



## AAH Policy

### AUSTRALIAN RESEARCH COUNCIL ERA EI REVIEW CONSULTATION PAPER OCTOBER 2020

The [Australian Academy of the Humanities](http://www.humanities.org.au) (AAH) welcomes the opportunity to respond to the Australian Research Council (ARC) consultation on the Excellence in Research for Australia (ERA) and Engagement and Impact (EI) exercises. The Academy's submission responds to select consultation questions and focuses on priority and systemic issues we think need to be addressed. We would be pleased to elaborate on this submission and convene further expert input.

Key priorities from the Academy's standpoint are as follows:

- > From the Academy's point of view, the value of ERA is that it assesses research quality, not the volume or quantity of research.
- > Into the future, the Academy would be concerned if ERA results are canvassed as a mechanism to disperse research funding to universities without requisite impact modelling being undertaken. We would also want to ensure that there is a review mechanism in place as a check against perverse consequences.
- > In ERA the growing divergence in 'performance' of peer review and citations disciplines has skewed the ratings and needs to be addressed as a matter of urgency. We think there is a need for 'normalisation' between peer review and citation results.
- > ERA outcomes and data are open to misinterpretation by media and universities. Common misinterpretations include using averages to compare across disciplines or making direct comparisons between peer review and citation disciplines. The ERA Report, as currently structured, does invite such comparisons. One way of mitigating this would be to commission authoritative commentary from the ERA Research Evaluation Committee (REC) chairs to accompany the release of results.
- > ERA is still not set up well to cater to interdisciplinary and emerging research.
- > In the EI exercise, the current income-based engagement metrics are insufficient, and only fit-for-purpose for select disciplines. We propose additional engagement metrics for humanities and arts fields. This will incentivise universities to collect data of significance on the most substantial patterns of community engagement.
- > Peer review is a fundamental element of both exercises. The frequency of the exercises has stretched the limits of the peer review system. We support ERA and EI at five-yearly intervals.
- > We support increased transparency and the release of data from both exercises so that robust analysis can inform policy and institutional decision-making.

AUSTRALIAN ACADEMY OF THE HUMANITIES

## 1.1 Operating context

The Consultation Paper asks for stakeholder views on how ERA and EI may need to be modified in light of the following current and recent reviews: [Research Sustainability working group](#); House of Representatives Review of Australian Government Funding Arrangements for non-NHMRC Research (2018); [Coaldrake Review of Higher Education Provider Category Standards](#) (2018-2019); and the [Australian and New Zealand Standard Research Classification \(ANZSRC\) Review](#) (2020).

At this point in the process, the Academy would be concerned if ERA results are canvassed as a mechanism to disperse research funding to universities without requisite impact modelling being undertaken. We would also want to ensure that there is a review mechanism in place as a check against perverse consequences. The Academy will develop up a more formal position when the outcomes of the Research Sustainability working group are public.

To date, ERA has achieved its ends and changed behaviour without having funding tied to it. From the Academy's point of view, the value of ERA is that it assesses research quality, not the volume or quantity of research. We would want to be assured that ERA is robust enough to be used in this way, especially given flaws with methodology that have emerged over time. We suggest to the ARC further targeted consultation into next year on the consequences of these reviews for ERA and EI.

Below we also raise some preliminary issues we have with the new ANZSRC field of research classifications and the composition of assessment panels.

## 1.2 'Excellence in Research for Australia (ERA)

### 1.2.1. Value and impact of ERA

ERA has a positive impact by enabling critical review of research strengths within and across individual institutions which has created the opportunity for research synergies by drawing out these pockets of research regardless of where they are situated organisationally – in a faculty, a research centre or institute, or a Centre of Excellence. The whole-of-institution assessment is a strength, allowing a more holistic view of institutional research.

ERA has forced a higher level of international engagement and it has reinforced the significance of research as part of an individual's performance, which has in the long run resulted in a higher level of research performance across the board (not necessarily more 5s but more 3s and 4s). To that end, we think that ERA has already largely met its objectives and effected cultural change within the university sector.

ERA is an invaluable data gathering exercise, but the ARC should go much further in making its data available to enable independent analysis and review, as we explain below.

An ongoing issue for ERA in need of redress is that universities have 'gamed' the assessment process, which has resulted in a shrinking of the support base for the full range of academic disciplines operating in the sector. In 2018, for instance, there was only one submission in Languages. As areas of strength consolidate, less comprehensive or even emerging areas have lost the support of their university and are in danger of disappearing entirely. In the humanities, structural diversity is important to the health of the system and we would not want to see a contraction of research to the Group of Eight universities. This would not be a good outcome for the nation.

### 1.2.2. How we use ERA outcomes

The Academy uses ERA data for policy and research purposes. Successive ERA reports have enabled longitudinal analysis of the humanities, arts and social sciences (the SHAPE disciplines)<sup>1</sup> against indicators of capacity, including investment, staffing, and ERA ratings. The Academy's [Mapping the Humanities, Arts and Social Sciences in Australia](#) (2014) report worked extensively with ERA data for a first-ever mapping of this research ecosystem in Australia.

### 1.2.3. Suggestions for improved reporting

ERA outcomes and data are open to misinterpretation by media and universities. One of the most common mistakes is using averages to compare across disciplines; another is comparing results in peer review and citation disciplines. These comparisons should not be made, but the ERA Report as currently structured, does invite such comparisons.

One way of mitigating this would be to commission authoritative commentary from the ERA Research Evaluation Committee (REC) chairs to accompany the release of results. The chairs are best placed to give a high-level overview of the results within domain areas, and contextualise results as relevant with respect to different funding sources, staffing profiles, different infrastructure, and publishing patterns. We would be mindful of adding to the workload on the chairs, and this would not solve the problem outright, but could contribute to improved communication and interpretation of results. This is done in the UK's Research Excellence Framework (REF) exercise – chairs do it by discipline and members of the panel do it by unit of evaluation.

We would also observe that ERA is limited by its FoR-based units of analysis – these are a limited lens through which to view areas of research not visible within the existing classification system. In the humanities, such areas include digital work, environmental and interdisciplinary studies, gender studies, and Indigenous studies (some of which have been included in the new ANZSRC FoR classifications).

To develop a more meaningful picture of humanities research and an account of the extent of research on important social and cultural challenges, we would suggest topic modelling as an approach to improve reporting. For example, in the Academy's *Mapping the Humanities, Arts and Social Sciences in Australia* report, we were able to work with the ARC to conduct keyword searches across the ERA publication dataset to map output of health and Asia-related priorities.

### 1.2.4. ERA methodology

Over time the growing divergence in 'performance' of peer review and citations disciplines has skewed overall results and needs to be addressed as a matter of urgency.

Since the patterns of the results coming out of citation disciplines are so different to those coming out of peer review, there is a genuine incommensurability in the respective set of outcomes. Citation disciplines skew high and peer review disciplines skew low. To some extent, this may be the consequence of different thresholds or interpretations of what constitutes 'world standard'.

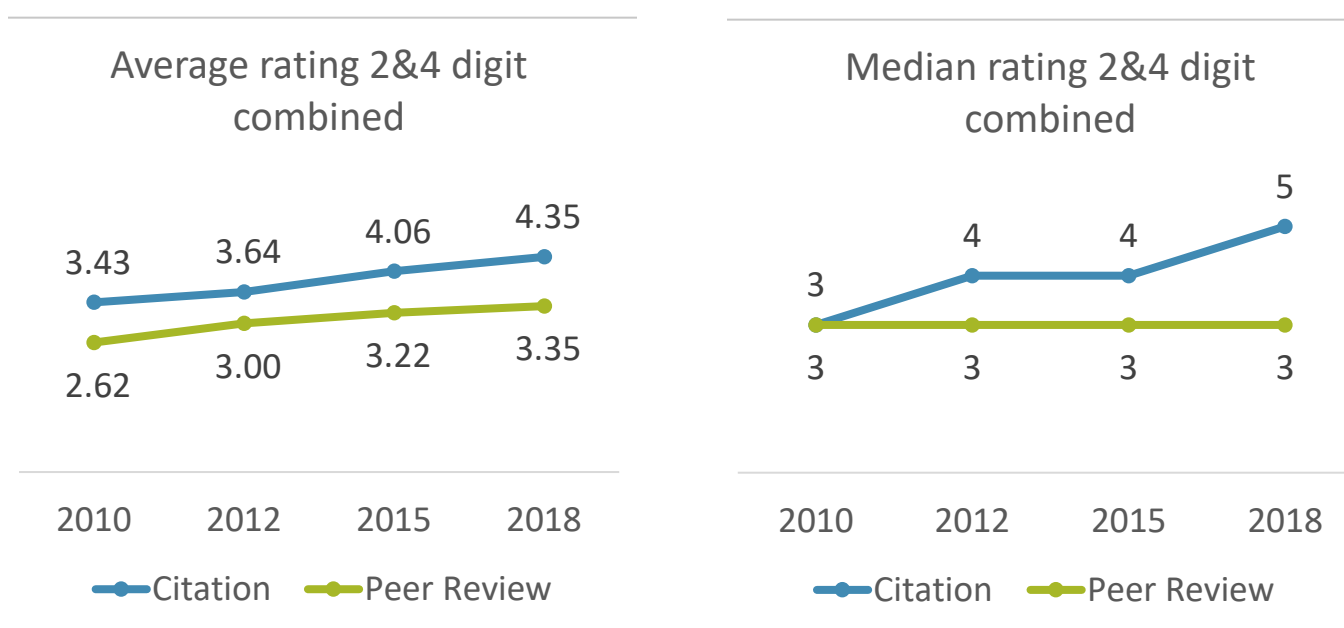
---

<sup>1</sup> 'SHAPE' stands for Social Sciences, Humanities and the Arts for People and the Economy, <https://thisishape.org.uk/>

Analysis by Jo Dalvean and Professor Mark Sanderson from RMIT University<sup>2</sup>, shared with the Academy, shows that in the 2018 ERA exercise, 59 per cent of STEM 4-digit submissions were rated a '5' compared to 15 per cent of SHAPE/ICT fields. Average ratings across Australia for the citation and peer review codes in 2018 were as follows (NB: omitting 01 and 10 because they are mixed citations and peer review):

- > The average score for the citation codes (02, 03, 04, 05, 06, 07, 09, 11, 17) was 4.17
- > The average score for the peer review codes (08, 12, 13, 14, 15, 16, 18, 19, 20, 21, 22) was 3.16

This divergence has widened over time, as shown in the following graphs produced by the RMIT team:



Source: Jo Dalvean, RMIT Research & Innovation, Presentation to ARMS Directors SIG, May 2019

More care needs to be given to the comparisons between the STEM and SHAPE/ peer review discipline results. The Academy suggests a normalisation process to bring them into better alignment, such as a multiplier of the average citation vs peer review rating.

A pilot project to test the use of citation analysis for the peer review disciplines could highlight potentials and problems and might provide the basis for considering alternative models. This would have to be carefully thought through.

<sup>2</sup> Jo Dalvean, Research and Innovation, RMIT University; and Professor Mark Sanderson, Professor of Information Retrieval and Director of the ISE Enabling Capability Platform, Computer Science, School of Science, RMIT University.

#### **1.2.5. Peer review**

The pool of available assessors within a relatively small research system is not inexhaustible. This is a critical issue for the fields of research where labour-intensive peer review, as opposed to citation analysis, is used to measure quality. These include the SHAPE disciplines, as well as a number of STEM fields such as pure mathematics. We are concerned that this resource is being stretched beyond capacity, with risks to the quality and credibility of the system. It is important that we maintain the quality of the peer reviewers and that universities put forward their best and most experienced staff members. Extending the time between ERA exercises will go some way to mitigating these impacts.

We also suggest more coordination between the ARC and the SHAPE sector on planning, recruitment, training and support for the peer review process. Feedback we have received as part of this consultation suggests that improvements could be made to how assessors are recruited and prepared.

#### **1.2.6. Contextual indicators**

The Academy is supportive of the publication of output volume information for all assessed Units of Evaluation (UoEs). Having said that, we would observe that counts of research outputs are crude measures that do not always reflect quality; nor does the volume of publications produced in a given institution necessarily reflect that institution's commitment to research.

For instance, while the ERA National Report is a very useful resource, the Academy believes that volume information could be better contextualised in ERA reporting to avoid distorting the picture due to differences in publication practice by disciplines. As it currently stands, regardless of the type of research output – book, journal article or conference paper – they are each ascribed the same value for this purpose as they are all counted as '1' output. Books are only weighted for determining low-volume threshold. The effect of this is evident in the ERA National Report where calculations are made about a discipline's share of national research output, which can under-state the real effort in disciplines such as History where 10% of its research outputs are books.

While noting that there are sensitivities about releasing some data publicly, we would point out that volume data is only part of the picture – research income data and staff (FTE) data disaggregated by UoE would give a more accurate indication of institutional investment and resourcing of research.

#### **1.2.7. Applied measures**

The applied measures as currently configured are of no use for the Humanities and Creative Arts (HCA) panel. While there is a commitment to tailoring this process to the disciplines, and not going with a one size fits all approach, the dashboard does need some further tailoring to provide more useful information to the HCA panel. We note that some peer review disciplines could benefit from a hybrid approach where the REC is provided with RCI data like the citation-based disciplines, for example Archaeology.

#### **1.2.8. ERA rating scale**

The rating scale is now effectively a 4-point scale rather than a 5-point scale. There are almost no units scoring a 1. The difficulty for the committees often occurs at the cusp of the 4 and 5 rating. It might be worth introducing a higher level on top of the existing 5 – which places the

unit at ‘well above world standard’. A category of ‘world leading’ (so, a 6) would be a way of distinguishing between those who are indeed operating well above the world standard and those who have gone beyond that and are driving progress and innovation in their fields (although the latter would be hard to see in the current design of ERA, where it is not at the individual or unit level, but by discipline).

#### **1.2.9. ERA interdisciplinary research and new topics**

ERA is still not set up well to cater to interdisciplinary and emerging research because it defaults to existing research areas.

#### **1.2.10. ERA and Indigenous research**

The Academy is concerned about how the next ERA will be set up to assess Indigenous research under the new 2-digit 45 code. This code involves a vast array of disciplines; dominated by health and medical research, it also includes community-based archaeology and history. Will there be a mixed peer review/citation panel, like pure maths?

Universities will also have to grapple with institutional decisions about how best to code research in Indigenous studies, i.e. whether, for example in archaeology, it is preferable to code everything to 4501 Aboriginal and Torres Strait Islander culture, language and history or retain in 4301 Archaeology.

#### **1.2.11. Collection of ERA data**

Much as ERA has been onerous as a reporting exercise, we would be concerned if there were attempts to streamlining or automate data collection at the expense of rigour.

Many universities have their own repositories of publication data and it would be good for some time to be spent harmonising the structure and content of these repositories with the demands of ERA submissions. It would limit the amount of time preparing the submissions and it could also limit some of the game playing that goes on around the allocation of FoR codes to specific pieces of research. Annual reporting of publication data might reduce gaming, depending what regulations were placed around coding changes.

#### **1.2.12. Publication of ERA data**

As above, we would note that quality not quantity matters, so volume needs to be contextualised. We would also urge more transparent data reporting and analysis.

### **1.3 Engagement and Impact Assessment**

#### **1.3.1. EI overview**

The first EI exercise has been a pilot effectively. Some of the issues raised with the Academy to date include: inscrutability of some of the decision making around the case studies and, as a result, the lack of consistency and rationale for some of the assessments made; issues around the guidelines for some of aspects of the process especially with regards to pathway to impact.

#### **1.3.2. Use of EI outcomes**

As with ERA, the Academy uses data generated through EI for its policy and research work in support of the humanities. The published EI case studies are a rich but under-utilised resource for understanding the nature, scope and reach of humanities research impact.

### **1.3.3. Engagement indicators**

The engagement indicators are totally inadequate as a means of capturing the engagement activities in the humanities, creative arts and, to a lesser extent, social sciences. Using income of various kinds as a proxy for engagement has very limited empirical justification, and at best it only has some relevance for some of the STEM disciplines. The Department of Education's recently released [consultation paper on the National Priorities and Industry Linkage Fund \(NPILF\)](#), itself recognises that care needs to be taken to 'prevent use of simplistic metrics which reinforce outdated engagement modes such as number of commercialisations' (p. 10).

There are no indicators which are designed to pick up engagement activity in the SHAPE disciplines and so that activity will not be captured. At present there is an inadequate attempt at identifying some of the most substantial patterns of community engagement. Public-facing activities in the arts, for instance – exhibitions and performances – and official engagement in government policy instrumentalities are among the things which might be considered for the future.

We would be pleased to convene further expert input on the development of appropriate indicators.

### **1.3.4. Co-supervision of HDR students**

This could be positive for HCA fields, including for example with galleries, museums, archives and libraries (GLAM) sector and other organisations in the cultural and creative industries.

### **1.3.5. Impact narrative**

The narrative approach is suitable for describing and assessing impact.

## **1.4 Overarching issues common to both ERA and EI**

### **1.4.1. Frequency of ERA and EI**

We suggest every five years. The three-year cycle is too short both for administrative reasons but also substantive ones (we do not learn much over such short periods) and so the recent extension of time between assessments is very welcome and should be retained for future rounds.

### **1.4.2. Streamlining and simplifying ERA and EI**

While ERA and EI measure separate things, we see value in combining the exercises and staging every five years. The resourcing implications for universities has been onerous and there has been pressure on peer review disciplines. Combining the process for ERA and EI could be a more efficient use of peer reviewers' time.

### **1.4.3. Utilising technological advances and pre-existing data sources**

We agree with the introduction of ORCID iDs, though note that ORCIDs are not universally used. Making them mandatory would require effective sector communication and education. For HCA, we would need to ensure that non-traditional research outputs (NTRs) are recognised to require different reporting options.