields.

*Individual meta-view holders and research assistants:* Digests of academic research relating to public issues on which policies or policy-driven interventions are focused are sometimes derived from individual ‘meta-view holders’. These individuals are usually professorial or emeritus researchers in their own right, globally connected, discovered through assiduous networking, and often on contract as advisers to department leaders. Digests are also derived by research assistants from hard- and soft-copy archives and databases.

*Market researchers:* Professional market research relating to public needs, desires or values with respect to a proposed service, presentation, exhibition, display or event, or relating to publicity surrounding a proposed or currently offered service, presentation, exhibition, display or event—or a change to an existing service or product—is derived from numerous sources, including the Government Communications Unit, the Australian Market and Social Research Society, and the Australian Tourism Research Institute.

## Q6 Which people do research for you?

The list of people referred to in interviews or questionnaire responses, or named in departmental or organisational annual reports, is very long. The characteristics of the people chosen as dependable providers of the required research services are outlined in answer to question 9.

As an indication, in its *Research and Evaluation Digest 2001–02*, the Department of Family and Community Services (FaCS) lists the following sources of research assistance:

Source	Number of projects
Department of Family and Community Services	
FaCS, drawing on staff from its own sections and divisions	21
External consultancies	
RPR Consulting	3
KPMG	3
Orima Research Pty Ltd	2
Wallis Consulting Group Pty Ltd	2
ARTD Management and Research Consultants	2
Jenny Pearson & Associates Pty Ltd	2
Phoenix Projects Pty Ltd	1
Focal Point Consulting	1
Consulting Insights	1
Healthabit Pty Ltd	1
NFO Donovan Research	1
Leanne Craze, Craze Lateral Solutions	1
Helen Connor, Helen Connor and Associates	1
AC Nielsen	1
Roy Morgan Research	1

Source	Number of projects
Logos Pleiades Consulting	1
Spice Consulting	1
Simpson Norris International	1
Thomson Goodall and Associates Pty Ltd	1
University-based centres and institutes	
Melbourne Institute of Applied Economic and Social Research—University of Melbourne	15
Social Policy Research Centre—University of New South Wales	14
Social Policy Evaluation, Analysis and Research Centre—The Australian National University	7
Australian Housing and Urban Research Institute—The Australian National University	4
Australian Housing and Urban Research Institute—Flinders University, University of Adelaide, University of South Australia, University of Tasmania and Northern Territory (affiliate)	3
Australian Housing and Urban Research Institute—Murdoch University and Curtin University	2
Australian Housing and Urban Research Institute—Swinburne University of Technology and Monash University	2
Australian Housing and Urban Research Institute—Sydney University	2
Australian Housing and Urban Research Institute—University of Western Sydney and University of New South Wales	1
Australian Housing and Urban Research Institute—The University of South Australia	1
Australian Housing and Urban Research Institute—University of Queensland and Queensland University of Technology	1
Centre for Research in Education, Equity and Work—University of South Australia	1
Australian Centre for Industrial Relations Research and Training—Macquarie University and the University of Queensland	1
The Demography Program—The Australian National University	1
The Communication Centre—Queensland University of Technology	1
Government departments other than FaCS	
Department of Workplace Relations and Community Services	1
Other bodies	
Australian Institute of Family Studies	2
Australian Council for Educational Research	1
Centre for Community Child Health Melbourne	1
Western Australian Council of Social Services	1
TVW Telethon Institute for Child Health Research.	1
Australian Federation of Homelessness Organisations	1
Individuals	

Over 70 individuals, affiliated with the bodies named above, are identified in the FaCS Research and Evaluation Digest 2001–02 as contracted researchers (consultants). Many of them are named as having been involved in more than one project.

A further 24 individuals are also named as contracted researchers (consultants), either without affiliations, or loosely affiliated as follows: 1 from 'United States', 1 from Edith Cowan University, 2 from the University of Western Sydney, 3 from the Royal Melbourne Institute of Technology.

## Q7 How do you, as an end-user, value the research you have access to?

Research was not commonly perceived by interviewees or questionnaire respondents as a discrete process with a single definable ‘end-point’, a single identifiable ‘end-user’ or a single definable way of ‘valuing’ research.

It was observed that research was ‘used’ in many ways by many different people, and that any piece of research would inevitably have different values for different users at different times.

Distilled, these observations suggest that any piece of research might have:

- an ‘inherent value’ (measured in terms of its methodological thoroughness, its validity and reliability, and the currency, relevance and comprehensibility of its findings)

and/or

- a ‘value-in-use’ (measured in terms of how usable the user finds it to be).

These observations also suggest that there are six main points of impact in the life of a piece of research and that each impact point marks the end of a phase in that life. The people who use the research at that end-point are, in terms of their particular research-related function, ‘end-users’. There may be as many as six end-users and six points of view from which to allocate ‘value’ to any piece of usable research.

The six potential end-users of published or archived research and research-in-hand are:

- 1 the commissioning or funding body’s research committee or evaluating panel
- 2 the policy developers
- 3 the policy implementers and program evaluators
- 4 those beneficiaries of policy-driven programs for political, economic, social or cultural change who become aware, or are made aware, of the research underpinning the program in question
- 5 researchers stimulated to undertake further research
- 6 publishers, owners of research archives, future policy developers and subsequent researchers

The inherent value of a piece of research will be captured (or assumptions about it will be made) at the first point of impact. Information about this inherent value may or may not be passed on, or it may be captured afresh by people at any or all of the five remaining points.

After the first point of impact, value-in-use predominates.

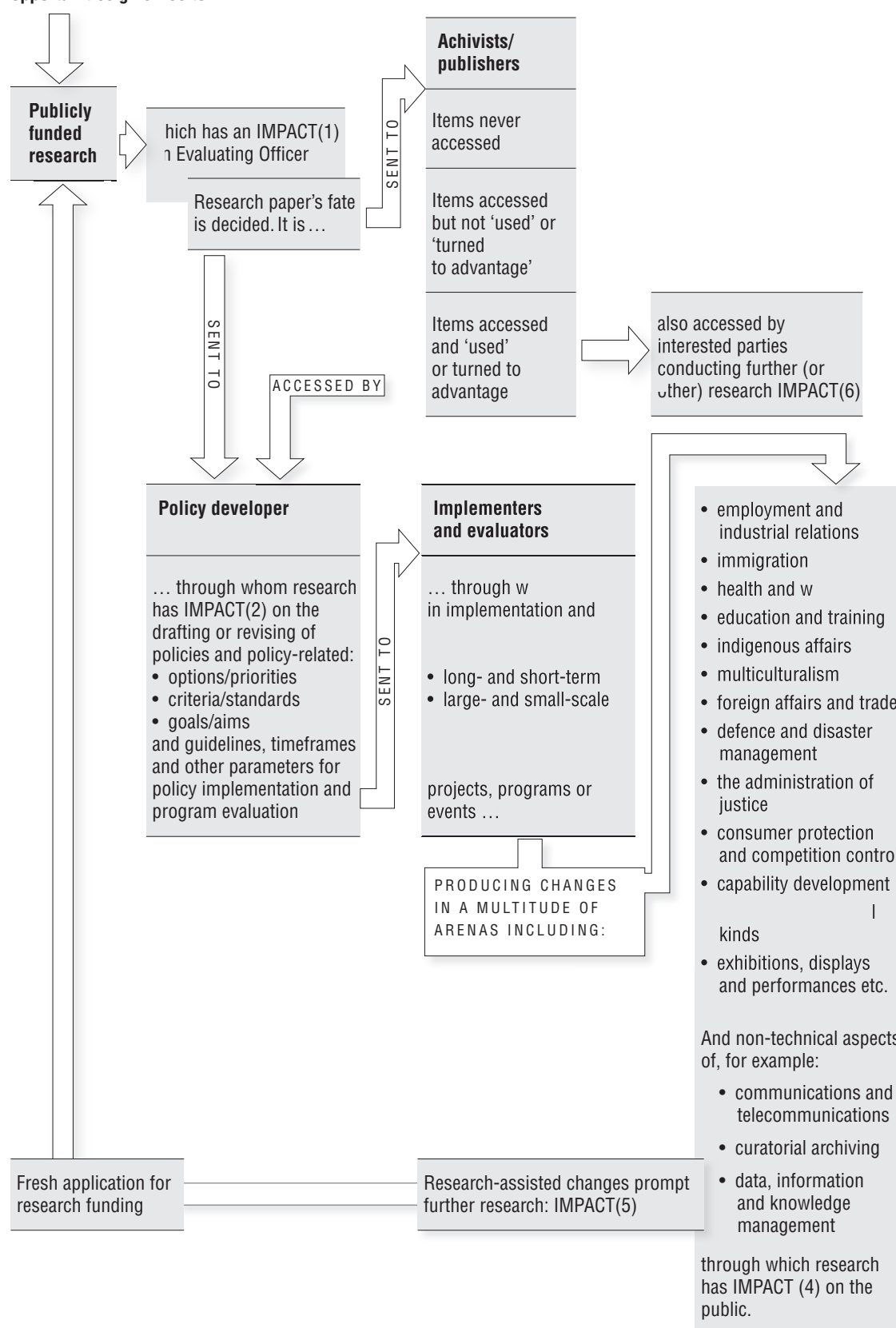
Value-in-use is relative. It is the value the research has:

- for the user
- at that point in time
- in the completion of their given task.

(See Figure B2: Impact points for publicly funded HASS research)

**Figure B2 Impact points for publicly funded HASS research**

White Papers budgetary issues, policy proposals, and manifest needs or opportunities give rise to ...



## Q8 How do you use research?

The research being discussed here is used to provide *guidance in action*, either *immediately*, or *subsequently* when it is accessed from library holdings or the internet.

Interviewees and questionnaire respondents generally see the guidance as being provided through

- the policies, procedural documents and manuals
- the lists of prioritised options, criteria, standards, aims and goals
- the guidelines, timeframes and other parameters for policy implementation and program evaluation

which the research itself helps people to develop.

In general terms, end-users value research primarily for its contribution to the development of the things listed above, and for the guidance which these research-influenced things give to the users in their work.

Work (or conscious reaction to the influence of the research) begins for each end-user at their respective point of impact.

*Value and use at Impact Point 1—Quality control:* Research is valued by the commissioning or funding body's research committee or evaluating authority *as evidence that terms of reference have been met and research funding has been satisfactorily acquitted*. The authority either disapproves of the research and returns it or dismisses it; or approves it and forwards it to policy developers and/or to hard- or soft-copy archivists and/or publishers.

*Value and use at Impact Point 2:* Research is used at this point in the development of policy. It is valued by the policy developer *for its comprehensibility, relevance, currency and usefulness in that task*. Note: The value ascribed to any piece of research at this point may relate to the speed and ease with which data, information, evidence or knowledge can be transferred directly from research report to policy document. The value ascribed to any given piece of research at this point may be very high, even if use is only made of one particle of the research output.

*Value and use at Impact Point 3:* Research is valued by the policy implementer and program evaluator *for the practical guidance it offers in implementation and evaluation*. The policy implementer and program evaluator place a high value on research the focus of which goes beyond the idea-content of policy to embrace the protocols of effective policy implementation and project/program evaluation. Where the research only focuses on the policy's idea-content, it has minimal value-in-use at this point, but it is still appropriate to make it available to the implementer. Where this does not happen, the implementer implements the policy document rather than the researched ideas behind the document.

*Value and use at Impact Point 4:* Research is valued by beneficiaries of policy-driven programs designed to achieve political, economic, social or cultural change *for its contribution to those changes*, but only where that contribution is known—i.e. where it is reliably attributable and openly attributed to the research. Where this does not happen, beneficiaries value the program or its implementers rather than the research.

*Value and use at Impact Point 5:* Fruitfully applied research is valued by the research community at large *for its power to stimulate and guide further research* or, to quote one questionnaire respondent, 'for its provocation to further scholarship'.

*Value and use at Impact Point 6:* Archived or published research is valued by publishers, owners of archives, policy developers and others *for its potential application in future policy development, review and implementation*; and by the research community *as material for further related research or for research in cognate but separate areas*.

*The additional value of individual and communal capability development:* A further and possibly immeasurable value is to be found in the progressive capability development in individuals and community groups that results from the funding of research that is intended to facilitate change.

Capacity to value research soundly with respect to its usefulness and to use it effectively in context is also affected by

- permanency or long-term stability of appointed staff
- having ‘maintaining current knowledge of research in the area’ as a term of employment.

Capturing the value of publicly funded HASS research to the end-user means identifying the ways in which each of the six end-users reacts to the research—especially in the work the research enables them to do.

## **Q9 What training/qualifications/experience/characteristics do you look for in your chosen/preferred researchers?**

Much of the work done by researchers on contract requires the researcher to be able to conduct interviews, focus group sessions, questionnaire surveys, case studies, literature searches, reviews of existing data or information, model creation, implementation and evaluation, assessments, evaluations, database analysis, data mining, factor analysis, multivariate analysis, iterative (plan–act–evaluate–reflect) action research projects, market testing, and community work involving observation and unobtrusive participation and focusing on human behaviour. The researcher is also expected to have a track record in completing tasks and presenting comprehensive, plain-English reports on time.

Chosen/preferred researchers need to have an understanding of the context in which their contracted work is to be done, of the issues being addressed, and of the possibility of achieving improvements through the work. They are expected to bring knowledge of similar work done earlier or currently being done elsewhere in the world, and to contribute and apply this knowledge in their work.

Ideally, the research *team leader*, or the project’s primary adviser, is an individual ‘meta-view’ holder in their own right. They will be a leader in their sector, a known opinion maker and decision maker, a professor or associate professor in academic service, a long-acknowledged expert running a component in an industrial complex, a clinical director in a hospital, a chief scientist working in industry in their primary discipline, or a senior NGO board member. Their status is genuinely matched by experience, understanding and expertise. Typically, this person understands research and its implications, can articulate this understanding and can give sound advice.

## **Q10 How do you find such people?**

Networking, word of mouth, recommendations made by the Minister, senior public servants, upper echelons in the hierarchical professional and academic networks.

## **Q11 What research questions do you typically ask?**

Interviewees and questionnaire respondents indicate that, while there will be a generic flavour to the questions being asked by any given government department or other purchaser of research, the actual questions asked are usually very diverse in detail.

However, there appears to be a typical pattern of thinking before the engagement of a researcher:

We need to draft a policy, create an implementation plan and an appropriate process of project/program evaluation. To do this, we need to know:

- what we already know experientially and from prior study
- what *more* than that our own library/archive can tell us
- what *more* than that we can get from our ‘meta-view holder’ / contracted adviser
- how *much* of that is usable in present circumstances
- what *more* than that amount we now need to find out
- how *much* of this extra information we can find out using our own in-house resources
- what *more* than that we still need to find out.



This final step yields the decision to employ or not employ a researcher, and gives rise to the prospective researcher's brief.

To illustrate the coexistence of generic flavour and specific diversity in a research purchaser's questions, the key priority research questions being posed by FaCS in its 2003–05 research program are as follows:

- What is social participation and what does it achieve?
- What are the dynamics of long-term reliance on income support, particularly among specific groups such as sole parents and people with disabilities?
- What are the financial circumstances of income support recipients and how does the level of wealth, investment and debt influence their capacity for long-term self-reliance? What interventions would increase their capacity for self-reliance?
- How do personal attitudes and behaviours impact on the level of wealth and debt among income recipients, and how do wealth and debt patterns affect people's choices and transitions?
- What are the factors influencing family choice with regard to child care and what are the different outcomes from these choices, particularly for groups such as people from culturally and linguistically diverse backgrounds, Indigenous clients, children with a disability, and children at risk?
- What are the drivers and disincentives to foster parenting?
- What are the factors that hinder or support young people to become self-reliant adults?
- Can we identify successful pathways for Indigenous children and youth?
- What changes are needed to existing FaCS programs to deal with the specific needs of Indigenous people?
- What are the individual and systemic incentives and barriers to people working for longer?
- What are the implications of declining fertility and structural ageing for key FaCS customer groups?
- What are the pathways to and from homelessness?
- What are the outcomes, including non-shelter outcomes (for example, health or education) for individuals making various housing choices?
- What service models work for remote communities
- What are the patterns of money flows in, within, and out of remote communities, including Indigenous communities? What critical factors affect the development of local market economies in remote communities, and how can existing assets and resources be better leveraged for economic development?
- What are the drivers of movement to and from different locations?
- In what ways is mobility related to the regional aspects of unemployment, long-term unemployment and the adequacy of income support and service provision?

## **Q12 What discipline areas do you draw research from?**

Representatives of government departments and instrumentalities, by definition, have specific fields of interest which map loosely across into academic discipline areas, but interviewees and questionnaire respondents indicated that they were neither constrained nor guided by this. When seeking research assistance, they thought less in terms of academic disciplines than in terms of the known research strengths of specific research centres or individuals.

Professional researchers are sought less for their subject-content expertise than for their expertise in the areas itemised in answer to Question 9.

### Q13 What difficulties lie in the way of securing funds for the research you need?

Some interviewees spoke of difficulties in securing funds for research projects which were difficult to align with the National Research Priorities and did not have political patronage.

No other difficulties in securing funds were spoken of, other than those that go with the competitive nature of bidding for funds from internal budgets and securing ARC (and other) research grants.

Statutory bodies unable to initiate (take a lead in) ARC grant applications, but who could only benefit from such grants by participating in joint projects with other organisations, experienced difficulties in securing funds for projects that would precisely target or completely meet their specific needs.

The difficulty most commonly spoken of was difficulty in securing *sufficient* funds.

### Q14 What difficulties lie in the way of getting high-quality (relevant, comprehensible, usable) research—and of having it delivered on time?

Interviewees seldom thought that quality was a problem.

Most interviewees work in organisations or departments that have a quality-control filter in the form of a research committee, or a formal research evaluation panel of some kind, which fulfils its quality-control function in the vetting of tenderers for research projects, in the selection of researchers or research organisations/institutes, and in making judgments on the acceptability or unacceptability of submitted research.

Typically, selection of researchers was said to be influenced by the proven ability of the researcher or research organisation to stick to their brief and meet contractual obligations in regard to deliverables and timelines.

## B3 Multiple end-users, multiple values, and individual and community capability development: a case in point

A joint research project into a sensitive topic, such as the policing of drug and alcohol use in Indigenous communities, is formulated. The project partners are the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) and the Australian Institute of Criminology (AIC). Funds are provided by the National Drug Law Enforcement Research Fund (NDLERF). A researcher is appointed through a national competitive tender process. The research involves participative involvement (protocol-compliant information gathering and internal reviewing of existing practices) from state police forces, health organisations, and Indigenous communities across the nation. The research is not done in isolation from other research. It is discussed and its inherent values and potential values-in-use are addressed while in progress. It is

- completed
- approved
- archived and/or published—which means its potential to inspire, guide or fuel further research is preserved
- applied—which means its findings give rise to numerous projects involving (among other things) reactive and proactive interventions and many interrelated forms of monitoring and evaluation.

All parties identified (including AIATSIS, AIC and NDLERF) are in a position to ‘use’ the research or ‘turn it to advantage’. Each party will attribute an end-user’s value to the research. All parties (including providers of research infrastructure) who are beneficially extended, or rendered more self-aware in the context addressed by the research and/or more capable of contributing to the changes facilitated by the research, may attribute their capability development, in part at least, to their involvement in the research or the activities the research gave rise to. There is agreement among interviewees and questionnaire respondents that the value of the research to *all* end-users warrants attention by funding agents.

## B4 Conclusion

The individuals and organisations drawn on for information collectively represent research; research assessment and evaluation; hard- and soft-copy archiving and publication; policy development; policy implementation; and evaluation of policy-derived programs and projects. As there were no points of disagreement among the informants, it has been possible to summarise and present their views as if they came from a homogeneous source.

There was universal agreement that publicly funded HASS research, or access to archives of previously completed publicly funded HASS research, was essential to their own work or the work of their organisation.

It has not been possible to represent fully all the information provided by informants. Further study is warranted, especially into the way networking strategies and the use of special interest groups repeatedly illustrate the accepted wisdom that ideas (no matter how potentially beneficial they may be) do not travel very far without being carried by people. This means that the full value to end-users of publicly funded HASS research cannot be captured unless it includes a study of how the ideas, perspectives, perceptions, and pieces of confirmed, valid and reliable knowledge emerging from that research are carried by people, by word of mouth, and by actions that go beyond policy-prescribed and policy-compliant human and social activity.

In the eyes of interviewees and questionnaire respondents, research lies between polarised values, variously identified as:

Pure	Applied
Academic	Professional
Methodologically rigorous	Pragmatic in selection of methods
Hypothesis testing	Problem solving
Curiosity driven	Brief driven
Broader scope and/or greater depth	Narrower focus and/or limited depth
High potential for anticipated and unanticipated results of high value in the longer term	High potential for specifically intended results of high value in the shorter term
'Inherent value' measured by peer assessment	'Value in use' measured by six end-users

Where research is intended to inform policy, and policy is politically driven, the short life of political administrations, short departmental budgetary timeframes, and the drive for return on investment conspire in favour of public funding for applied research over 'pure' research.

Applied research is also favoured even when policy appears to be driven by social necessity, as the prioritising of social necessities is politicised and the timeframes within which the needs are to be met are influenced by the same factors of short political and budgetary cycles and the required return on investment.

The purchaser of research is not seen as the only or the most important end-user. Neither is the beneficiary of research-informed programs for social improvement.

It is thought that inherent value and value-in-use have (or should have) equal importance.

Capturing the value that publicly funded HASS research has for the end-user means identifying the ways in which each of the six end-users reacts to the research:

- 1 in the quality-control process enacted on receipt of the research report
- 2 in the application of the research in policy development

- 3 in the guidance it provides in implementation (or ‘operationalisation’) and in the evaluation of whatever projects or programs spring from policy
- 4 in the reactions of policy-aware, research-aware beneficiaries of projects and programs
- 5 in the research-informed project/program’s provocation of further scholarship
- 6 in the recourse researchers and others have to the initial research through material and electronic archives and publications.

The value that end-users place on *access to research records* is of particular importance. Interviewees and questionnaire respondents emphasise the belief that the existence of archives has a significant but probably immeasurable ‘inherent value’ in its own right, and that the development of archives is, perhaps, the most important outcome of publicly funded HASS research. It is thought that there is value in funding agents prioritising the capture of the value-in-use of material and electronic archives—possibly by variant forms of long-term, continuous archive-visitor research.

Interviewees and questionnaire respondents believed that the ‘capability development’ effect of publicly funded HASS research warrants the full attention of research funding agents.

## B5 Interview format

This enquiry into the value to the end-user of publicly funded humanities, arts and social sciences (HASS) research was a sub-project of a broader enquiry into the quality and impact of publicly funded HASS research.

The sub-project was to proceed, using focus groups, interviews, and possibly a questionnaire/survey with groups and individuals drawn from research-using corporate bodies and government departments and institutions located in Canberra.

A list of questions was proposed as starting points for interviews and focus group sessions and for inclusion in the possible questionnaire. In the event, this list was augmented by the inclusion of questions which respondents wanted to answer and by the rephrasing of some questions. Answers to the rephrased questions are folded into the summary of main responses to interviews and questionnaires. The augmented list was:

*End-users:*

- 1 How would you define/describe the kind of work you do?
- 2 What do you understand ‘research’ to be?
- 3 What kinds of research do you wish you had access to, but cannot get?
- 4 What kinds of research do you have access to, and use?
- 5 Where do you go for research expertise (e.g. do you do it yourself, employ consultants, go to professional research institutes, go to universities and other academic sources)?
- 6 Which people do research for you?
- 7 What training/qualifications/experience/characteristics do you look for in your chosen/preferred researchers?
- 8 How do you find them?
- 9 What research questions do you typically ask?
- 10 What discipline areas do you draw research from?
- 11 How do you, as end-user, value the research you have access to?

- 12 How do you use it?
- 13 What difficulties lie in the way of securing funds for the research you need?
- 15 What difficulties lie in the way of getting high-quality (relevant, comprehensible, usable) research—and of having it delivered on time?

*Researchers:*

- 1 Who funds your research?
- 2 Who uses your research?
- 3 How do they use it?
- 4 How do they value it?
- 5 How do you know how they value it?

The sub-project was to generate a brief report summarising main responses from the community consulted. The final version of this report was to be delivered by 12 June 2005.

Interviews and focus group sessions were conducted with individuals and small groups from:

The National Library of Australia  
 The National Museum of Australia  
 The National Gallery of Australia  
 The Australian War Memorial  
 The Department of Health and Ageing  
 The Department of Family and Community Services  
 The Department of Agriculture, Fisheries and Forestry  
 The Centre for Aboriginal Economic Policy Research  
 The Centre for Applied Philosophy and Public Ethics  
 The Faculty of Economics, ANU  
 The National Graduate School of Management  
 The Australian Institute of Aboriginal and Torres Strait Islander Studies  
 OXFAM, Community Aid Abroad and The World Neighbourhood Education Program  
*The Canberra Times*  
 The PACE Consultancy Group  
 The Australian Institute of Management.

Wherever possible, interviews were conducted with two or more informants from each organisation. Second interviews were separated in time and place from first interviews. Second interviews sought supplementary and/or complementary information and/or alternative views.

After the augmented set of questions was established, it was sent as an email attachment or by post as a hard-copy questionnaire to members of the PACE Consultancy Group, to whom anonymity was guaranteed. They were asked to write responses in their own words. Twelve usable responses were received. These responses are folded into the summary of main responses to questions and questionnaires above.

## B6 Informants

Name	Position	Organisation
<b>Archives</b>		
Ms Linda Ferguson	Evaluation and Visitor Research Manager	Australian War Memorial
Ms Helen Hyland	Research Librarian	National Gallery of Australia
Ms Robyn Holmes	Director, Music Archive	National Library of Australia
Mr Warwick Cathro	Assistant Director-General, Innovation	National Library of Australia
Ms Susan Tonkin	Manager, Evaluation and Visitor Research	National Museum of Australia
<b>Government departments</b>		
Ms Elizabeth Williams	Director, Research Library	Department of Agriculture, Fisheries and Forestry
Mr Michael O'Hara	Ex-Director, Quality and Evaluation Section	Department of Health and Ageing
Mr David Webster	First Assistant Secretary, Portfolio Strategies Division	Department of Health and Ageing
Mr Andrew Whitecross	Assistant Secretary, Research and Data Management	Department of Family and Community Services
Ms Heather Evert	Manager, Research Management Section	Department of Family and Community Services
Ms Denise Ryan	Acting Manager, Research Support and Publications	Department of Family and Community Services
Ms Helen Moyle	Research Projects	Department of Family and Community Services
<b>Charitable organisations</b>		
Dr Emma Rooksby	ACT Committee Member	Oxfam and Community Aid Abroad
Ms Mandy Nearhos	National Board Member	Oxfam, Community Aid Abroad <i>and</i> World Neighbourhood Education Program
Ms Liz Wilson	Education Officer	Oxfam, Community Aid Abroad <i>and</i> World Neighbourhood Education Program
<b>Academic groups</b>		
Professor Russell Craig	Director of Education	National Graduate School of Management, ANU
Professor Keith Houghton	Dean of the Faculty of Economics and Director of the National Graduate School of Management	Faculty of Economics and National Graduate School of Management, ANU
Professor Steve Dowrick	ARC Senior Fellow	Faculty of Economics, ANU
<b>Research groups</b>		
Dr Peter Veth	Deputy Principal (Research)	Australian Institute for Aboriginal and Torres Strait Islander Studies
Mr Brendan Delahunty	Project Officer	Australian Institute for Aboriginal and Torres Strait Islander Studies
Dr John Taylor	Senior Fellow	Centre for Aboriginal Economic Policy Research
Ms Marie Davern	Executive Officer	Centre for Aboriginal Economic Policy Research
Dr Emma Rooksby	Research Fellow	Centre for Applied Philosophy and Public Ethics
Professor John Weckert	Program Manager, IT and Nanotechnology	Centre for Applied Philosophy and Public Ethics

Name	Position	Organisation
<b>Professional consultants</b>		
Mr Ian Carmody	ACT Regional Executive, Australian Institute of Management	Australian Institute of Management
Mr Graham Durant-Law	Director, Durant-Law Consulting Pty Ltd	Co-founder and Member of the PACE Consultancy Group
Dr Lyndal Thorburn	Managing Director and Principal, Innovation Dynamics Pty Ltd	PACE Consultancy Group Member
<b>Media</b>		
Mr Cameron Ross	Editor, the Opinions Page	Canberra Times



# Appendix C

## The case-study trial

During May and the first weeks of June 2005, seven academic departments and some twenty assessors participated in a trial of an assessment process closely akin to the new system recommended in this report.

### C1 Organisation

The principal purpose of the exercise was to determine the usefulness and validity of a wide range of measures of research quality, impact and capability. The seven units were drawn from across the various types of university, and across the range of disciplines within the HASS sector:

- Faculty of Law, ANU
- History, University of Queensland
- Conservatorium of Music, University of Newcastle
- School of Architecture and Design, RMIT
- Department of Politics and Public Policy, Griffith
- Department of Marketing, Griffith
- Trans/Forming Cultures research centre, UTS.

Each unit prepared a document outlining its case for excellence in the three areas of quality, impact and capability. The case was made by presenting concrete evidence in the form of a range of quantitative and qualitative indicators (see below). In preparing its case, each unit had a certain degree of flexibility. It was able to specify the weighting given to each of the three categories within defined limits:

- Quality: 40–70%
- Impact: 20–50%
- Capability: 10–30%.

Once the minimum levels of 40% (quality), 20% (impact) and 10% (capability) were met, each unit was able to allocate the remaining 30% at its discretion. This approach followed the suggestion in the 2005 RQF Issues Paper to allow flexibility in the assessment framework in order to accommodate the different natures of specific disciplines and the different strategic goals of various research units.

The document prepared by each unit was referred to a panel of experts to assess the research effort of the discipline as a whole. The panels comprised experts from academia, research end-users and research managers. The academics, ideally both Australian and international, were judged to be in the best position to make assessments of the quality and academic standing of the work within the discipline. Experts from industry, government, the professions or other end-user groups were best placed to assess the impact of the unit; and experienced research managers, for example deputy vice-chancellors or pro vice-chancellors (research) were best placed to assess research capability.

The criteria the panels used are specified below. Part of the case study process was to develop discipline-specific criteria and indicators where necessary.

The assessment panels did not physically meet. Instead, they conferred electronically via a secure website. This contained a repository of the research materials to be assessed, the case made by each unit, background documentation, and a web forum for the assessment panel to discuss both the research and the process.

The assessment panel rated the units (on a five-point scale) in each of the three categories of quality, impact and capability. These scores were aggregated to form an overall research rating, and then entered with an accompanying commentary on a custom-built (and password-protected) website: <http://sts-dev.anu.edu.au/moodle/chass>.

At the conclusion of the exercise, the assessment panels and the research units discussed the outcomes, both separately and together. The focus of these discussions was:

- the overall validity of the approach
- the usefulness of each individual measure
- discipline-specific issues that arose in the course of the exercise.

The timeline for the process was as set out below. Despite the constricted time and the failure of the ANU website over a critical weekend, the process generated useful discussion on the process and the measures.

21 March	Identification and invitation of research units participating in the exercise
4 April	Selection and invitation of assessors
11 April	Briefing of the research units; research units begin to collect the evidence used as the basis for their case for excellence
18 April	Briefing of the assessment panels; assessment web tools operational
25 April	Research units provide the case for excellence, with supporting data and evidence, to the panels for assessment via the website; assessors begin to evaluate the evidence
9 May	Final assessments collected and made available to research units
16 May	Feedback on the process gathered from assessors and research units.

## C2 Outcomes

### Evidence presented

Extensive work was done by the participants in assembling each case for excellence. These cases provide very useful examples of how such a document might look in a genuine RAE, as well as a good range of examples of the disciplinary differences (as well as similarities) that need to be accommodated in any assessment of research in the HASS sector.

Table C1 summarises the evidence presented. One of the issues that emerged from the trial—and one of the lessons learned—was that it was not always possible or desirable to distinguish or prescribe which indicators applied primarily to ‘quality’, which to ‘impact’ and (to a lesser extent) which to ‘capability’. This was often context-dependent. In general, assessors found it relatively easy to discriminate and to synthesise their judgments in each of the three categories on the basis of the whole body of evidence presented. Table C1 broadly gives the categories ‘Q’, ‘I’ and ‘C’ in order of priority. For example, the entry ‘IQ’ suggests that a particular indicator might be useful mainly in judging impact, but might also be useful for judging quality. The entries are suggestions and should not be taken as prescriptions.

**Table C1:** Summary of evidence presented

	Architecture and design	Multidisciplinary cultural studies	History	Politics and public policy	Marketing	Law	Music
<b>Overview data</b>							
Narrative describing the research and research-training activities of the group, research management, plans and organisation	QIC	QIC	QIC	QIC	QIC	QIC	QIC
Quantitative data: numbers of staff, students, research projects, and details of research infrastructure and equipment	QIC	QIC		QIC		QIC	
Individual researchers' statement of research impact				I		I	I
Financial management data		C					
<b>Research production and dissemination</b>							
Traditional (e.g. current DEST-eligible) academic publications, and academic conference papers	QI	QI	QI	QI	QI	QI	QI
Professional publications (usually not currently DEST-eligible), and professional conference papers	IQ					QI	QI
Non-written research outcomes (e.g. creative works, designs, documentaries, multimedia websites)	QI	QI					QI
Policy and other reports				IQ		IQ	
Self-assessment of publications and scholarly/artistic outputs (e.g. 'best five' publications by researcher, journal rank, etc.)				QI	Q	Q	
Media appearances, articles and reviews	I	I	I			I	QI
Public awareness programs		IQ					
<b>Bibliometric data</b>							
Citation indices				Q		I	
<b>Research funding</b>							
ARC grants awarded (Discovery grants, fellowships)	QI	QI	Q	Q	Q	Q	
ARC Linkage grants awarded	IQ	IQ		QI		QI	
Other government (local, state, federal non-ARC) funding for research	IQ	IQ	QI	QI		QI	
Industry research funding	I			QI		IQ	I
Competitive research funding internal to the institution	QC	QC	Q		Q		QI
Industry funding for education-linked research	CI	CI					
Competitive funding internal to the institution for education-linked research	C						
Data on successful <i>and</i> unsuccessful grant applications		QC					
Assessors' reports on grant applications				Q			Q
<b>Awards, prizes, fellowships and markers of esteem</b>							
Fellowships of academies and learned societies			Q	Q	Q	Q	Q
International lectures, keynote addresses, invitations to address government and non-government organisations	QI	QI	Q	QI		QI	QI
Academic prizes and awards		QI	Q	Q	Q	Q	
Industry/professional prizes and awards	IQ			Q	QI	QI	
Editorships and memberships of editorial boards	QI	QI	QI	Q		Q	Q
Convenorships of academic conferences	QI	QI	QI		QI		QI
Convenorships of professional conferences, exhibitions and/or artistic festivals.	IQ						
Offices held in academic societies and organisations			Q	Q		Q	
Offices held in professional or industry organisations or government consultative bodies	IQ	IQ	IQ	Q		IQ	
Academic standing of scholars-in-residence and adjunct researchers	QIC	QIC	QIC	QI		Q	
Professional standing of advanced students and scholars-in-residence from industry	CIQ						
<b>Partnerships</b>							

	Architecture and design	Multidisciplinary cultural studies	History	Politics and public policy	Marketing	Law	Music
Collaborations with other academic institutions and networks (national and international)	QI	QI		IC		QI	QI
Collaborations with government	IQ	IQ	IQ	I		IQ	
Collaborations with industry	I		IQ	I	I	IQ	QI
Collaborations with community organisations	I	I	I	I	I	IQ	
<b>Research training, education, and early career researcher development</b>							
New researcher recruitment and advancement strategies		CQ	CQ				
Research degree enrolments and completions	C	C		C	C	C	C
Destinations of research students and postdoctoral fellows			CQ	CQ			
Research-based education initiatives	C			CQ			CQ
Research-student publications	CQ	CQ			CQ		CQ
Prizes and grants for teaching and learning in research-based fields of education	CQ						
<b>Equipment and infrastructure</b>							
Self-assessment of research equipment and infrastructure	C		CQ	C		C	CQ

## Assessor comments

The following is a digest and summary of the major comments on the trial advanced by the assessors. These are grouped into broad categories.

### Format of the case for excellence

‘The unit’s own overview, including its caveats about assessment evidence, was critical. Those without such an overview, or with a limited one, were much harder to assess, particularly by a non-expert. This also suggests you should be pretty flexible about how the detailed evidence is provided. But such evidence is important.’

‘There were huge differences in the way information was provided in each of the submissions I read ranging from lists of pub[lications], research grants, etc to the totally descriptive. This points to the need for proformas like the NZ PBRF otherwise it is very difficult for the assessors.’

‘Among other things, the exercise reveals that there is a danger of what is supposed to be an assessment of a department’s research quality being distorted into an assessment of a department’s ability to make a good case for itself. They are not the same thing.’

‘But this judgment is all based on my prior knowledge or preconception—or prejudice in its literal sense. The information in the portfolio would not readily lead to this sort of conclusion—and I’m fairly confident that it is the portfolio presentation that is the major problem, not my broad preconception. If I’m right, then this matters a lot for any research assessment exercise.’

I’m sure that it was difficult for the Department to work out how to pitch its covering statement and the attached addenda. As presented, the documents seem to try to do a number of different things and end up doing none of them as well as the Department deserves. The statements fall somewhere between (a) trying to present neutral information that speaks for itself at least re quantity and possibly re quality; (b) trying to present some hints about why the assessors might be impressed, i.e. something of a case for a good assessment; and (c) trying to present some advice to assessors about what a generic case for high quality might look like. It ends up doing none of these particularly systematically or clearly.’

**Narrative overview vs. ‘hard data’**

‘The essentially descriptive cases were much more difficult to assess and I kept having to ask myself how easy it would be for any unit to describe one of their key projects in similar terms—too easy is the answer.’

‘I would have liked more concrete data about size of groups I was assessing and membership. Submissions seemed to concentrate on top researchers but what about the rest.’

‘Some of the information presented in summary form is very useful. Reporting DEST points per capita is a reasonable and useful statistic, and ought to be benchmarkable. I suspect that 2.67 is pretty good—and I would like benchmarks to test this hunch. A distributive metric would help—e.g. standard deviation or better still some graphical representation of all staff from lowest to highest score. (I appreciate that the shortcomings of the DEST points system are one of the reasons for this RQF coming in—but equally the DEST points system is far from meaningless, and it is how Departments have been asked to measure themselves for some years now).

‘This suggests the utility of considering other potentially benchmarkable statistics—e.g. competitive external research grants (number, dollars) per capita.’

**Panel composition and format**

‘On the issue of the expertise of the assessor, I think you need generalists and experts. But I also feel some face to face discussion (or at least by phone) is essential. I know there is some danger of group-think that way, but my experience is that it helps to clarify misunderstandings and ensures more consistency. Also, where there is a hint of conflict of interest, the expert having declared it may be able still to contribute in some useful way, the others being fully aware of the need to distinguish between that interest and the expertise they are looking for.’

‘While I have made an assessment based on the written material provided, I am conscious that a subject matter expert would be able to calibrate some of the evidence. I suspect my ratings would mostly be within close range, not knowing the standing of particular journals etc could affect weightings. Most importantly, I had to set aside my personal preference for other fields of work when considering ‘impact’, and focus on the evidence provided of external interest through commissioned work, collaborations etc.’

**Outputs and processes**

‘I am reminded of some work a few years ago on evaluating policy advice. That cautioned against emphasising too much the extent to which advice is accepted and, when it is, what happened. These things are relevant (particularly the latter obviously), but it is also important to assess the inputs (e.g. the information and the expertise) and the process (e.g. consultation, timeliness, accessibility, assessment of risks and implementation issues). In a similar way, I think impact and capability in particular need to include evidence of inputs and processes to complement the (probably limited) evidence of final societal impact or effectiveness.’

**Criteria**

‘My main reservation about exercises of this kind, based on very limited experience of the receiving end of the UK RAE and some limited experience of evaluating projects in the UK, Austria & Norway, is that the more complicated and formalistic the criteria the harder it is to distinguish between the quality of the case made and the quality of what’s being assessed. In other words, rather like evaluating exam answers with varying quality of handwriting.’

‘I don’t like the International/National/less-than-National wording at all, and the explanation was inconsistent. The first sentence is fine: “the intrinsic academic or artistic excellence”. Why is there any need to do have anything other than an International Best Practice standard? Why is there an assumption that “the best work in the field in Australia” is necessarily lowering the bar vis-à-vis “the best work in the field [internationally]”? Then the second paragraph suddenly switches to “international profile”—but rather than saying flatly that this criterion is NOT a test of “profile” (i.e. it is “intrinsic excellence”, and for profile

hold your fire until you get to the Impact criterion) it confuses things by presenting an unnecessary and irrelevant excuse for Australians perhaps not having an international profile. There is a danger here of an assessor equating Quality with Impact, when these are supposed to be separate criteria. It should be obvious that something unequivocally local—e.g. a splendid history of North Queensland—can be absolutely of the highest international standards within its genre of local/regional history, yet have pretty minimal “international significance”. It is the intrinsic Quality that ought to matter under this criterion.

‘So why not just use one measure—something like “International Best Practice”?’

### **Researchers put forward for assessment**

‘This group ... is indeed one of the two largest groups with this mix of programs in the country. The documents presented, in the very way in which they focus on a small number of cases, actually undermine for me the impact of the group. This may be because I suspect that most ... schools in the country would have one or two world class people, and two or three high profile and innovative projects they could point to. Yet most schools of [this discipline in] the country are less than one third of the size of [this one]!’

‘How should an assessment handle the performance of recently departed staff members? This Department has recently lost a stellar performer or two. That is presumably a negative in terms of assessing the current staff group. But it is a positive in assessing the Department as a corporate entity over a given period of time. And it is arguably a positive in assessing Capability—i.e. this is a Department that demonstrably can attract and/or develop excellent people, notwithstanding that some of them naturally move elsewhere in the course of their career trajectories.’

### **Size of groups assessed**

‘Scale is both a positive and a negative. The large schools/units have major resources (money, space, equipment, staff) to support very strong research groups. Yet, on a per capita basis (i.e. research seen as average output per person) can be far lower than a small efficient unit elsewhere (which is something they found in the New Zealand experience). The approach taken in the present assessment which I focused on tends to emphasise only a section of a school or group which may be completely unrepresentative (or not). A much smaller school might be able to make a similar type of submission but they will be immediately disadvantaged through having a smaller range of cases to choose from, or less opportunity for internationalisation, or less resources. I wonder if any serious research assessment exercise has to get to the level of individuals for it to have any meaning?’

### **Quality**

‘Quality was very difficult to assess as a broad rating. The documents clearly show that there are some pockets of international, best practice, quality. A copy of the 2004 snapshot would flesh this out but I am not sure by how much.’

‘After looking at a number of the different cases I found any “hard evidence” in the form of lists of things (grants, publications, prizes etc.) was relatively straightforward to assess.’

‘On quality, I accept the argument that effort to build international connections should be recognized, but the number and standing of publications is still central to comparative assessments. Much of this core data is in the Annual Report, showing good growth in output and a strong recent record on grants, but little else to make me confident of quality (two researchers appear to have very good standing).’

‘On quality therefore I still would have liked to see data on publications and citations. That said the standing of the senior staff appears very impressive from both grants and external respect.’

‘On impact, the range of indicators was useful and impressive. I give weight to the fact that outsiders pay money for research and consultancies, particularly outsiders who have a reputation for canny purchasing or for the quality of their own research.’

‘The listing of “ten career-best publications” (Addendum C) is very useful. Should there be a time limit—e.g. is a (very good) book from 1981 still admissible? The listing of ARC Grants (Addendum D) is likewise useful. The quotes from the ARC assessors’ reports, where provided, are interesting—though I assume that they are also very selective. Why only ARC Grants? And why present only the project summary at the time of application—why not a summary of outcomes where possible?’

‘The reference to an international rating of political science research centres was interesting and frank. It points to the limits of such exercises in fields where relevance and impact are most likely to be domestic (for similar reasons national journals may have higher standards than international ones as they require greater depth of understanding of institutions, culture etc). This contrasts sharply I suspect with assessments in the scientific and medical research fields.’

‘This has a bearing on the rating scale for quality, though I think the guidance for assessors generally allows sufficient flexibility. It also is relevant to impact. My own experience in setting frameworks for prioritizing research (in defence, health and public administration) has been to identify three elements: first, research that is excellent at the international level contributing to knowledge of global use/interest; second, research that can only be expected to be done in Australia (because of environmental or social or institutional factors etc); and third, research that ensures we are informed purchasers of overseas research, and able to adapt it to our own circumstances (this justifies a broad base of research, but not necessarily at great depth). The balance between the three depends upon the field.’

### **Assessing the work**

‘It has become clear to me that assessors definitely need to have actual publications to peruse—i.e. the four pieces of work per researcher as foreshadowed in the covering documentation but (unless I have missed something) not part of this trial exercise. Lists of publication outputs provide a good handle on quantity and some handle on quality—but not as much as actually being able to skim over the four pieces of work per researcher that those researchers themselves claim to be their “highest quality” pieces. Without this material, the assessment exercise is too sterile.’

### **Bibliometric indicators**

‘The feedback effects have been insufficiently reflected on, I think—for example the skewing of publication patterns towards US journals, which is probably a danger for Australia as well as the UK.’

‘In [this field] bibliometric data is largely meaningless (very long journal half-lives, very small citation counts) but most informed assessors could quickly tell which journals are significant, what the expected pattern of publication is etc.’

### **Journal quality**

‘Academic journals and publishers often have strong affiliations (or are indeed owned by) particular institutions. This is a complex factor which is often at the back of my mind when I assess research—but I don’t have any answers. For instance, the key architectural scientists in Australasia work for the institution that publishes the only journal on architectural science in Australia and hosts the only annual conference in the region. In some senses they are only publishing their own work (this is slightly unfair I know but consider the case in isolation). How can we assess such work? If most of the publications from a particular institution come from that same institutions’ own press does it matter? If there is a long association, say between the University of Melbourne and MIT press in Boston, because of regular academic exchanges should we disregard publications from Melbourne uni academics from MIT press because they possess an unreasonable and uncompetitive advantage? A simple database search reveals that many of these publishing anomalies exist. Again small universities often do not have either a tame university press or close ties to international publishers and are once more disadvantaged.’

‘Because of close knowledge of the field, I am able to accept (or reject) claims such as whether international journals are important indicators of quality. But there is also a risk of conflict of interest (serious if this were a proper assessment not a trial) and of reading more into the material than is presented (not necessarily biasing the result, but certainly compensating for the relative quality of the documentation).’



### Prizes and awards as indicators of quality

‘The awards and prizes list is very valuable in referring to full-time academic staff awards and distinctions. It is slightly less valuable when referring to student prizes and the long list of design awards to honorary, casual or adjunct staff, just seems to me a bit misleading—sorry! If every university listed the achievements of its honorary professional staff many other groups would have similar scale lists to [this institution] and this would simply dilute the impact.’

‘While I have some problem with the prizes section and the industry impact sections they can be, if carefully compiled, good evidence for a research assessment exercise.’

### Impact

‘I found the prizes and industry projects sections from [this institution] made a useful contribution towards my understanding of the high impact of their work. The prizes (ignoring my earlier comments about honorary staff for the moment), perhaps because they are quantitative, were easy to grasp as evidence of impact. The industry case studies ... were more difficult to quantify and I suspect that most schools of [this discipline] could come up with examples that, superficially at least, looked similar in quality and impact.’

‘On impact, the case examples were quite powerful, illustrating that the multidisciplinary approach is real and useful. I was uneasy about the press evidence as it seemed to indicate (at least in part) a vagarious interest in why this research was being done rather than demonstrate a respect for the expertise in an area of undoubted public importance. My impression was that a more substantial impact was expected as the current work advances, and the group is not yet able to demonstrate the effect of some others.’

‘On Impact. I’m not sure what to suggest here, but the criteria are again a bit unsatisfactory because they are again worded as if geography matters. Why distinguish between international or national over regional or local? I can see a case for something like Scope—but why should a bit of research that impacts on Perth count for less than a bit of research that impacts (say) on Singapore? What is the intrinsic impact difference between them?—I can’t think of any. Another example. I could imagine a piece of social research has absolutely transformed the delivery of health services to specific Indigenous communities. Should this sub-national focus matter? Could, for example, intensity of impact make up for a narrower geographical focus?’

‘This case is clearly, and properly, focused on research which actually affects the development of the [practice of the discipline], rather than research on esoteric questions of little practical importance. However, impact is still difficult to pin down. Do they have any other potential measures of it?’

‘Certainly I agree with [this institution] that impact must go beyond the scholarly effect (which is mostly picked up under quality). I also agree there is a qualitative as well as quantitative aspect. The former is nonetheless harder for an assessor without expertise in the field to handle with confidence. Judgments are assisted by evidence such as funding by users (e.g. via linkages partnerships), consultancies and invitations to present expert advice, particularly by major government authorities. Any evidence of the expert advice influencing decisions is particularly useful (as provided in both the case examples from [this institution]). Media contributions can be a useful indicator, but the key here is that the contribution is based on the researcher’s expertise, not notoriety—generally even a non-expert can judge that from the evidence presented.’

‘The societal impact has a time horizon, as [this institution] highlights, that makes it particularly hard to judge both new research organizations and the more theoretically focused ones. But there are ways of informing judgments of the impact even of these, such as through their strategic partnerships and sponsors. Citations are important for more theoretical centres.’

### Capability

‘Curiously, from the documents provided, I felt better able to rank capability than quality or impact. I found quality very difficult to rank just on the documents provided. For instance, the list of ARC grants is really the work of 1 primary member of staff and two secondary ones. This tells me that there is at least one world class staff member it doesn’t tell me much more about a group which the overview suggests is quite extensive.’

‘I was disappointed in the limited information on capability. There was some recognition of weakness at the middle level, and I liked the idea of taking into account the successful employment of past research students. But to be confident of capability into the future I need more information on student trends, strategic planning, governance, building new relationships etc.’

‘On capability, there was a refreshing honesty in the Annual Report even if it took effort to go through it. The growth in student numbers and the level of grants are impressive, and substantial effort is being put into governance and strategic planning etc. The risk summary was informative. It seems clear this is a work in progress, but with some useful runs on the board.’

‘This is one of the reasons that I have ranked the Department high on Capability. I think its history does demonstrate a model of best practice in the strategic development and management of research capabilities—i.e. getting a lot of value and prominence out of whatever resources are at hand. But will other assessors be using a different conceptualization?’

‘Understandably, we must draw heavily on the track record of the key players, particularly in assessing quality and impact. But in this situation I think we need much more on their strategic planning, their governance model, their intentions about partnerships and collaborations etc., and how the new organization will take advantage of the elements brought together. Some sense of this is in the material, but not enough to give me confidence in rating their capability into the future.’

‘The ... postgraduate programs at [this institution] are, to my personal knowledge, not only world class but also world leading. The documents that describe these programs cannot capture the full importance, richness, or innovation, of this case. However, the documents do go part way towards explaining to an assessor some of this significance. At the end of that case document I was left with some clear impressions about impact and capability (but not quality).

### C3 Documentation for the trial process

The seven participating research units were issued instructions which set out the process to be followed in assembling their cases. The assessors were similarly given an outline on how they should approach the task of assessing the work.

These instructions are set out below.

#### a) Criteria

For each of the categories of quality, impact and capability, assessment will be made on a five-point scale, as follows:

##### Quality

The assessment of quality is of the intrinsic academic or artistic excellence of the work, and its significance and originality within its field or discipline. The rating to be awarded is against a scale broadly comparable with that of the UK Research Assessment Exercise. The body of work assessed, in whole or in part, should be judged according to its significance: internationally significant work is of a quality comparable to that of the best work in the field; nationally significant work is of a quality comparable to that of the best work in the field in Australia.

Where a field is concerned principally with Australian issues or those of the local region (for instance, Asia-Pacific economics, or Indigenous Australian music), it is difficult for research to achieve an international profile. This should be taken into account by assessors, who should explicitly identify the criteria by which seminal work in such fields is equivalent to internationally significant research.

In the creative arts, the outcomes of artistic or creative practice are regarded as ‘research’ if the work demonstrates originality, or a level of influence in its field, analogous to similar-quality academic research.

Likewise, non-standard media for the dissemination of research outcomes (for instance, documentaries, policy reports, or action research processes) can be judged for quality by academic peers in the same way as more traditional outcomes, although necessarily on a discipline-specific basis.

The body of work as a whole is then to be awarded a rating for quality, on the basis of the following distributions:

	Internationally significant	Nationally significant	Less than nationally significant
<b>5</b>	50–100%	0–50%	0–10%
<b>4</b>	20–50%	30–70%	0–20%
<b>3</b>	0–20%	50–100%	0–30%
<b>2</b>	0–20%	20–50%	30–60%
<b>1</b>	0%	0–50%	50–100%

### Impact

The assessment of impact is to do with the contribution the work has made to society-at-large outside the academic discipline as practised in universities and other publicly funded research institutions. The term has the potential for confusion, as ‘high impact’ often refers (for instance, in the bibliometric literature) to publication in prestigious or widely read journals.

The evaluation of impact is almost entirely discipline-specific. For instance, in law, work that has high impact will be work that has had demonstrable widespread influence on judgments; on commercial and professional legal practice; on the behaviour of government and community organisations dealing with legal issues, and so on. For musical performance, high-impact work is that which has received widespread notice through, for instance, performance at venues of major significance, recording with wide distribution through sales or broadcast, or critical acclaim. For work in areas of social significance, impact can be evaluated through take-up in policy formation, influence on professional and government practice, media exposure, and so on.

The case for impact will therefore necessarily be made on a discipline-by-discipline basis. However, regardless of discipline, evidence of successful collaborative ventures between academic units and other organisations—government, industry, the professions, charitable and cultural organisations, to name a few—constitutes a solid basis for demonstrated impact.

Assessors are asked to judge both the proportion of the work that has had some degree of impact, and also as to whether that impact is on a national or international extent, or more local. For instance, legal research that has informed legislation or judgments that have established significant precedent, or has influenced corporate legal behaviour at a national level, would be good examples of the former. In musical performance, the equivalent would be national broadcast, performance at nationally recognised venues, or critical reception in the national press.

The assessors will award a rating on the following scale:

- 5 The majority of the work has had an impact on society at a national or international level.
- 4 The majority of the work has had demonstrable impact on society; less than half has been at a national or international level.
- 3 The majority of the work has had a demonstrable impact on society.
- 2 Some of the work has had a demonstrable impact on society.
- 1 No appreciable impact on society has been demonstrated.

## Capability

‘Capability’ is a term that has been adopted during the course of the consultation phase of this project. It draws together a number of related themes that have been identified by stakeholders as significant in the assessment of research, but that are not captured by ‘quality’ or ‘impact’. Significantly, many of these themes have been identified in the RQF Issues Paper (March 2005) as important aspects of the research effort that possibly should be considered in any research assessment framework.

Broadly, ‘capability’ refers to the capacity of a research unit to enhance research skills, capacity and ability: those of itself, its partners, and of Australia in general. It represents the success or otherwise of a unit in adopting a strategic approach to research management. Capability is measured by a number of pieces of data and evidence that will give a reliable indication of a unit’s likelihood of producing high-quality, high-impact research in the future.

A representative number of indicators of capability are given in Attachment 2. It will be seen that these cover, broadly, the strategic management and planning of the research unit; its success in finding appropriate collaborative partners inside and outside the publicly funded research sector; its management of finances and infrastructure that support the research effort; its successful recruitment and development of early-career researchers; and its success in research training and in providing education that is led by and focused towards research in a way that demonstrably enhances the research capability of the discipline.

As opposed to quality and impact, which are, for the purposes of this exercise, explicitly attributes of research outcomes, capability revolves around people: their research skills, their deployment, the way they are resourced and supported, and the way in which they collaborate. In a healthy and dynamic national research environment, it must be acknowledged that people move around, and therefore that investment in the development of research capabilities is an investment that pays dividends not just to the unit making that investment, but to the research capacity of the nation as a whole. For instance, a unit that has a flourishing research training program that delivers high-quality researchers and research-capable professionals into the workforce—both the academic and non-academic workforce—has made a valuable contribution to the research capability of the nation, irrespective of whether those researchers ultimately contribute research outcomes to that particular unit.

Given the wide variation in size of units, in the hugely variable costs involved in supporting the research efforts of different disciplines, and in the differing strategic aims of research units in different parts of the sector, the assessment of capability is best made relative to opportunity. Therefore, the criteria have been framed in this way.

Relative to its size, its strategic orientation, and the particular nature of its disciplines, the research unit:

- 5 provides a model of best practice in the strategic development and management of research capabilities, making it a national leader in its field
- 4 demonstrates a highly successful approach to the strategic development and management of research capabilities, above average in its field in national terms
- 3 demonstrates a successful approach to the strategic development and management of research capabilities, broadly consistent with practice nationally
- 2 demonstrates an adequate approach to the strategic development and management of research capabilities
- 1 displays some significant inconsistencies or instances of unrealised potential in its strategic development and management of research capabilities.

## b) Indicators and evidence

Research units will be asked to provide a brief overall statement about the context in which the research takes place, the strategic aims of the unit, its organisation, and its orientation within the discipline(s) it practises. Then, it will be asked to make specific claims to excellence against the criteria—a

self-assessment—within each of the three categories of quality, impact and capability. These claims should be supported by evidence and indicators covering as many of the following as are appropriate to the discipline(s).

For the purposes of this exercise, research to be assessed will be limited to that published since January 2000.

As the purpose of the exercise is not to arrive at genuine assessments of the research of a unit, but to test the validity and usefulness of specific measures, units are encouraged to provide any other data or evidence which they feel supports their claims to excellence.

### Quality

Evidence used to assess the quality of research should include:

- four pieces of work from each researcher participating in the exercise (including creative work and non-standard research outputs where appropriate to the discipline); this work must have been published since January 2000
- any reviews of these four pieces of work; the standing of the referees for publication; or the works' critical reception in the case of creative works.

For *all* research outputs of the participating researchers completed since January 2000 (i.e. not simply the four selected sample works):

- a fully referenced list of publications
- bibliometric indicators including the quantum of publication and the citation impact of the research
- evidence of publication or dissemination that attests to quality: for instance, publications in high-quality journals (or equivalent dissemination of creative works through venues or media of high artistic profile).

For all participating *researchers*:

- data on the academic standing and esteem of participating researchers, including:
  - membership of learned societies and other academic bodies
  - editorships of journals and series
  - prizes, fellowships and awards
  - invitations to participate in significant academic fora such as international conferences and public lecture series
- statistics (number and value) and qualitative accounts of *successfully acquitted* contestable research grants received by the researchers.

### Impact

Evidence used to assess the impact on society of the research will be highly discipline-specific. Nevertheless, the following broad categories will help to frame discipline-specific indicators. Units should give an account of the impact on society of their research. Indicators of such impact should be given using data since January 2000; however, the research itself may have been completed before that time. The account should be supported by evidence, for instance, evidence of:

- the take-up, recognition or influence of research or creative work by users in industry, government, community and cultural organisations
- indicators of social, cultural, artistic and economic benefits to the community that flow directly or indirectly from the research or creative work, where such a link can be demonstrated conclusively
- media profile or other public exposure (e.g. online accession) of the research or creative work
- *successfully acquitted* collaborative projects with partners outside the academic sector

- membership of professional, industry or cultural organisations or advisory groups, where relevant to the discipline
- commercial returns on research outcomes and creative products.

### Capability

The unit should provide a narrative account of its strategic research planning and management processes, drawing particular attention to how these will enhance the capabilities of the unit, its partners, and the nation. This should be supported by evidence and indicators from the period January 2000 to the present that include:

- budgetary data
- number of staff, and their academic position
- research students' enrolments, completions and graduate destination
- evidence of the recruitment and development of early-career researchers
- details of research infrastructure and plans for its development and management
- statistics and accounts of *ongoing* externally funded research projects
- details of *ongoing* collaborations both with other academic units and with partners in industry, government, the community and cultural sector
- alignment of research plans with National Research Priorities.

### c) Evaluating the assessment of quality, impact and capability in research— instructions to assessors

Seven academic departments have each made a case for excellence for their research. These are in the disciplines of history, music, politics and public policy, architecture and design, marketing, law, and (broadly) cultural studies. As an 'assessor' in this process, you have been asked to examine this case for excellence in three different categories:

- the *quality* of the research, in terms of academic standing and intrinsic excellence
- the *impact* of research, in terms of its influence on society at large
- the *capability* of the research unit, and its contribution to developing the research skills and capabilities of the nation.

A fuller definition of these, together with some criteria against which they may be assessed and some examples of the sort of evidence which might be presented to make the case, is contained in the document 'The assessment process' which is on the website.

As part of this exercise, we are asking each assessor to enter a numerical rating for each of the three categories. However, this is purely to simulate a research assessment process—there is no intention of actually attempting to achieve meaningful assessments of the cases presented. The point of this exercise is not to evaluate the research, but to *evaluate the nature of the evidence* presented for research quality, impact and capability. In other words, we are attempting to elicit answers to questions such as:

- What evidence needs to be presented to come to a reliable judgment on research quality?
- What are the best indicators of the impact of research? How discipline-specific are these?
- What evidence needs to be presented to demonstrate a flourishing research culture that enables the development of research students and early-career researchers, and thus contributes to the development of research capability in the future?

To achieve this, we are asking each assessor to follow these three steps:

- 1 Read the ‘case for excellence’ presented, with an eye to making an evaluation in each of the three categories: quality, impact, capability.
- 2 Enter a numerical rating for each in the area of the website entitled ‘Assessment’.
- 3 In the website area entitled ‘Case study discussion’, post some comments regarding the nature of the evidence presented. For instance, some typical questions or issues might be:
  - In an assessment of quality, how important is it to actually see or read the work itself?
  - If bibliometric or other quantitative indicators of research quality were presented, how useful were these in making an assessment?
  - How convincing was the case for impact? How could it be tested for validity? What other or better indicators of impact might there be in your discipline?
  - What further information or evidence would you have liked to have seen that would have improved the confidence you had in your judgment?

If you would like to make any general observations about research assessment, or ask questions for discussion, they can be posted in the section of the website entitled ‘Questions and provocations’. Any purely practical observations about this exercise (for instance, technical issues to do with the website) can be posted in the section called ‘Feedback on the process’.

Finally, although you have been invited to take part in this process because of your expertise in one of the disciplines that has put forward a case for excellence, we are also interested to determine how comfortable you as an assessor would feel in making judgments, particularly about ‘impact’ and ‘capability’ in disciplines further afield. For instance, assessors in law might wish to examine the case made in public policy; assessors in history might wish to examine the case in cultural studies; and so on. Feel free to examine, assess and comment on any of the cases presented, not simply the one in your discipline.

## C4 A postscript

CHASS would like to thank all the participants and assessors, who in many cases put in a vast amount of work for no reward other than to do justice to this complex and fraught issue of research quality and research impact assessment. The cases and assessments constitute hugely valuable sources of information for this project in their own right, and deserve a wider audience.

The trial took place on a custom-built website at <http://sts-dev.anu.edu.au/moodle/chass>. In the fullness of time, it would be desirable for this website to be made available as an adjunct to this report.

However, for this to happen the permission of all participants and assessors will be required, to protect their privacy and intellectual property rights. Also, participants should have the opportunity to reply to some of the assessor comments. An ongoing dialogue of this kind will provide further analysis of the issues dealt with in this report.

Once these requirements have been achieved, the site will be made available for wider access.



# Appendix D

## Acknowledgments

CHASS is grateful to the hundreds of scholars in the humanities, arts and social sciences, and many from other areas, including the Department of Education, Science and Training, who have contributed to the work through the case studies, focus groups and the reference group.

### Reference group

Name	Organisation
Ms Sarah Barns	Australia Council
Professor Tim Brown	Australian National University
Dr Linda Butler	Australian National University
Professor Stewart Clegg	University of Technology Sydney
Associate Professor Tony Dalton	RMIT University
Professor Mark Dodgson	University of Queensland
Dr Claire Donovan	Australian National University
Professor Alistair Fox	University of Otago
Professor Cindy Gallois	University of Queensland
Professor Malcolm Gillies	Australian National University
Professor John Hartley	Queensland University of Technology
Ms Elizabeth Hoole	National Health and Medical Research Council
Professor Helen Fullgrave	National Health and Medical Research Council
Professor Bryan Horrigan	University of Canberra
Dr John Howard	Howard Partners
Professor Lesley Johnson	Griffith University
Ms Michelle Leggo	National Health and Medical Research Council
Professor Terry Lovat	University of Newcastle
Dr Robert O'Connor	Australian Vice-Chancellors' Committee
Professor Elim Papadakis	Australian Research Council
Dr Jonathan Powles	Australian National University
Professor Sue Richardson	Flinders University
Mr Ken Richardson	University of Queensland
Professor Linda Rosenman	University of Queensland
Professor Susan Rowley	University of Technology Sydney
Associate Professor Huib Schippers	Griffith University
Professor Deryck Schreuder	University of Western Australia
Professor Margaret Seares	University of Western Australia
Professor David Siddle	University of Queensland
Dr Mandy Thomas	Australian Research Council

## Respondents

Name	Organisation
Professor Fred Affleck	Curtin University of Technology
Professor Pauline Allen	Australian Catholic University
Dr Bryant Allen	Australian National University
Professor Margaret Alston	Charles Sturt University
Professor Ien Ang	University of Western Sydney
Professor Anona Armstrong	Victorian University
Mr David Armstrong	Murdoch University
Dr Michael Arthur-Kelly	University of Newcastle
Professor Andy Arthurs	Queensland University of Technology
Dr Eril Bailly	University of Sydney
Associate Professor Su Baker	University of Melbourne
Professor Bonnie Barber	Murdoch University
Dr Liz Baynham	University of Newcastle
Ms Carolyn Bennett	Curtin Business School
Mr Paul Bentley	Wolanski Foundation
Dr Santina Bertone	Victoria University
Professor Paul Boreham	University of Queensland
Ms Petra Bouvain	University of Canberra
Mr Ben Bradley	Charles Sturt University
Mr Kevin Brophy	University of Melbourne
Emeritus Professor Lois Bryson	RMIT University
Dr Philip Bull	La Trobe University
Professor Mark Burry	RMIT University
Dr Melissa Butcher	University of Sydney
Dr Laurie Buys	Queensland University of Technology
Dr John Byron	Australian National University
Mr Ian Carmody	Australian Institute of Management
Associate Professor David Carter	University of Queensland
Mr Warwick Cathro	National Library of Australia
Ms Bobbi Cerini	Questacon
Professor Wendy Chaboyer	Griffith University
Professor Don Chalmers	University of Tasmania
Associate Professor Ping Chen	University of Queensland
Professor Eugene Clark	Charles Darwin University
Professor Thomas Clarke	University of Technology Sydney
Professor Kevin Clements	University of Queensland
Professor Michael Coper	Australian National University
Professor Russell Craig	National Graduate School of Management, ANU
Professor Peter Cryle	University of Queensland
Professor Joy Cumming	Griffith University
Professor Stuart Cunningham	Queensland University of Technology
Dr George Curry	Curtin University of Technology
Dr Michael Cuthill	University of Queensland
Dr Michael Darcy	University of Western Sydney

<b>Name</b>	<b>Organisation</b>
Ms Marie Davern	Centre for Aboriginal Economic Policy Research
Mr Russell Dawe	Australasian Research Management Society
Professor David de Vaus	Latrobe University
Dr Ann Deden	Open Universities Australia
Mr Brendan Delahunty	Australian Institute for Aboriginal and Torres Strait Islander Studies
Dr Stephanie Donald	University of Technology Sydney
Professor Ian Donaldson	Australian National University
Ms Kim Douglas	Queensland University of Technology
Dr Peter Downton	RMIT University
Professor Steve Dowrick	ARC Senior Fellow Faculty of Economics ANU
Mr Graham Durant-Law	PACE Consultancy Group
Associate Professor Pam Dyer	University of the Sunshine Coast
Dr Wendy Earles	James Cook University
Professor Victor Emeljanow	University of Newcastle
Professor Philip Esler	University of St. Andrews, Scotland
Professor Norman Etherington	University of Western Australia
Mr Harry Evans	Clerk of the Senate
Associate Professor Michael Ewans	University of Newcastle
Ms Linda Ferguson	Australian War Memorial
Associate Professor Jeremy Finn	University of Canterbury, New Zealand
Ms Diana Forward	Queensland University of Technology
Professor Majella Franzmann	University of New England
Professor Lucy Frost	University of Tasmania
Professor Liz Fulop	Griffith University
Ms Debby Gairns	Australian National University
Professor Brian Galligan	University of Melbourne
Dr Debjani Ganguly	Australian National University
Professor Stephen Garton	University of Sydney
Mr Toss Gascoigne	Council for the Humanities, Arts and Social Sciences
Dr Jochen Glaser	Australian National University
Professor Brendan Gleeson	Griffith University
Professor Lyn Gorman	Charles Sturt University
Dr Anne Grahame	Southern Cross University
Mr Jonathan Grant	RAND Europe
Professor Lelia Green	Edith Cowan University
Ms Anne Gregory	Department of Education, Science and Training
Dr Greg Hainge	University of Queensland
Dr Rosalind Halton	University of Newcastle
Mr Ian Harris	House of Representatives, Parliament House
Dr Yasmin Haskell	University of Western Australia
Professor David Hensher	University of Sydney
Ms Katharine Heron	University of Westminster, UK
Professor Tim Hirst	University of Sydney
Professor Bernard Hoffert	Monash University

<b>Name</b>	<b>Organisation</b>
Associate Professor Allyson Holbrook	University of Newcastle
Ms Robyn Holmes	National Library of Australia
Professor Ross Homel	Griffith University
Mr Stephen Horton	Council of Australian Postgraduate Associations
Professor Keith Houghton	Faculty of Economics and National Graduate School of Management, ANU
Ms Helen Hyland	National Gallery of Australia
Professor Ivor Indyk	University of Western Sydney
Professor Dick Johnson	Australian National University
Associate Professor Phillip Jones	University of Sydney
Associate Professor Roberta Julian	University of Tasmania
Professor Ilan Katz	University of New South Wales
Ms Bronwyn Kelly	Deakin University
Professor Susan Kenny	Deakin University
Ms Kerry Kilner	University of Queensland
Professor Max King	Monash University
Professor Sue Kippax	University of New South Wales
Dr Natalie Klein	Macquarie University
Associate Professor Jeri Kroll	Flinders University
Associate Professor Tess Lea	Charles Darwin University
Associate Professor David Lemmings	University of Newcastle
Rev. Dr Lee Levett-Olson	Flinders University
Rev. Professor Bill Loader	Murdoch University
Ms Carmel Lutton	University of Newcastle
Associate Professor Andrew Lynch	University of Western Australia
Dr Lenore Lyons	University of Wollongong
Professor Martyn Lyons	University of New South Wales
Dr Gaynor Macdonald	University of Sydney
Professor Stuart Macintyre	University of Melbourne
Professor Alison Mackinnon	University of South Australia
Ms Gail Mahon	University of South Australia
Associate Professor Bill Martin	Flinders University
Ms Fiona Martin	Southern Cross University
Mr Gavin McCarthy	University of Melbourne
Associate Professor John McDonald	University of Ballarat
Associate Professor Sue McGinty	James Cook University
Dr Paul McKechnie	University of Auckland
Mr Dominic McKenna	Australian National University Choral Society
Dr Andrew McNamara	Queensland University of Technology
Mr Doug McNaught	Emmaus Ministries Australia
Associate Professor Joanna Mendelssohn	University of New South Wales
Professor Bill Merrilees	Griffith University
Ms Jenni Metcalfe	Econnect Communications Pty Ltd
Professor Clive Moore	University of Queensland
Dr Gerard Moore	Sydney College of Divinity

<b>Name</b>	<b>Organisation</b>
Professor John Moorhead	University of Queensland
Professor Sally Morgan	University of Western Australia
Associate Professor Gillian Mort	Griffith University
Professor Bruce Muirhead	Institute for Educational Research, Policy and Evaluation
Professor Kerry Mummery	Central Queensland University
Associate Professor John Murphy	RMIT University
Associate Professor Peter Murphy	Monash University
Associate Professor Shane Murray	RMIT University
Professor Martin Nakata	University of Technology Sydney
Ms Mandy Nearhos	Oxfam, Community Aid Abroad and World Neighbourhood Education Program
Professor Richard Nile	Curtin University of Technology
Associate Professor Alanna Nobbs	Macquarie University
Dr Louise Noble	University of New England
Mr Michael O'Hara	Department of Health and Ageing
Mr Tim Oakley	La Trobe University
Professor Graham Oppy	Monash University
Professor Michael Ostwald	University of Newcastle
Professor William Outhwaite	Essex University
Ms Darlene Oxenham	Curtin University of Technology
Ms Sue Page	University of South Australia
Professor Andrew Parkin	Flinders University
Dr Josko Petkovic	Murdoch University
Dr Malcolm Pettigrove	Australian National University
Dr Robert Phiddian	Flinders University
Associate Professor Kerry Philip Green	University of Canberra
Associate Professor Gail Phillips	Murdoch University
Dr Maggie Phillips	Edith Cowan University
Mr Andrew Podger	Department of the Prime Minister and Cabinet
Dr Adam Possamai	University of Western Sydney
Professor Wilfrid Prest	University of Adelaide
Ms Barbara Preston	Barbara Preston Research
Professor Huw Price	University of Sydney
Rev. Dr Duncan Reid	University of Melbourne
Professor Peter Reiman	University of Sydney
Dr Matthew Rimmer	Australian National University
Dr Victoria Rogers	University of Western Australia
Dr Emma Rooksby	Centre for Applied Philosophy and Public Ethics
Mr Cameron Ross	Canberra Times
Associate Professor David Rowe	University of Newcastle
Associate Professor John Ryan	University of New England
Justice Ronald Sackville	Federal Court
Professor Charles Sampford	Griffith University
Professor Penelope Sanderson	University of Queensland
Associate Professor Angela Scarino	University of South Australia

<b>Name</b>	<b>Organisation</b>
Professor Andrew Schultz	University of Wollongong
Professor Krishna Sen	Curtin University of Technology
Professor Thomas Shapcott	University of Adelaide
Mr Tony Sheil	Innovative Research Universities Australia
Professor June Sinclair	University of Sydney
Ms Robyn Sloggett	University of Melbourne
Ms Claire Smith	Flinders University
Professor Paul Spoonley	Massey University, New Zealand
Ms Aliya Steed	Australian National University
Associate Professor Kate Stevens	University of Western Sydney
Dr Iain Stewart	Macquarie University
Dr John Taylor	Centre for Aboriginal Economic Policy Research
Dr Julian Thomas	Swinburne University of Tehcnology
Dr Lyndal Thorburn	PACE Consultancy Group
Ms Susan Tonkin	National Museum of Australia
Associate Professor David Tripp	Murdoch University
Professor Graeme Turner	University of Queensland
Dr Elizabeth Van Acker	Griffith University
Honorary Associate Professor Walter Veit	Monash University
Dr Peter Veth	Australian Institute for Aboriginal and Torres Strait Islander Studies
Ms Catrina Vignando	Craft Australia
Dr Robert Vincs	Victorian College of the Arts
Ms Jill Waterhouse	Canberra Museum and Gallery
Associate Professor Jennifer Webb	University of Canberra
Associate Professor Ruth Webber	Australian Catholic University
Mr David Webster	Department of Health and Ageing
Professor John Weckert	Centre for Applied Philosophy and Public Ethics
Professor Phillipa Weeks	Australian National University
Professor Patrick Weller	Griffith University
Associate Professor Mark Western	University of Queensland
Dr Margaret White	Macquarie University
Mr Andrew Whitecross	Department of Family and Community Services
Ms Gillian Whitehead	Independent Musician
Ms Katherine Wilkinson	RMIT University
Ms Elizabeth Williams	Australian Bureau of Agriculture and Resource Economics
Dr John Wiltshire	La Trobe University
Ms Tamara Winikoff	National Association for the Visual Arts
Dr David Woods Curtin	University of Technology
Dr Suzettew Worden	Curtin University of Technology
Dr Claire Wyatt-Smith	Griffith University
Dr Mary Journazi	University of Wollongong

# Appendix E

## Abbreviations and acronyms

ABC	Australian Broadcasting Corporation
ABS	Australian Bureau of Statistics
AIATSIS	Australian Institute of Aboriginal and Torres Strait Islander Studies
AIMS	Australian Institute of Marine Science
ANSTO	Australian Nuclear Science and Technology Organisation
ANU	Australian National University
ARC	Australian Research Council
CHASS	Council for the Humanities, Arts and Social Sciences
CSIRO	Commonwealth Scientific and Industrial Research Organisation
DEST	Department of Education, Science and Training
FaCS	Department of Family and Community Services
GERD	gross domestic expenditure on research and development
HASS	humanities, arts and social sciences
HCA	humanities and creative arts
HERG	Higher Education Research Group (Brunel University)
IGS	Institutional Grants Scheme
IP	intellectual property
ISI	Institute for Scientific Information
NGO	non-government organisation
NHMRC	National Health and Medical Research Council
OECD	Organisation for Economic Co-operation and Development
PFRA	publicly funded research agency
PNG	Papua New Guinea
R&D	research and development
RAE	Research Assessment Exercise
RFCd	research fields, courses and disciplines
RQF	Research Quality Framework
RTS	Research Training Scheme
SBS	Special Broadcasting Service
STEM	science, technology, engineering and medicine



Council for the Humanities, Arts & Social Sciences

28 Balmain Crescent Australian National University ACT 2601

phone (02) 6249 1995 fax (02) 6247 4335 [www.chass.org.au](http://www.chass.org.au)

