

ERA 2023 BENCHMARKING AND RATING SCALE CONSULTATION SUBMISSION

April 2022

Universities Australia (UA) welcomes the opportunity to make a submission to the Australian Research Council (ARC) consultation paper on the ERA 2023 benchmarking and rating scale.

UA is the peak body for Australia's 39 comprehensive universities. Our member universities are spread across Australia in both regional and metropolitan areas. They educate more than a million students each year, undertake all of the university research in Australia and engage globally to add to this country's stock of knowledge, as well as Australia's economic and social wellbeing.

RECOMMENDATIONS

Universities Australia recommends that the ARC:

- extend the consultation period and then pilot the approach to arrive at a measurement system that achieves the policy intent of robustness and granularity;
 - Option A could be used as a starting point for further development and piloting
- work with UA to develop a solution to the inconsistencies resulting between the peer review and citation approaches as a result of methodological differences; and
- consider how to ensure that the following research is not disadvantaged as a result of these reforms:
 - Indigenous research
 - National security research
 - Commercially oriented research

EXECUTIVE SUMMARY

The issues outlined in this consultation are complex and require considerable modelling, testing and refinement. When ERA was introduced, a piloted approach was employed which allowed for methodological challenges to be worked through and unintended consequences to be identified (notwithstanding existing issues).

UA notes that further work is required on the Options, to increase the precision in terminology and to better address the differences in methodology between peer review and citation fields.

UA strongly recommends that a piloted approach is adopted. UA sees Option A as the better starting point for this further development and piloting. Option A provides a higher chance of achieving robustness because it differentiates with five levels rather than six and is less of a change with previous ERA rounds (which may provide potential for at least some consistency of interpretation against previous results).

UA stands ready to work with the ARC to allow a co-design of a system which achieves the policy intent, is robust and meets the obligations of the ERA exercise. The current timeline does not afford sufficient time to achieve this and places at risk the reputation of Australia's university research system.

INTRODUCTION

The government is seeking to revise the ERA benchmarks and rating scale. Its stated aim is to account for changes in the global research environment and to provide greater granularity to enable differentiation between institutions and identification of areas of research strength.

The objectives of ERA are:¹

1. **Promoting Excellence:** Rigorously assess research quality to promote pursuit of excellence across all fields and all types of research.
2. **Informing Decisions:** Provide a rich and robust source of information on university research excellence and activity to inform and support the needs of university, industry, government and community stakeholders.
3. **Demonstrating Quality:** Provide government and the public with evidence of the quality of research produced by Australia's universities.
4. **Enabling Comparisons:** Allow for comparisons between Australian universities and against world standard for all discipline areas.

Australia's university research system has been demonstrated to be amongst the best in the world, through the ERA exercise, verified through results in a number of international ranking systems. As ERA is an internationally recognised measurement system of research excellence, the reported quality of Australia's university research has enhanced the global competitiveness of our universities, allowing greater talent attraction and investment, which in turn have led to better research and flow-on benefits to all Australians.

It is essential that any reforms to the measurement system of research quality through ERA is a demonstrated enhancement of the current system; is internationally credible; and provides sufficient information to policy makers, the public and institutions to allow for continuous improvement of the research system. These principles are well encapsulated in the current ERA objectives.

¹ <https://www.arc.gov.au/excellence-research-australia>

A ROBUST, RIGOROUS AND GRANULAR MEASUREMENT SYSTEM THAT ALLOWS THE ASSESSMENT OF RESEARCH EXCELLENCE AND STRENGTHS

Pilot

UA supports a well-tested, robust measurement system that allows differentiation and identification of research strengths beyond the current approach. In making any changes to ERA, however, it is important to be aware of the ways in which results are used by government, researchers and universities, industry and those observing from overseas.

Piloting changes, either through conducting subsets of citation and peer review disciplines alongside the existing methodology in 2023 or a more comprehensive pilot, would allow a comparison of outcomes and an indication of the robustness of the changes being proposed. Given the already significant changes in ERA 2023, such as the changes in Field of Research (FoR) codes and the introduction of the Indigenous Studies FoR Code 45, it would be difficult to identify what has caused change if we do both sets of changes at once.

Definitions

Both options A and B outlined in the discussion paper seek to provide granularity through differentiation at the top of the scale. It will be difficult to truly differentiate at three levels at the highest level. Currently, the definitions are conceptually vague and would potentially produce less information than is currently the case. It will be crucial before implementing changes to ERA to ensure that definitions are tight by refining the language used in them and that there is sufficient differentiation and definition to allow them to be applied consistently. Feedback from assessors and panels on how results have been reached would be useful to understand how the benchmarks and rating scales are being applied.

More granularity may not yield better information

For some fields, top citations can lead to some very small institutions being identified as a global leader. A program such as ERA, given the importance of the outcomes, needs to consider how to manage 'outliers' in terms of size (ie understanding how very large and very small groups or institutions can be accurately placed on the same rating scale).

It is quite possible that for a large number of fields of research, no Australian university would meet the standard required to be in the top ten per cent in the world (option B, AAA ranking); or option A (world leading). In particular, the AAA rating is highly sensitive to just a few papers. Even the 'well above world standard' (option A) and AA (option B) ratings suffer from this issue.

The current ERA model includes a metric for income, which assists in understanding the result on investment in research. If similar metrics are available in the new system, then that would be helpful in calibrating return on investment.

It is unclear how a 'high performing institution' in a global sense would be defined, particularly the selection criteria and sources of information that might be appropriate for making these decisions and how cultural biases and assumptions might be taken into account in the selection of comparator institutions. In making these assessments, the process will need to ensure that assessment is against the content of all the research presented for evaluation and comparison purposes (noting the potential, though, to increase workloads on assessors).

Aim to recognise strength and depth, not only scale

The focus should be research strength and depth, rather than to try and fractionate the upper levels. The ERA should be able to identify and differentiate between a smaller group of researchers who are all performing at a high level and a larger group of researchers of whom a small percentage are performing at that high level.

ERA should be encouraging excellence at a depth and/or scale that produces rich research environments.

Mind the impact on early career researchers, emerging fields and interdisciplinary fields

Consolidating the lower categories of standard provides room for more granularity at the top of the scale, which is a helpful development. However, concerns have been raised about how the implications for early career researchers, emerging fields and interdisciplinary fields have been taken into account in making this change. UA would be interested in any modelling that ARC has undertaken to understand the impact on these groups. The language used to label the lower categories is also important and 'emerging' or 'developing' could be considered.

It is also not clear how the government's policy drivers towards commercialisation would interact with this proposed change in scale. A move to increasingly applied research and experimental development is in the main, likely not to be consistent with the maximisation of citation scores.

PEER REVIEW AND CITATION APPROACHES

This is an opportunity to fix emerging and legacy issues with the peer review

It has been long known that the methodological difference between the peer review and citation analysis approaches leads to a distorted picture of research excellence. The issues articulated above are amplified for the peer review approach. This review may wish to examine whether the two approaches should be brought onto a common basis or treated separately. In either case, it is an issue that requires significant further work, as identified by the recent ERA and EI review.

UA is concerned by the divergence between peer review and citation disciplines that has become evident over successive rounds of ERA and what mechanism there is within the changes proposed to address this divergence.

It would be interesting to understand further why there has been this divergence: is it an accurate reflection of the quality of research, a function of the peer review process or assessment panel interpretation of the data, or perhaps a reflection of the relative level of investment in these disciplines? The discussion paper points to a solution that focuses on improved guidance to the peer review panel, increasing the need for guidance and language to be consistent and unambiguous.

It would be worth considering an analysis of any divergence between ERA outcomes and those from a range of international benchmarking and ranking systems. The challenge is to address the issue that some units of evaluation have hit a ceiling without risking an effect where most of the units of evaluation fall lower. Only some Fields of Research have demonstrated that ceiling effect. It is important to avoid creating a scale that artificially repositions sections of the sector as underperforming.

Take care that analytics does not subordinate expert judgment

The design of ERA intentionally sought to use the assessment panels as committees able to exercise judgement about what constituted excellence in a discipline and to use the data (peer assessment and citation) as indicators to support those judgements.

A risk in the proposal is that the use of Relative Citation Index (RCI) data to identify highly performing institutions and the excellence of FoR at Australian universities will displace or downgrade the importance of panel judgements.

Greater dependence on citation indices appears to provide an efficient means of judging quality but is based on the premise that the RCI are indicators of excellence. They may well be indicators of dominant trends in a discipline while still failing to track excellence.

The proposed use of RCI will need to be critically assessed to remove outliers, informed by panels of experts who understand the discipline. Again, a pilot would allow for an assessment of whether the outcomes determined through RCI are aligned to or consistent with the judgement of panels.

OTHER CHALLENGES

Citation Measures – the measurement is highly dependent on citation providers

UA is concerned that the proposed reforms will place a high amount of power into a few companies to determine what is world leading through the way in which they collect and process bibliometric data. Given the sensitivity of the measures to only a few papers, it is also highly dependent on the way in which the data is collected and treated by providers.

They have also raised that measuring scientific quality using the proxy of citation can be problematic. It may push researchers to only publish papers that will be highly cited. It has the danger of skewing the literature towards positive / statistically significant publications, when it is also important to understand what does not work, so as not to repeat experiments or trials that are negative or not statistically significant.

It would not be useful for ERA to lead researchers to restrict their publishing options to certain journals with particular impact factors, given the important of reporting rigorous research in a broad range of outlets.

It would be useful to discuss the sector and the ARC's views on these matters.

Indigenous research

It is not clear how Indigenous research would be treated under the proposed system or how the guidance provided translates to an internationally benchmarked rating scale. The government has made great strides in recognising the importance of Indigenous research through the revision of the ANZSRC classification system by the ABS and the introduction of the new division level code. It would be counter to this intent if in turn these incentives unintentionally marginalised Indigenous research.

Again, piloting before full implementation would allow an examination of how indicators of excellence, impact on community and how Indigenous lead research are taken into account through the ERA exercise.

UA is concerned about the number of questions proposed as prompts for reviewers may unintentionally lead to a higher workload for reviewers, particularly for FoR Code 45 where two sets of guidance questions are provided. It will be important to frame these guidance questions in such a way as to set a reasonable expectation for workloads, so an unintended consequence of fewer peer review volunteers is avoided.

National security research and commercially oriented research

It is also unclear in the consultation paper whether consideration has been given to how national security research and commercially oriented research – which do not lend themselves to standard metrics but which government policy is encouraging – would be recognised.

Measuring the cost and benefit of conducting the ERA exercise

The cost of conducting the ERA exercise across universities in the sector and in the ARC is considerable. While some of the proposed changes may reduce some of the administrative burden of conducting ERA, UA strongly suggests that the ARC collect and share information with the sector about the full costs involved in ERA submissions. This will allow all those involved to understand the cost-benefit equation involved with ERA and assess its place as a fit-for-purpose program in the Australian and global research measurement landscape.