

## **Feedback from the Faculty of Science UNSW ERA and EI Review**

### **ERA**

#### **Q3.1. To what extent is ERA meeting its objectives to**

**a. ...give assurance of the excellence of research conducted in Australia**

A small amount

**b. Provide a national stocktake**

A moderate amount

**c. Identify excellence...**

A moderate amount

**d. Identify emerging research areas and opportunities for further development.**

Not at all

The review uses metrics that are firmly in the rear view mirror – so is not effective in identifying emerging trends.

**e. Allow for comparisons of research in Australia, nationally and internationally, for all discipline areas.**

A moderate amount

Good for comparisons across the sector domestically and value in offering a snapshot of research strengths at different institutions. It is however based on expert opinion, and doesn't allow meaningful comparisons with the best universities internationally.

#### **Q 3.2 Meets the needs of the stakeholders**

Disagree

The lack of international comparison makes it less valuable. It is backward looking, based on expert opinion not data. ISI Web of Science is based on quantitative data and is more relevant to academics. Academics see no relevance, and are unlikely to use ERA to determine where to publish. Neither ISI nor ERA are useful for determining the value of new journals.

#### **Q3.3 What impacts has ERA had on....**

Universities now invest money and considerable time in compiling the data and manipulating where outputs are positioned. At present, ERA provides the government with comfort that research money is well spent, but it might be more meaningful if decisions about resource allocation were made regarding truly "world class" research.

#### **Q3.4 and Q3.5 How do you use this and are the outcomes beneficial....**

Not used by academics in any meaningful way

#### **Q3.6 Suggestions for enhancing ERA**

Use ISI, consider alternative metrics

#### **Q3.7, 3.9. 3.10 The current methodology meets the objectives of ERA. Strengths, weaknesses and does the discipline specific approach work**

The methodology is very open to "strategic manipulation". Using metrics like citation analysis is helpful, but flawed. For example, in a field like materials science there are some sub-fields where citation rates are traditionally much higher or lower than other areas, eg papers on graphene attraction huge citations levels, but steels hardly any, so some skewing

there. Focusing on graphene and moving out of steel research, would surely lift that ranking. There are also (often low impact/ranked) journals (commonly in life sciences, public health etc) that publish “Australian” research that may have important impact locally. These papers may not attract high numbers of cites and would struggle to be published in the best international journals, but may have important impact at a local level or in clinical care and may be important in changing policy or guidelines. Discipline specific approaches don’t really work. Where a FOR code has 2 disciplines eg Ophthalmology and Optometry, where an institution has different performance in each of the disciplines, there is often a low score overall.

There is also some capacity to skew evaluations by including and excluding visiting academics or hiring academics specifically to boost ERA metrics.

Reliance on citations and publication metrics talks only of publication quality. The actual impact of a study is often not addressed. “Applied measures” doesn’t look at the effect of research on social policy, medical practice etc. Perhaps place more weight on non-Cat1 research income? Alternative metrics – could be assessed in a way that is less onerous re the EI process., white papers etc.,

### **Q3.11-3.15 Citation analysis and peer review**

'Grade inflation' over time has been noted in both the UK REF, e.g.

<https://www.mendeley.com/catalogue/ad97cd0f-4bfc-371b-80a5-5145d3b5c9fa/>

and in the Australian ERA

<https://campusmorningmail.com.au/news/teaching-and-research-belong-together-why-we-dont-need-two-eras/>

The first figure in the latter is taken from the Consultation document and is already concerning, but the second figure is striking. Both the above references suggest improbable improvement in fields assessed by bibliometric means. So, ERA is probably now compromised by the very coarse scale it uses for grading (Q3.23, Q3.24) and by the broader well-known technique of optimisation with respect to bibliometric indicators. The huge difference in improvement between citation-judged disciplines and peer-reviewed disciplines is a cause for concern and for incredulity (Q3.11 vs Q3.15).

In the UK REF2021, Mathematical Sciences are subpanel 10 of Panel B. Para 279 of the REF2021 guidelines says "Sub-panels 10 and 12 believe that citation data in their disciplines cannot be used to provide sufficient added value to inform the assessment of output quality. They therefore will not receive nor make use of citation data."

In ERA all but Pure Mathematics within the Mathematical Sciences are assessed using citations. It seems unlikely that ERA would increase peer review but the disconnect between REF and ERA in Pure Mathematics is striking. Given that REF has real implications for funding, one could conclude that they regard bibliometric analysis as insufficient to make big calls on paper quality.

### **Q3.22 Alternative metrics**

Agree. Should there be one process EI/ERA rather than 2?

### **Q3.23 and 3.24 Scale is suitable for assessing research excellence**

If the purpose of the evaluation is only to assess if a body of research is “world class”, then the ratings scale is appropriate. If the purpose is to determine quality on a more granular level, then bands 4 and 5 should be further sub-divided to make finer adjudications.

### **Q3.25 Low volume threshold appropriate?**

The low volume threshold needs to be raised. There are some disciplines at some institutions where the assessment for a FoR code uses very small blocks of research (often 1 or 2 people in effect), this should be raised. Under these circumstances, institutions may claim a ‘5’ for their School on the basis of one or two people and by ignoring outputs from a number of research inactive staff.

### **Q3.27 By-line or Census date?**

Use by-line not census date. The census date promotes transfer behaviours between institutions to manipulate outputs. This also happens with REF.

### **Q3.31 Interdisciplinary fields**

This is an area where non-standard metrics might be beneficial. ERA does a pretty poor job in interdisciplinary fields.

### **Q3.35, 3.36 Annual collection – advantages/disadvantages**

Given the amount of time and effort in massaging data that takes place in the current cycle, an annual collection would be unworkable, every 5 years is fine.

There are tremendous costs involved for an annual collection performed on the same basis as now. In the current climate, universities may struggle to comply an annual request with fewer staff. If universities simply handed over their AROC collection, then annual capture would be straight-forward.

### **Q3.37. Should volume of output be published?**

Publishing the number of outputs or more specifics would give more indication of how important a discipline is at a university. Again, just knowing a university met the quota in a discipline is not much of a measure of scale.

## **E&I**

### **Q4.1-4.3 General comments:**

1. The UK combines the exercises of ERA and EI, as the consultation document states. This is probably a better way to give a snapshot of research quality. The work required to generate this information also tough for Schools, it was certainly pretty draining to back up for EI after ERA the year before.
2. EI didn't work well for enabling sciences. Many applications for example in maths and stats are counted in a different FOR code so the true impact is perhaps diluted.

3. Limiting universities to a very small number of case studies for a field as a means of demonstrating engagement/impact across that field is limiting (especially in Engineering and Medicine that are across one and two FoR codes only compared with 7 in Science).
4. The narratives requirement seemed unconvincing without definitive metrics.
5. This was a new exercise and there was huge confusion about what the ARC was after.
6. Most institutions now place more emphasis on “E&I” with their staff in general. So, most academics think more about who their end users and stakeholders might be. But, the “E&I” process has not directly impacted on that.

#### **Q4.4,4.5,4.6 How are the EI outcomes used**

As with ERA, but more so, these are necessarily backward looking so not helpful in driving investment in new areas. They do make good news stories for universities, however.

#### **Q4.7,4.8,4.9,4.10,4.11, 4.12,4.13. Definitions?**

This is generally OK. Academics are still confused which is impact and which is engagement. 2 digit FOR codes are not granular enough for engineering and medicine.

#### **Q4.14,4.15 Low volume threshold appropriate?**

There should be a higher threshold – again localised examples of engagement are not necessarily representative. Having the university approach to engagement perhaps mitigates this to an extent.

#### **Q4.16,4.17,4.18,4.19,4.20,4.21,4.22,4.23 Metrics appropriate?**

The metrics used to assess these need a lot of work. Dollars are easy to measure, but every other metric was difficult to capture and not necessarily quantitative. The metrics need to be re-designed to include only those that are clearly measurable and comparable across the sector. This requires a change in systems for data capture.

Metrics are different across disciplines and no university captures this information well – industry HDR supervisors, number of people impacted at a public lecture, are examples of metrics not broadly captured. Where industry partners as authors on papers is evidence of engaging end uses, many of these people may be so well engaged they have visiting appointments for example so are not visible as end users.

#### **Q4.24 Narrative suitable?**

The narrative process was ineffective and not metrics based. In many ways a process more about the quality of the writing than the research.

#### **Q4.25 on Format of narrative and number of submissions?**

Small number of FOR codes for engagement/impact across is limiting, especially in Engineering and Medicine that are across one and two FoR codes only compared with 7 in Science.

#### **Q5.1,5.2,5.3,5.4**

Every 5 years – see comments above

**Q5.5,5.6,5.7,5.8**

There is merit in combining, see UK approach but there are challenges. The two assessments assess quite different measures and over a different time frame. ERA is measuring where the research is “world class”, E&I is measuring whether the local steel works are engaged – these are quite different criteria.

**Q5.9**

From an EI perspective, not having data routinely collected and available to compile the reports. Reliance on individual schools who have very different methodology and approaches for compiling this information made the process more of a lottery.

**Q5.10**

Agree will help to capture data and perhaps this helps in knowing whether you have 2 academics or 50.

**Q5.11**

Yes, as much automation as possible (and with minimal opportunities for gaming by institutions as possible).

**Q5.12**

DOIs are helpful for in press items or non-standard eg web-based publications

**Q5.13**

Standardise metrics, as much automation as possible and provision of raw data to the ARC.