



Research Evaluation and Policy Project
Research School of Social Sciences

**A Review of Current Australian and International Practice in
Measuring the Quality and Impact of Publicly Funded Research in the
Humanities, Arts and Social Sciences**

Claire Donovan

**REPP Discussion Paper
05/3
November 2005**

Donovan, Claire

A Review of current Australian and International Practice in Assessing the Quality and Impact of Publicly funded Research in the Humanities, Arts and Social Sciences

E-mail: claire.donovan@anu.edu.au

REPP Discussion Papers are scholarly papers that report research in progress. They can be downloaded free of charge (PDF)

© by the author

Research Evaluation and Policy Project

Research School of Social Sciences

The Australian National University

Canberra ACT 0200

Tel: (02) 6125 4849

Fax: (02) 6125 9767

<http://repp.anu.edu.au>

CONTENTS

1 INTRODUCTION	1
2 CURRENT MEASURES OF QUALITY AND IMPACT OF PUBLICLY FUNDED HASS RESEARCH IN AUSTRALIA.....	1
2.1 University block funding – the IGS and RTS formulas	1
The research publications component	2
Conclusion.....	3
2.2 Measuring research production: a closer look at publication indicators and HASS	4
The social sciences and humanities	4
Creative arts and design	5
Conclusion.....	6
2.3 Publicly funded research agencies and research funding agencies – current measures of quality and impact.....	6
2.3.1 Publicly funded research agencies	7
Measures of research quality	7
Measures of research impact.....	8
2.3.2 Publicly funded research funding agencies	10
The Australian Research Council.....	10
<i>Measures of research quality</i>	11
<i>Humanities and creative arts</i>	11
<i>Measures of research impact</i>	12
<i>Humanities and creative arts</i>	13
The National Health and Medical Research Council	14
<i>Measures of research quality</i>	14
<i>Measures of research impact</i>	15
<i>Conclusion</i>	16
2.4 Measuring research quality: a closer look at citation indicators and HASS.....	17
Citation indicators, science policy and HASS	17
Coverage of HASS publications.....	18
Conclusion.....	19
3 CURRENT INTERNATIONAL MEASURES OF QUALITY AND IMPACT OF PUBLICLY FUNDED HASS RESEARCH	19
Measures of research quality	20
Measures of research impact.....	22
Conclusion.....	23
4 POTENTIAL MEASURES OF QUALITY AND IMPACT OF PUBLICLY FUNDED HASS	23
4.1 Lessons from Australian practice.....	24
4.2 Lessons from international practice.....	25
4.3 New ways of measuring quality and impact currently being postulated.....	26

Measures of research quality	26
Measures of research impact.....	27
4.4 Measuring research quality: a closer look at non-bibliometric indicators.....	30
4.5 Measuring research impact: a closer look at qualitative modelling	31
REFERENCES	33
TABLES	
Table 1: Payback categories for arthritis research	29
FIGURES	
Figure 1: CSIRO - Overview of Selected Achievements, 2003-04.....	9
Figure 2: Example of 'payback profile' impact assessment	29

1 Introduction¹

The purpose of this review is (1) to describe current practice in measuring the quality and impact of publicly funded HASS research in Australia and internationally; and then (2) to consider what it is that we may glean from Australian and international practice in assessing both HASS and SET, 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 HASS research. While these terms 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.

2 Current measures of quality and impact of publicly funded HASS research in Australia

This section of the review describes current Australian practice in measuring the quality and impact of HASS research. Section 2.1 reveals that there is no existing direct measure of research quality or research impact tied to government block funding of Australian universities. Section 2.3 examines measures of quality and impact used to assess SET and HASS research supported by publicly funded research funding agencies. We find that despite advances made in the SET sector, the notion of research impact is not operationalised in respect of specific HASS benefits, and 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 applying these SET-derived publication and citation-based quality measures to HASS streams.

2.1 University block funding – the IGS and RTS formulas

The Institutional Grants Scheme (IGS) and Research Training Scheme (RTS) were introduced as a means 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 2004-5 estimates, the IGS and RTS formulas which constitute ‘Performance Based Block Funding’ guide the distribution of the greatest single financial outlay within the Commonwealth Science and Innovation budget, accounting for \$1203.5m or 22.5% of spending.² However, these measures do not directly gauge research quality or research impact.

¹ This research was funded by the Council for Humanities, Arts and Social Sciences (CHASS) with the financial assistance of the Australian Government, through the Department of Education, Science and Training. This report appears as Appendix A in Council for the Humanities, Arts and Social Sciences (2005) *Measures of Quality and Impact of Publicly Funded Research in the Humanities, Arts and Social Sciences*, CHASS Occasional Paper 2, Canberra.

² http://www.dest.gov.au/science/analysis/pdf/Science%20and%20Innovation%20Budget%20Tables%201%20to%205_2004-05.pdf.

IGS formula weightings
60% = Research income
30% = Commonwealth-funded research students
10% = Research publications
RTS formula weightings
50% = All higher research degree completions
40% = Total research income
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 predicate the quality or impact of research outcomes. There are two output measures: higher degree completions, which are 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.

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 DEST categories: (A1) a scholarly book produced by a commercial publisher (worth five ‘DEST points’),³ (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), and (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 the Australian case, the effect of distributing funds on this basis has run contrary to government intentions because increased research production has been rewarded rather than improvements in research quality. 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 SET and HASS fields alike.⁴

The 2004 evaluation of the *Knowledge and Innovation Reforms* discussed the publications component of the IGS and RTS formulas, and identified three weaknesses (2004, 10). The first, as already identified above, was that the publications count rewarded quantity rather than quality. The second was that the publications component

³ Funds tied to publication ÷ all publications (weighted by type) = value per ‘publication unit’ or ‘DEST point’.

⁴ 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 2001-2.

was highly correlated with the other elements of the IGS and RTS formulas, particularly research income, and thereby provided limited extra value. The third issue was that the count is open to criticism in terms 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.

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. where 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.^{5 6} 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 audio-visual recordings, computer software, technical drawing, designs and models, and other creative works including recorded works and individual exhibition or representation of original art. The current four DEST categories came into effect in 1995, although in 2002 a reduced list of creative arts and design was reintroduced based upon 2001 returns, but was dropped by 2003. The inclusion or absence 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 beg the question 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.

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 SET or HASS research if the aim of government policy is to measure research quality and research impact.

⁵ Section 2.2 below takes this discussion further and draws out particular considerations for publication measures applied to HASS generally, by stream and by field.

⁶ 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 as they do not follow the required refereeing or publication protocols.

2.2 Measuring research production: a closer look at publication indicators and HASS

Section 2.1 above 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 upon specific issues involving publication-based measures and HASS research.⁷

‘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 upon 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 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 upon 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 2.4 below for an overview of citation indicators and HASS).

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 upon 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 thereby differently structured and that this has bibliometric consequences (Glänzel and Shoepflin 1999, 31; Hicks 2004, 473; Luwel *et al.* 1999, 13; Nederhof *et al.* 1989, 427). Publication-based measures were originally developed for SET using journal articles and conference proceedings as the basis for scholarly communication. When applied to HASS, such a model excludes core categories of publication within humanities and social science fields, such as books and book chapters, and will ignore non-scholarly ‘enlightenment’ literature aimed at practitioners and the

⁷ 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>

public.⁸ In fact, the idealised model of scientific communication does not equally serve all SET 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).

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 vastly 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’ (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’ (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’ (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 instance, in the case of music, a text is something that can be read, ‘whether it is in visual, audio or written forms’ (1998, 54) and an opera ‘could be performed on stage, recorded on compact disc and filmed all at the same time’ (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.

⁸ Luwell *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, which will be excluded from standard publication counts.

The view that only scholarly texts count as publication clearly excludes the central activities of creative arts and design. For example, a draft review of 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 2.1 above), Australia's current counting system when applied to arts, media and design is inadequate and lags far behind other international practices.

Conclusion

To sum up, if an indicator is designed to measure SET publications it will detect the most SET-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 served to illuminate 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 HASS 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 SET-friendly model of journal publishing. HASS research will not be favoured by such arrangements, a subject to which we shall turn in Section 2.4 below. 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.

2.3 Publicly funded research agencies and research funding agencies – current measures of quality and impact

Public money is used in Australia to fund three SET-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).⁹ A brief overview of their current measures of research quality and research impact is included because, as demonstrated above, it is the norm for SET-based measures to be applied wholesale to HASS, and it is therefore important to consider what are the potential consequences of this. There follows an assessment of

⁹ The Defence Science and Technology Organisation (DSTO) is excluded here as it falls outside of the remit of DEST, being under the leadership of the Chief Scientist as part of the Department of Defence.

current measures of quality and impact employed by Australia's publicly-supported research funding agencies: the Australian Research Council (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 SET-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 SET research to measuring research impact in economic, policy, social, intellectual, artistic, cultural, educational and community terms.

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

2.3.1 Publicly funded research agencies

It may seem strange to discuss AIMS, ANSTO and CSIRO in a HASS-centred report as these are SET-focussed institutions, yet there is an important reason for briefly doing so: SET practice in evaluating research has traditionally provided the template for national research policy. The RQF consultation process considers measures of quality and impact which are to be applied to university-based research *and* 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 upon HASS. 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 while we write this report these institutions are considering measures of quality and impact for SET research, so it is prudent to ascertain whether they will advocate traditional SET-friendly measures or alternative indicators which may be more complementary to HASS, developments which shall be discussed in Sections 4.1 and 4.5 below.¹⁰

Measures of research quality¹¹

All of these publicly funded research agencies use 'number of publications' as a performance indicator, although we have established that measuring research output does

¹⁰ To further bolster the case for considering current (and potential future) measures of research quality and impact used by these SET bodies, it is worth remembering that CSIRO received an estimated \$576.5m or 10.8% of the Commonwealth Science and Innovation Budget for 2004-5 and 'Other R&D Agencies' (which include AIMS and ANSTO) received \$398.6m or 7.5%, compared with \$1203.5m or 22.5% allocated for university block-funding under the IGS and RTS formulas. SET dedicated Cooperative Research Centres attract \$193m or 3.6% of Commonwealth funds, and in 2003-4 the CSIRO alone participated in 50 of the 71 centres. So in other words, at 21.9%, funding for these three areas of dedicated SET 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 for Commonwealth Science and Innovation expenditure figures).

¹¹ See AIMS (2003, 15), ANSTO (2004, 18-21) and CSIRO (2004, 23ff).

not unearth any truths about research quality. AIMS does not specify what kinds of publications they count, while ANSTO and CSIRO together 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. 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 upon 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 (2003, 18).¹²

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. Prizes, honours and peer assessment are measures of quality which can in principle be applied equitably to SET 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, an issue elaborated in depth in Section 2.4 below.

Measures of research impact

There is less cross-organisational consensus to be found in approaches towards 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 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’.

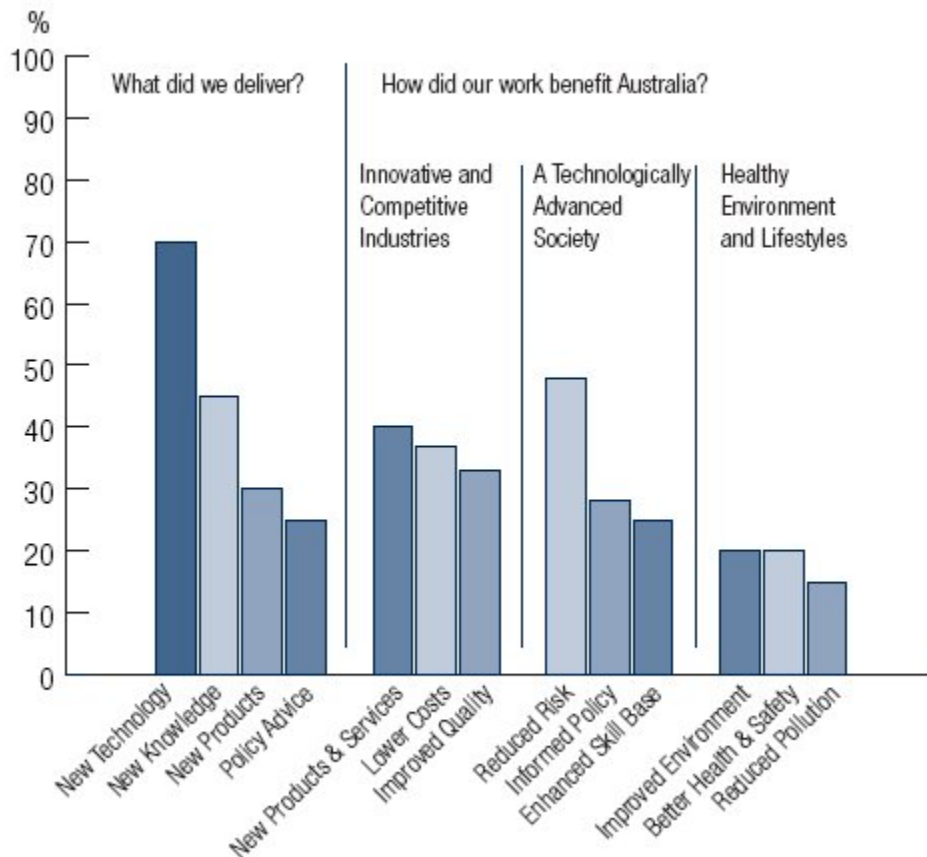
All three institutions gauge research impact through the advice they give to government, ANSTO in terms of its level of involvement in international policy developments by number of projects, person years and money expended; CSIRO through submissions, verbal reports and presentations, advice given for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

¹² 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.

ANSTO uniquely thinks about its impact in terms of the number of visitors that 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 (SEI) 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 Chart 1 below) and individually (2004, 38-41).¹³

Figure 1: CSIRO - Overview of Selected Achievements, 2003-04¹⁴



¹³ CSIRO’s four Research Groups are Information Technology, Manufacturing and Services; Sustainable Minerals and Energy; Environment and Natural Resources; and Agribusiness and Health.

¹⁴ Reproduced with permission from CSIRO (2004, 39).

Although the precise methodology is not presented, we can see that impact ‘outputs’ cover new knowledge, technology and products, and policy advice, while benefits accrue to Australian business, the technology society, and environment and lifestyle. It is noted that when taken individually, there are different outcomes for the four main strands of CSIRO research.

To sum up, we again find that there are no specific arrangements for evaluating the impact of any HASS research supported. 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 SET’s lead and variously demonstrate the impact of research upon 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-Outputs 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 4.1 and 4.5 below.

2.3.2 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 the Social, Behavioural and Economic Sciences Program, and through the NHMRC which funds some medical and health-related HASS research. 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 there are distinct arrangements for evaluating HASS research. We find that the ARC is quite naturally our most important consideration and that current evaluation practices of quality are publication or output focussed (particularly in terms of citations gained), which will not favour HASS, as explained in Section 2.4 below. 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 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 financial year 2003-04 the ARC received \$482.4m or 9.0% of the Commonwealth Science and Innovation budget. Total ARC research funding for 2002-06 gives over a fifth of funds to HASS research, representing the single largest Australian source of competitive HASS funding.¹⁵

¹⁵ The HASS share of expenditure on Discovery Grants is 23.3%, Research Fellowships, 24.2%, Federation Fellows 20%, and Linkage Projects 22.2% (Source: 2002 grant statistics in ARC 2004, 153-7).

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 SET and HASS 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 SET-relevant. However, as end of grant reports may be submitted not later than six months after completing the research, 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 total countries involved.

The ARC also commissions bibliometric studies of journal publications attributable to its funding (Butler 2004a),¹⁶ and measures used cover 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 least served, and that the coverage of the indexed database employed¹⁷ clearly favours SET over HASS, a point elaborated further in Section 2.4.

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).¹⁸ While largely discussing research ‘impact’ several areas fall into the remit of research ‘quality’ as defined in Section 1 above. The ARC report does not formally present measures or indicators of HCA research, but does examine major areas of concern for the HCA sector, identifies potential data sources (ARC applications, grants, end of grant reports), and although not discussed in these terms reveals likely novel indicators which may be supported through mining ARC databases. While the focus is upon HCA and its distinctive features, these indicators can readily be applied to the social sciences and to SET.

¹⁶ Again commissioned from REPP at ANU.

¹⁷ Comprising the Institute for Scientific Information’s (ISI) Science Citation Index, Social Science Citation Index (SSCI) and Arts and Humanities Citation Index (A&HCI).

¹⁸ 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.

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’ (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 audio-visual recordings which are simply not counted because they fall outside the core DEST block-funding publication categories (2004, 11). As for citation-based indicators which depend upon 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 a disproportionate level of high-impact international HCA journals are not indexed by ISI (2004, 21).

However, several potential indicators of HCA research quality are revealed in a survey of 1996 Large Research Grants and Research Fellowships (2004, 64-66), which are supported by existing ARC data. 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/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/examine postgraduate research theses, and invitations 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.

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 SET research.

The ARC commissioned a report, *A Wealth of Knowledge: The Return on Investment from ARC-Funded Research* (The Allen Consulting Group 2003a), which studied the returns on investment in ARC funded research. While the economic approach is an

important case to make, there are missed opportunities for HASS when the report addresses the ‘social rate of return’ of ARC research, as we find that this is in fact defined as ‘the permanent increase in GDP as a percentage of the dollar cost of the investment that led to this increase’ so is 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 be [*sic*] quantified in this study (2003a, 1), ‘but which are nevertheless valued by the community’ (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’ (2003a, 6). Due to the confines 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 Sections 4 and 4.5).

Humanities and creative arts

The ARC’s draft review of Humanities and Creative Arts does address research impact but does not discuss any measures. Yet the range of the discussion combined with the descriptions of data held by the ARC points the way towards developing future qualitative indicators of research impact. The report identifies international collaboration, and linkages with industry, government and community already outlined above. However, what is 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 (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 developing qualitative models of research impact informed by academic and ‘user’ perspectives, a point discussed in further detail in Sections 4 and 4.5.

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 financial year 2003-04 ‘NHMRC and Other Health’ received \$428.3m or 8.0% of the Commonwealth Science and Innovation budget. The NHMRC is dedicated to research, ethics and advice in health and medicine. Some HASS research will fall within its 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 (PMF)¹⁹ using the Department of Finance and Administration’s preferred model for performance measurement, the ‘outcome-output’ framework, which demonstrates what the NHMRC produces (‘outputs’) and the consequences (‘outcomes’) for the community. Various headings include ‘Creating new knowledge’, ‘Utilising knowledge’, ‘Enhancing capacity to innovate’, ‘Ensuring high ethical standards’ and ‘Strengthening communications and collaborations’, which are supported by a range of indicators discussed below in terms of 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, nor technical details discussed. Under the category of ‘Creating new knowledge’ lies the subcategory ‘Growth in internationally competitive knowledge’ which employs measures of the number of publications and citations per publication, presumably based upon journal-based output which allows for comparison with national and OECD averages. A further category ‘Enhancing capacity to innovate’ includes ‘Increased international recognition of Australian researchers’ which employs new measures, and data will be supplied and benchmarked in 2005 covering non-bibliometric indicators (awards, prizes and other honours, keynote addresses, presentations and speaking engagements).

¹⁹ See <http://www.nhmrc.gov.au/aboutus/pmf.htm>.

The NHMRC also commissions bibliometric studies of journal publications attributable to its funding (Butler 2003b), and 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 more 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 SET 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, 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; and in terms of '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 (this 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, and community awareness of issues addressed by the NHMRC and changes in community attitudes on issues addressed by the NHMRC.

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 within the category of 'Creating new knowledge' there is a heading entitled 'Research results with outstanding scientific impact or potential to impact highly on public health policy and practice', 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 adverse to the idea of qualitative benchmarking, currently being considered by its Research Committee on Measures of Research Impact and Achievement (MORIA) 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 Sections 4 and 4.5 below).

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

Conclusion

In terms of measuring research quality we find that all agencies collect data on the number of publications produced, and records do often extend beyond the core DEST categories, and 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 Humanities and Creative Arts, and in detail in Section 2.4 below – vital considerations if citation indicators are proposed to replace the current publications element in any future funding arrangements.

AIMS, ANSTO, CSIRO and 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 SET and HASS sectors.

Turning to 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 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 the CSIRO and NHMRC have taken with their ‘Outcome-Outputs’ models, which potentially combine quantitative and qualitative data with peer assessment, and present HASS with the opportunity to design similar models of research impact which can highlight what HASS shares in common with SET and what also makes HASS research distinctive and worthy of public financial support. The ARC’s report on research in the Humanities and Creative Arts lists in detail the impacts of its research upon 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 NHMRC. It is curious that it is SET which is most actively promoting its wider social and community benefits when it is HASS and its spectrum of activities which has most to gain by adopting this approach. The Humanities and Creative Arts are

a step closer towards this, but the social sciences remain silent to date, the very HASS stream which has 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 4.1, 4.3 and 4.5 below.

2.4 Measuring research quality: a closer look at citation indicators and HASS²⁰

We have seen that AIMS, ANSTO, ARC, CSIRO and NHMRC use journal-based citation indicators derived from The Institute for Scientific Information's (ISI) Science Citation Index, Social Science Citation Index (SSCI) and Arts and Humanities Citation Index (A&HCI) as a proxy for measuring the quality of publicly funded research which 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 citations to articles in a journal in a particular period). This review has observed that practice in evaluating SET 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 further using insights gained from the specialist bibliometrics literature on citation indicators.

It is first important to note that the bibliometrics literature insists 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 upon subsequent literature, although purely in internal academic terms (i.e. academic papers that are cited by other academic papers) and not the broader definition of 'impact' adopted by the RQF consultation process and hence used in this review.

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

²⁰ 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), <http://repp.anu.edu.au/Literature%20Review3.pdf>, from which Sections 2.4 and 4.4 of this report are derived.

research assessment tools in the natural and life sciences' (Luwel *et al.* 1999, 13). However, while Nederhof *et al.* (1989, 425) earlier acknowledge that 'bibliometric tools have proved their usefulness as monitors of developments in the natural and life sciences', they caution that 'evidence on this point is almost completely lacking for the humanities and many of the social science disciplines', concerns that persist today.

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 from Law (4%), Arts (11%), Political Sciences and Public Policy (23%), Historical Studies (40%), Population Studies (44%), Economics (55%), Sociology (61%) and Psychology (62%). Here HASS compares badly with Australian SET 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 SET coverage is generally high,²¹ we must question whether bibliometric approaches suited to measuring research output or research impact in the 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 where 'non-standard' publications (books, monographs, chapters and other non-ISI literature)²² 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' (2004, 473).²³

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

²¹ It is interesting to note that engineering is less well-served by journal-based indicators as this discipline relies heavily on technical papers which are not indexed by ISI.

²² 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.

²³ For further concerns about A&HCI and SSCI coverage see Nederhof and Zwaan (1991, 323-33), Nederhof and Noyons (1992, 255), Burnill and Tubby-Hille (1994, 148), Glänzel (1996, 292), Cronin *et al.* (1997, 264; 270) and Hicks (1999, 197; 210-2).

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, which 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' (1996, 54). Royle and Over make the strongest criticism relating to the social sciences in Australia when they argue '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' (1994, 86).

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 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 SET.

For example, if database coverage changes to include a wider range of publications, then HASS may be better served by citation-based indicators. In this respect, there is some interesting work in progress (Butler and Visser 2005) 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 4.3. There is also much scope for developing non-bibliometric quality indicators (also discussed further in Section 4.3), which are currently adopted or collated to some extent by AIMS, ANSTO, ARC, CSIRO and NHMRC. Bibliometricians would advocate that citation and other 'quality' measures should not be used in isolation, but rather selectively combined and applied to appropriate sub-fields of research within an overall context of expert peer judgement.

3 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. 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 research quality framework. 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 then 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 with, for example, France, Germany and the USA 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 UK several initiatives are in development, ranging from standard and non-standard bibliometric measures through to refining 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.

Measures of research quality

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

In the Netherlands, national assessment of research is carried out internally by universities, and is at present comparable but not rigorously comparative between departments. Self-assessment is carried out every three years, and external evaluation every six years. This exercise presently 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 socio-economic 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 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 UK it has been recommended that quantitative performance indicators should play a significant role in the 2008 Research Assessment Exercise (Roberts 2003), but while there is a list of metrics in the guidance given to panels for 2008 RAE, it does not mention bibliometrics or any other indicator of quality or impact (HEFCE 2005). To date, in terms of esteem measures departmental 'units of assessment' or UoAs 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 specify what 'metrics' are, but say that these are still in development. Some light is shed in a discussion on performance indicators: 'esteem indicators or bibliometric measures, might also be incorporated where they were more closely approximate to quality' (2003 §210), and it is recommended that the UK 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' (2003, p. 50). So while quality 'metrics' are a work in progress there is, nonetheless, sensitivity shown to HASS research with the acknowledgement 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' (2003 §212). The weight of such measures is expected to vary between subject areas, 'such variation a healthy reflection of the genuine differences between disciplines' (2003 §213).

The UK Arts and Humanities Research Council (AHRC) 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'. It has, however, in the face of outcry from its constituency 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 28 January 2005). The UK Economic and Social Research Council (ESRC) 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, e.g. citations to government reports.

In other countries we find that Flanders (Belgium) has produced the most concerted effort to introduce bibliometric methods for the purposes of 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, but is not including scientific publishing in a funding formula due to the lack of reliable data, but is producing a database for number and type of publication. Finland similarly 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.

Measures of research impact

Again it is New Zealand, the Netherlands and the UK 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 and reports commissioned by users under contract or formal arrangement, presentations where users are in attendance or 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', the number of these relationships are counted and a brief description of each is sought. Here, 'partnership' is defined as a formal relationship with another party (i.e. organisation or individual) with whom there is some form of alignment of objectives or strategy, or some close association; '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 UK 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). UoAs are free to provide information on relationships with research users, including industry and commerce. The UK AHRC 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 USA, 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 SET-based research only.

Conclusion

In terms of measuring research quality 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 any 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, highlight 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.

4 Potential measures of quality and impact of publicly funded HASS

The final section of this review 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 4.1 and 4.2 present lessons which we may glean from Australian and international practice. Section 4.3 introduces promising new ways of measuring research quality and impact being postulated which are not currently adopted by standard international or Australian practice. Section 4.4 takes a closer look at measuring research quality using non-bibliometric indicators, and Section 4.5 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 taken 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 can be very disadvantageous for HASS, particularly standard journal-based citation studies, 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 SET has currently stolen the march in this area.

It appears that for measuring HASS research quality, the most promising areas to develop are 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 SET, this is an area where 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 the sector is still unfortunately wedded.

4.1 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 PFRAs (AIMS, ANSTO and CSIRO) and the research councils (ARC and 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 drawn from existing extended university DEST returns (see Section 4.3 below). 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 honours, awards and prizes. A range of quality measures

therefore already exist, 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 the Research Quality Framework 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’ (2005, 20).

The above bodies employ indicators of research impact relating to commercialisation and technology transfer, and advice given to government; and there are various measures distributed between these bodies 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 NHMRC have taken the lead with their respective ‘Outcomes-Outputs’ 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 which can highlight what this sector shares in common with SET and what also 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 2003a), and this alternative strategy presents the opportunity to explore qualitative analyses and impact visualisation models already embraced by Australian SET. A more poignant lesson is that HASS has been unnecessarily conservative in its thinking about promoting the value of the sector’s research through restricting its thinking to presenting ‘impact’ in purely economic or commercial terms.

4.2 Lessons from international practice

In OECD countries we find that science policy founded on science principles provides the framework for regulating both SET and HASS research, and this policy resonates with the terms ‘utility’, ‘commercialization’ and ‘wealth creation’. The idea is that governments believe that scientific discovery creates social and economic progress, and wish to harness research towards 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 where academics self-regulate their activities. A consequence of such wealth-maximising, SET-dominated and indicator-led thinking is that the distinctive qualities and benefits of HASS are likely to be lost within standardised evaluation

systems. This has indeed been the case in several countries where HASS is unreflexively judged by SET-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 UK where it is hoped that these ‘metrics’ will feed into the 2008 RAE.

Australia is in the 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 sensitive to a variety of disciplines and fields, and which are tailored to identifying and stimulating the best research outcomes Australia can offer in SET and HASS alike.

4.3 New ways of measuring quality and impact currently being postulated

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 2004, 4-6) and suggests an internal university points allocation system for various musical creative outputs, ranging from published musical scores, commercially published or broadcast original compositions, 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, 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.

There is current work in progress mentioned briefly above in Section 2.4 which investigates non-standard citation analyses that promise to be more HASS-friendly than standard ISI journal-based citation studies (Butler and Visser 2005). This research uses existing information provided by universities in their DEST returns, which often extends beyond the four main categories, and mines the Web of Science for citations to books, book chapters, non-ISI journals and conference proceedings. This approach has been

applied to some SET areas and has so far been extended to economics, history and philosophy, although more testing is required.

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 having SET-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 (The Allen Consulting Group 2005). While this report is largely SET-focused HASS is mentioned, particularly the social sciences. Impacts are divided into four main societal values: material, human (largely health and well-being related), 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’ and indicators include ‘marriage and divorce rates, persons living alone, participation in voluntary work’ (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 ABS 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 interesting in occasionally hedonistic ways. For example, that 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) (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 therefore raising community awareness in the importance of protecting certain legal and political rights (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 head 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 (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, like the CSIRO 'Outcomes-Outputs' model HASS may plunder this approach and adapt it to its purposes.

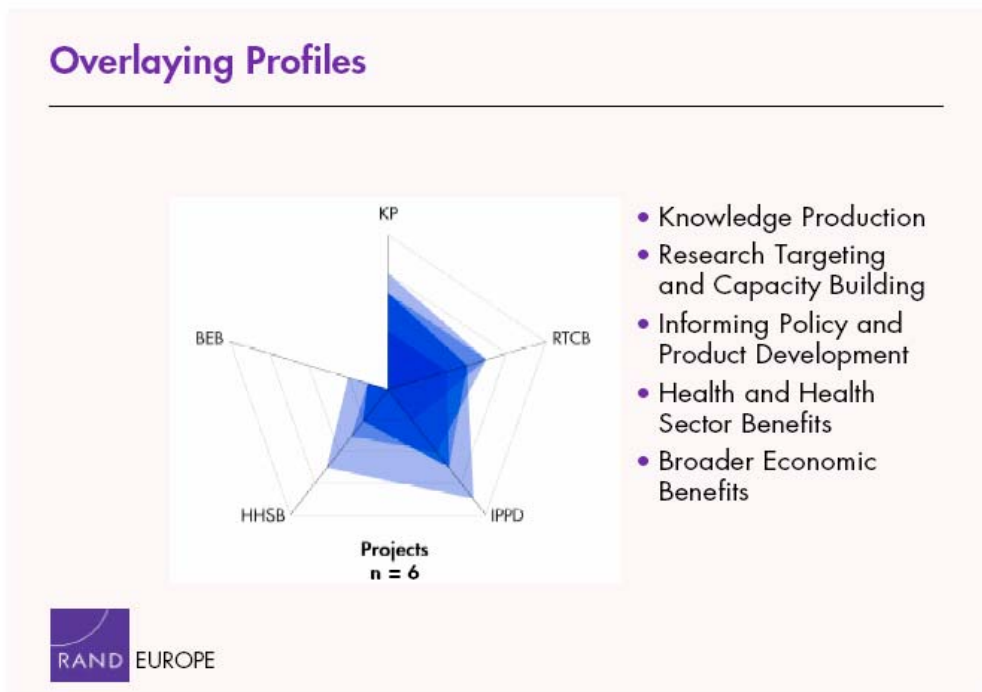
The attraction of this approach is that measures of research quality and impact are combined with peer and user evaluation to produce a logic model of the research process (see Chart 2). Category A in Table 1, 'Knowledge' includes bibliometric analyses of publications resulting from research funding, 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, and this is then mapped resulting in various profiles like that illustrated in Chart 2. The other categories are 'Benefits to Future Research and research Users' (which is largely forward looking and capacity building in nature), '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).²⁴

²⁴ 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 judgements about research group activities.

Table 1: Payback categories for arthritis research²⁵

A.	Knowledge
B.	Benefits to future research and research use: <ol style="list-style-type: none"> 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.
C.	Political and administrative benefits: <ol style="list-style-type: none"> i. improved information bases on which to take political and executive decisions; ii. other political benefits from undertaking research.
D.	Health sector benefits: <ol style="list-style-type: none"> i. cost reduction in the delivery of existing services; ii. qualitative improvements in the process of service delivery; iii. increased effectiveness of services, eg increased health; iv. equity, eg improved allocation of resources at an area level, better targeting and accessibility; v. revenues gained from intellectual property rights.
E.	Broader economic benefits: <ol style="list-style-type: none"> 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.

Figure 2: Example of 'payback profile' impact assessment²⁶



²⁵ Reproduced with permission of the HERG/RAND group.

²⁶ Reproduced with permission of the HERG/RAND group.

Like the CSIRO ‘Outcomes-Outputs’ approach discussed in Section 2.3.1 above, we can imagine similar logic models being developed for HASS research, and a particular attraction is the emphasis upon research impacts which can be tailored to HASS stream and discipline. This approach would also appeal to the RQF desire for peer and user-informed research assessment.²⁷ The potential of this approach for evaluating HASS research is discussed further in Section 4.5 below.

4.4 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 being indicators of esteem such as awards, honours and prizes. Such qualitative indicators of esteem are recognised as particularly important for those parts of HASS where 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 yet comprehensive study of law and linguistics in four Flemish universities by Luwel *et al.* (1999) 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’ (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.²⁸

The mainstream literature tends to discuss non-bibliometric indicators in terms of esteem indicators and the social sciences and humanities, and non-written outputs and the creative arts have generally been overlooked. Yet due to 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 noted that publication in the creative arts includes:

- *public performance* for dancers, actors, choreographers, musicians, playwrights;
- *building or manufacture* for architects and designers;

²⁷ 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

²⁸ 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>

- *written works* for academics, researchers and creative writers;
- *exhibitions* for visual artists, craftspeople; and
- *computer software* for designers, musicians, visual artists and so on (1998, 55)

a list which again can be extended and refined but serves to give a flavour of the potential for using non-bibliometric indicators.

A novel take on non-bibliometric indicators is suggested by a Royal Society of Engineering Report which under the heading of ‘Scholarship’ employs what it terms ‘independence indicators’ which denote research quality to the extent that ‘it is free from financial and political control’ (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’ (2000, 28). Non-bibliometric independence indicators can be applied equally to SET and HASS research.

Non-bibliometric indicators may also be applied to investigating policy impact. For example, Cave *et al.* produced a report for the UK ESRC which discussed potential performance indicators for ESRC-funded research. They examined, but did not put into practice, the possibility of evaluating the relevance of social science research by using ‘policy usefulness’ as a novel non-bibliometric performance indicator. 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’ (§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’ (1988, iii), although 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.

4.5 Measuring research impact: a closer look at qualitative modelling

Sections 2.3.1 and 4.3 of this review describe models used by SET and the biomedical sciences to assess research impact, namely the CSIROs ‘Outcomes-Outputs’ 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 specialized 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 HASS' 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) identify a range of interesting research benefits, but due to the confines of the study only quantify these in financial terms, although 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 common threads emerging such as knowledge creation, educational benefits and contributions to policymaking.

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 (2004 15; 145-152). The following areas emerge, 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',²⁹
- 'Understanding other societies and cultures'
- 'Informed public debate, improved policy-making and economic impact'
- Knowledge/education
- Cultural/economic

Similarly, The British Academy (2004, vii-ix) examines in detail five key functions of the arts, humanities and social sciences:

- Contributing to cultural and intellectual enrichment
- Contributing to economic prosperity and well-being
- Contributing new knowledge and understanding of major challenges facing both the UK and beyond
- 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 SET 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 SET and HASS share, and which are also suited to HASS, and its streams and disciplines. For research evaluation purposes we may for once adapt SET to the use of HASS rather than HASS to the use of SET.

²⁹ These areas are defined in detail in Section 2.3.2 above.

References

Association of the Universities in the Netherlands (VSNU), Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NOW) and Koninklijke Nederlandse Akademie van Wetenschappen (KNAW). 2003. *Standard Evaluation Protocol 2003-2009 for Public Research Organisations*, January 2003.

Australian Institute of Marine Science (AIMS). 2003. *Research Plan 2003-2006*.

Australian Nuclear Science and Technology Organisation (ANSTO). *Annual Report 2003-04*.

Australian Research Council (ARC). 2004. Draft of *Review of the Humanities and Creative Arts Discipline Grouping 2004*, dated 14 October 2004.

Australian Research Council (ARC). 2004. *Annual Report 2003-04*.

Bourke, P., L. Butler and B. Biglia 1996. *Monitoring Research in the Periphery: Australia and the ISI Indices*, REPP Monograph Series No. 3, Research School of Social Sciences, ANU.

Burnhill, P. and M. E. Tubby-Hille. 1994. 'On Measuring the Relation Between Social Science Research Activity and Research Publication,' *Research Evaluation* 4(3): 130-152.

Butler, L. and M. Visser. 2005. 'Extending citation analysis to non-source items,' *Scientometrics*, submitted.

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 H. F. Moed, W. Glänzel and U. Schmoch (eds), *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems*. Dordrecht/Boston/London: Kluwer Academic Publishers.

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.

Cave, M., M. Kogan, S. Hanney and D. Burningham. 1988. *Performance Indicators for the Economic and Social Research Council: Report on Preliminary Project* prepared for UK Economic and Social Research Council.

Commonwealth Scientific and Industrial Research Organisation (CSIRO). 2004. *Annual Report 2003-04*.

Cronin, B., H. Snyder and H. Atkins. 1997. 'Comparative Citation Rankings of Authors in Monographic and Journal Literature: A Study of Sociology,' *Journal of Documentation* 53(3): 263-273.

DEST. 2004. *Evaluation of Knowledge and Innovation Reforms Consultation Report*. Canberra: Department of Education, Science and Training.

DEST (year?) *Science and Innovation Budget Tables 2004-05*, http://www.dest.gov.au/science/analysis/pdf/Science%20and%20Innovation%20Budget%20Tables%201%20to%205_2004-05.pdf

Donovan, C. 2005. 'The Governance of Social Science and Everyday Epistemology,' *Public Administration* 83(3): 597-615.

Foundation for Research, Science & Technology (FRST). 2004. *Outcome Indicators 2003/04: Guidelines for Data Collation by Research Providers*. 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 U. Schoepflin. 1999. 'A Bibliometric Study of Reference Literature in the Sciences and Social Sciences,' *Information Processing and Management* 35: 31-44.

Higher Education Funding Council for England (HEFCE). 2005. *RAE 2008: Guidance to Panels*, January 2005.

Hicks, D. 2004. 'The Four Literatures of Social Science,' in H. Moed, W Glänzel and U. Schmoch (eds) *Handbook of Quantitative Social Science and Technology Research*. Dordrecht/Boston/ London: Kluwer Academic Publishers.

Hicks, D., H. Tomizawa, Y. Saitoh and S. Kobayashi. 2004. 'Bibliometric techniques in the Evaluation of Federally Funded Research in the United States,' *Research Evaluation* 13(2): 78-86.

Hicks, D. 1999. 'The Difficulty of Achieving Full Coverage of International Social Science Literature and the Bibliometric Consequences,' *Scientometrics* 44(2): 193-215.

International Federation of Arts Councils and Culture Agencies (IFACCA). 2004. *Statistical Indicators for Arts Policy: Discussion Paper*, July 2004.

Katz, J. S. 1999. *Bibliometric Indicators and the Social Sciences* prepared for UK Economic and Social Research Council.

Luwel, M., H. F. Moed, A.J. Nederhof, V. De Samblanx, K. Verbrugghen, L.J. van der Wurff, 1999. *Towards Indicators of Research Performance in the Social Sciences and Humanities*. CWTS: Leiden.

Martin, Ben R. 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.

National Health and Medical research Council (NHMRC). 2004. *The NHMRC's Performance Management Framework 2003-2006*.
<http://www.nhmrc.gov.au/aboutus/pmf.htm>

Nederhof, A. J. and E. C. M. Noyons, 1992. 'International Comparison of Departments' Research performance in the Humanities,' *Journal of the American Society for Information Science* 43(3): 249-256.

Nederhof, A. J. and R. A. Zwaan. 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, A. J., R. A. Zwaan, R. E. De Bruin and P. J. Dekker. 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.

Research Evaluation and Policy Project (REPP). 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>

Department of Education, Science and Training (DEST). 2005. *Research Quality Framework: Assessing the Quality and Impact of Research in Australia – Issues Paper*, March 2005.

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), May 2003.

Royal Academy of Engineering. 2000. *Measuring Excellence in Engineering Research*, January 2000.

Royle, P and Over, R. 1994. 'The Use of Bibliometric Indicators to Measure the Research Productivity of Australian Academies,' *Australian Academic and Research Libraries* 25(2): 77-88.

Strand, D. 1998. *Research in the Creative Arts*. Canberra: Department of Employment, Education Training and Youth Affairs.

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

The 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.

The 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.

The Australian National University. 2004a. *ANU Capabilities & Performance Statement*, Report for the Committee of Review of ANU Quality, July 2004.

The 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.

The 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.

Times Higher Education Supplement (2005). *Journals 'Top Ten' Sparks a Rebellion*, Phil Baty, 28 January 2005.

Toutkoushian, R. K., S. R. Porter, C. Danielson, and P. R. Hollis. 2003. 'Using Publications Counts To Measure An Institution's Research Productivity,' *Research In Higher Education* 44: 121-148.

Van Raan, A. F. J. 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., K. Debackere, M. Luwel and E. Zimmermann. 2002. 'Measuring Progress and Evolution in Science and Technology – 1: The Multiple Uses of Bibliometric Indicators', *International Journal of Management Reviews* 4(2): 179-211.

Von Tunzelmann, N. and E. Kraemer Mbula. 2003. *Changes in Research Assessment Practices in Other Countries Since 1999*, Final Report to the Higher Education Funding Council for England, 28 February 2003.

Webber, P. 1994. 'Creative Arts as Research', in C. Cusack (ed.) *Creative Arts and Research*, Centre of Creative Arts: University of Sydney.

Wissler, R. 2004. 'Research Outputs – Up For Auction', in R. Wissler, B. Haseman, S. Wallace and M. Keane (eds) *Innovation in Australian Arts, Media and Design: Fresh Challenges for the Tertiary Sector*. Flaxton: Post Pressed.

Wooding, S., S. Hanney, M. Buxton and J. Grant. 2004. *The Returns from Arthritis Research Volume 1: Approach, Analysis and Recommendations*, RAND Europe.
http://www.rand.org/pubs/monographs/2004/RAND_MG251.pdf