

Submission: ERA EI Review Public Consultation

12 October 2020

FACULTY OF ARTS AND EDUCATION AT DEAKIN UNIVERSITY: RESPONSES TO THE REVIEW QUESTIONS

This submission is the result of a group discussion in the Faculty of Arts and Education at Deakin University by a group of academics, all of whom have been involved in preparing the Faculty's portfolios for both the ERA and the Engagement and Impact assessments. These are our views and not necessarily those of Deakin's Research Office. They are offered here as a contribution from the HASS disciplines to the Review.

As a group, we have the following points to make in terms of the general aims of ERA.

1) In terms of question 3.1 c and 3.1.d and Q3.23 and 3.24 we felt that the scheme as a whole took a wrong turn in establishing a ranking system that equated excellence with an absolute idea of an international standard. Standards can shift and not recognising that invites confusion between relativity vs notion of absolute. There is also a serious question as to the methodology used for establishing what this international standard is, not least because majority of reviewers, at least in the peer based disciplines, are Australian. To be blunt, it smacks of cultural cringe. Far more meaningful would be to ask reviewers to evaluate the originality, significance and rigour of the actual work within disciplines (FOR) as indeed is done by the British system and to have a scale associated with each term.

This also means that ERA has little value in terms of ranking Australian work vis-à-vis the same discipline elsewhere in the world (question 3.1 e) as the benchmarks being measured are not the same. Britain does not ask if its research is at or better than, international standard, for example. Many other nations do not measure at all. There is no international ranking system for research excellence in HASS disciplines despite world rankings whose methodologies are questioned by many as they are driven by citation/metric measures that are simply not as relevant in HASS for determining excellence.

The point also speaks to our firm belief that it is important to keep the process of peer review for the HASS disciplines. Citation metrics say little about the quality or excellence of the research. Attention could be because it is generally considered by the academic community in question, to be a bad piece of work. The attention received by, for example, Keith Windshuttle during the history wars would be an absolute case in point. Establishing the quality of a piece of work requires judgement based on knowledge of the field and reading the work in question. Basing judgement on metrics is unreliable. It is perfectly possible for a brilliant piece to be published in a journal that isn't even ranked in Scimago, particularly in the humanities and in Australian related content. Peer review recognises this. The problem for HASS disciplines however, lies in the fact that the selection of material for peer review is often made on metrics which may or may not actually capture the best work in a given discipline within HASS. This is because Universities tend to develop university wide policies that reward citation based methodologies as these are the methodologies used in university rankings. This will work for STEM but not, we suggest for HASS. Thus, the selection of the 30% of peer review material will often be made on external criteria such as citation, quality of journal defined by impact factors and scimago rankings, quality of publisher with international publishers, including international university publishers ranking more highly. Such criteria for selecting what goes up for peer review does not, necessarily, reflect the quality of the work. A case in point at Deakin is Anthropology. Here researchers with excellent national and international standing within that field published in journals not defined as Anthropology journals within the ERA system. Apart from being a problem for ERA in recognising the interdisciplinarity and cross-over of many fields, what this meant was that Anthropology did not go up for review. A solely peer review system would not have led to that decision being made and there would have been a good chance of doing well. Publishing in an ERA listed

journal under a particular FOR or a Q1 journal is not, in HASS, a determiner of quality. Expert judgement in the field is.

The impact of these issues on Australian Universities, and on HASS researchers in particular is not uniformly positive (questions 3.3). The system strongly discourages investment in local debates which may in fact be quite pertinent to the value of that discipline within a local environment – take Australian History as an example. Should we make a demand that all debates within Australian history connect with international debates? Why do we assume that international is better or more relevant? This is not to say that Australian historians working on Australian history cannot make a significant contribution to the discipline of history at an international level but creating an expectation that only work that does so is of an international standard is mistaken and even a negative force for the field. Peer review mitigates these issues because it allows for work published in local journals to be judged as of high quality. Choices and judgements made purely on metrics however, would often miss this work because those journals/book publishers would not be highly cited by international standards. While the impulse to make our research more international is not wrong, it is wrong to set up a system that equates the local with a lack of quality, particularly for disciplines, or parts of disciplines where the focus is Australian content. The question then, is whether it is possible to set up criteria for HASS that use internationally recognised indicators in HASS disciplines but which are nevertheless, also sensitive to local contexts? If we followed this line of argument, criteria which require judgement of experts undertaking peer review would bring us back to questions of originality and significance of contribution. What this means is that the system has to allow for peer review as it presently does but back this up with meaningful criteria and indicators – which it currently does not.

With regard to contextual indicators (3.35). The issues here are similarly complex. One issue is that there is no space to take account at the two digit level, of institutional narrowness within a given field. Thus, it is easier to achieve a four or a five with a very narrow band of work by many people than it is to be recognised for excellent work across a broader range of four to 6 digit level specialisations. This opens the system up for gaming. The indicator of funding needs to be nuanced for discipline as well. It is much harder for philosophers, for example, to do well in income given they have neither the opportunity to get cat.2-4 funding or indeed cat. 1 level funding as say history within the HASS disciplines. It is therefore inherently unfair to compare History and Philosophy or to have the same indicators for both disciplines. Applied Measures (patents etc) are mostly irrelevant in the HASS area while research income is a valuable indicator for particular kinds of work but not others such as the more traditional humanities or the creative arts. Publishing indicators are too driven by STEM based assumptions when it comes to the Humanities. More recognition needs to be given to both the significance and the quality of work that simply is not published in American journals that form the basis of international ranking systems but which are not, in themselves, the only measure of research quality.

Another set of issues have to do with the location of researchers and its impact on what work gets counted where come census date. (q.3.25-3.26). There is absolutely no doubt that allowing research to follow the researcher has led to serious levels of poaching and that the institution that invests in particular researchers is not necessarily the institution that reaps the reward. A far more accurate picture of the research standing and the investment of institutions in their researchers and their fields of discipline would be gained if the research was counted at the institution where it was done rather than the institution the researcher was at at the time of the census cut-off date. A by-line is a far more honest system for acknowledging the location of the research in question.

On questions of interdisciplinarity there is, in our opinion, no doubt that ERA does not adequately capture such activity. One solution to this problem is to change the limitation of 66% rule on how many journal articles can be moved between disciplines which are increasingly interdisciplinary in nature. This is an acute problem in HASS where it is not unlikely for example, that 2102 can be published in 1201, 2103 or in 1601 or 1603 and still be 'Cultural Heritage studies'. In addition, a review of FOR codes for journals is needed to more accurately reflect the large quantity of interdisciplinary journals as well as the multitude of areas of field studies that specific disciplines can contribute to – for example, both sociology and anthropology contribute to Science Technology Studies, which does not have an FOR code because it is not

a discipline as such. The 66% rule is preventing very good work in particular disciplinary contexts from being visible.

2. Data collection

Another set of issues is around processes for data collection (questions 3.4) Our feeling here is to be wary of automated systems such as Orchid which may not necessarily automatically catch all relevant publications for HASS given their harvesting mechanisms. If there is an annual collection system this should be by returning to a HERDSE style collection of publications. The advantage would be the ability to stop gaming for all universities. The disadvantages, the workload increase. There would need to be more resources in managing the monitoring and classification of outputs.

We do not think that the evaluation of the quality of that research should happen annually. Indeed extending the timeline for evaluation of quality out to five years might actually help with getting a clear picture of the depth, breadth and quality of investment in research by any one university in any given field.

Publication of data as part of each ERA exercise would certainly help with transparency and the minimising of gaming.

3. On Engagement and Impact.

1) The general feeling from the Faculty on this aspect of the Inquiry is that there were some problems with the assessment only being done at the two digit level. This is because it is very hard to actually see the discipline based contribution to both engagement and impact. On the other hand, given the lack of systems for documenting both engagement and impact, the amount of work required to change this is too high. One possible solution is to have the ability to produce more than one case study per two digit FORs so that it would be possible to showcase more than one discipline if there was a desire to do so. However, this would not be mandatory.

2) There is a lack of clarity between any possible connection/disconnection between ERA results and E & I results. Our sense is that in some cases excellent results in ERA do not necessarily translate to excellent results in E & I. The relationship was not necessarily a causal one. Some felt that E & I might actually be better measure against SEOs rather than FORs which would also allow for interdisciplinary work to be counted more easily.

3) General feeling was that narrative rather than metrics was key here. A lot of important work done in this space that did not necessarily have a lot of dollars attached to it. Defining what could count as a case study by the amount of dollars attached to it ruled out a lot of meaningful engagement with real world impacts. The general feeling was that there was a lack of clarity as to what the indicators for engagement could be and that a reduction to metrics such as dollars was not helpful. In addition, the feeling was that dollars should be counted once, probably in ERA and not in E and I.

4) If choice of case studies for impact is dependent on dollars earned from industry partners, there is a question of institutional ownership. In the case of a case study on impact from a research project funded by a Linkage scheme across two or more universities, who owns the case study? It should be possible for the same case study to be used across universities. The length of time for a project to have impact is also problematic as researchers move across institutions. The problem here is the reverse of ERA. In the latter, the outputs follow the researcher. In E and I, if a researcher moves, neither institution can count the impact.

5) Unlike with ERA, general sentiment was that the rating scales in E & I were appropriate.

CONCLUSION AND RECOMMENDATIONS

Our group recommends the following:

Recommendation 1:

Abandon the current rating scale and develop one that is not vague and which does not use abstract measures that invoke absolutes where these are not easily substantiated. International world standard is neither an accurate measure of quality nor something which is easily evidenced or measured.

Recommendation 2:

Devise a system for peer review that does not encourage universities to use a citation based metrics system to drive publication practices in HASS and prioritise the international over the local. Instead, have meaningful measures of quality that require expert judgement. Criteria such as originality, significance, scholarly rigour.

Recommendation 3:

Remove limitation of 66% rule to help with interdisciplinarity issue. At the same time, allow finessing of data sets so that publications clearly out of the field are not counted in the field – eg. Any article discussing pedagogy being counted in education, or any article with policy implications being counted in policy studies.