



Excellence for Research in Australia and the Engagement and Impact Assessment schemes consultation

Submission by Deakin University

October 2020

Excellence for Research in Australia and the Engagement and Impact Assessment schemes consultation

Submission by Deakin University, October 2020

Introduction

Deakin University is pleased to contribute to the ERA consultation process in 2020.

Overall, we support the continuation of the ERA process and framework and the Impact and Engagement Exercise. We do, however, make some suggestions which would prioritise efficient data management, reduce the large burden on the university sector and ensure greater transparency.

The ERA process enables objective review and reflection of our research impact and quality. It highlights high-quality work and outcomes at a national level against global standards for the most part.

Our concerns are centred around the very significant burden on academic and research support staff and university to produce the data and related materials and the frequency of the process – although the move from every three years to every five years is to be welcomed, based on previous ERA exercises, perhaps every six years would be more appropriate for detailed comparisons with past data sets (i.e. 3 + 3 years).

There are significant concerns that a lack of clarity regarding some of the rules and criteria may allow some institutions to gain inequitable advantage. In addition, while 90 per cent of all research across Australia is now considered to be at or above world-standard, there should be some further nuancing at the top end of the scale.

Furthermore, the silo-based assessment of codes can disadvantage applied and multi-disciplinary research teams that conduct research across disciplines. A greater focus on the strengths of cross-disciplinary research would enhance innovation and creativity across disciplines within Australian institutions.

We have drawn on the most relevant ARC ERA review questions to frame our responses, but in summary, Deakin's overall conclusion is that ERA provides a sector wide benchmark that allows for disciplinary comparison both at the national and international level, and with identified universities.

It should be noted, however, that ERA does have a disproportionate impact on those disciplines that cannot as easily be judged by the more usual metrics of publications and citations, as is typically applied across the STEMM disciplines. It is Deakin's firm belief that it is important to keep the process of expert peer review particularly for the HASS disciplines. Nevertheless, there are some fields currently peer reviewed where the metrics of publications and citations could be reasonably applied rather than peer review and this should be considered and potentially validated by comparisons in the next ERA.

Deakin's detailed response to the Terms of Reference follow.

Deakin University's Responses to the Review Questions

SECTION 3: EXCELLENCE IN RESEARCH AUSTRALIA

ERA Policy:

Q3.1: To what extent is ERA meeting its objectives

Overall ERA provides a sector wide benchmark that allows for disciplinary comparison both at the national level and with identified universities. Internally ERA can be applied to set aspirational targets for the development of research, teams, and disciplines. It does, however, have a disproportionate impact on those disciplines that cannot as easily be judged by the more usual metrics of publications, citations, typically applied across the STEMM disciplines. It is Deakin's firm belief that it is important to keep the process of peer review for the HASS disciplines. Nevertheless, there are some fields currently peer reviewed where the metrics of publications and citations could be reasonably applied rather than peer review and this should be considered and potentially validated by comparisons in the next ERA.

Internationally the picture is somewhat different and more difficult to accurately compare. While the comparison of international publications and citations used across the STEMM disciplines used in Australia is generally applied elsewhere e.g. in the REF Exercise used to evaluate UK Universities, it is important to note that different funding mechanisms for research exist, with different priorities, as well as different levels for funding for research exist in other countries e.g. USA, China, Europe etc. In addition, 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 may not be as relevant in HASS for determining excellence.

A key issue with the publication and citation methodology is a need for robustness and reliability that the process captures all of the publications and published reports and outputs associated with the academics at the University. It is key that future exercises work with ISI Clarivate or with Elsevier (SciVal/Scopas) to ensure that all outputs are captured.

It is also important to note that ERA is an exercise in looking backwards, not forwards. So while Deakin would agree that it does identify excellence across the full spectrum of research activities, it does not directly identify emerging research areas or areas and opportunities for further development. Similarly, it does not particularly inform a range of policies or initiatives. As an exercise it may highlight areas of research focus where Australia does not perform well or participate at all but the connection between ERA and policy government research policy and incentives such as funding, is relatively tenuous.

Engagement and Impact (EI) I was measured for the first time during the 2018 ERA Exercise. One of the clear outcomes from this exercise was the finding that the timeline from laboratory-based research to industry or societal impact is generally significantly longer than the period chosen in the first EI exercise. Different types of outcomes can take several years if not decades to be adopted as standard practice, particularly in medical and animal practices where the lives of people are impacted.

Q3.2: The ERA objectives are appropriate for meeting the future needs of its stakeholders. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree.

Agree. ERA has certainly moved the research agenda and focus toward research quality rather than quantity and brought together a picture of research in Australian universities that ranking bodies around the world find useful and easier to apply. It is certainly one of the factors in the improved ranking of Australian universities over the past decade.

As an exercise, it is also very clear and visible feedback to the funding bodies (ARC/NHMRC/Industry/other) as to the outcomes from the use of awarded research funds, milestones achieved, publications and reports, intellectual property and licences and compliance with guidelines would provide insight into the alignment of university research activities with the Government's science, research and innovation agendas.

Q3.3: What impacts has ERA had on:

- a. the Australian university research sector as a whole**
- b. individual universities**
- c. researchers**
- d. other?**

ERA has certainly moved the research agenda and focus toward research quality rather than quantity and brought together a relatively objective picture of research performance and activity in Australian universities. For individual universities, it has allowed them to drive the research agenda towards positive outcomes. For individual researchers it has allowed a focus on research quality and quality outcomes – publications in the highest impact factor publications because there is an intrinsic knowledge that the quality for their research outputs will be judged every few years.

It has also driven a change in publication attitude by academics, for example, particularly away from conference proceedings and conference publications which unless they are in certain fields (e.g. Computer Science or Mathematics) are no longer acceptable. Also see answers to Q3.1 & Q3.2.

Q3.4: How do you use ERA outcomes?

There are several ways in which ERA outcomes are used. The first is to advertise the quality of the research across the Deakin – for attracting Higher Degree by Research (HDR) students, new academic staff and undergraduate students.

It also helps an institution to develop a university research strategy in that it allows institutions to take a forward look and clearly identify which research areas will be affected over the period before the next ERA, in terms of staff who may leave or retire or where there are gaps in what are developing areas. In the past Deakin has used this information to inform areas of research that require strengthening with new appointments or areas of research that we were prepared to allow to come to a natural end point.

Q3.5: ERA outcomes are beneficial to you/your organisation. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Agree. ERA provides a sector wide benchmark that allows for disciplinary comparison both at the national level, and to some extent at the international level, with identified universities. Internally ERA can also be applied to set aspirational targets for the development of research, teams and disciplines.

Q3.6: Do you have any suggestions for enhancing ERA's value to you/your organisation? Please explain your answer.

Enhancing the value of ERA would be by defining ways to undertake this assessment without such a large staff time and effort. By modifying the process as far as possible to be data driven with AI collection and analysis would provide the evidence the process delivers now but without the huge staff efforts across the sector.

ERA Methodology

Q3.7: The current methodology meets the objectives of ERA. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Agree. The approach of the current methodology meets the objectives of ERA. The low volume output threshold of 50 indexed apportioned journal articles for citation analysis for a particular discipline group is appropriate for those disciplines that publish journal articles. Similarly, peer review of 30 per cent of outputs for disciplines that do not routinely publish journal articles is also a reasonable cut off point. Consideration is also given to research outcomes that cover categories 1-4 grants and applied measures. Traditional and non-traditional output types are also assessed, which takes into account the full repertoire of outputs from diverse disciplines.

Q3.8: What are the strengths of the overall methodology? Please describe

The overall methodology is comprehensive and has enormous breadth and depth. It takes into account the large variety of research activities undertaken by different disciplines. As a process it is particularly strong at grading the STEM disciplines, less so for the HASS disciplines, hence the need to retain peer review (potentially with some exceptions).

Q3.9: What are the weaknesses of the overall methodology? Please describe

Extensive time is required for peer reviewers to assess the outputs of disciplines groups. The impact narrative has fairly diffuse guidelines in its current form and the creation of these impact narratives is very time consuming. There were situations in previous ERA assessments where discipline groups obtained ratings for universities where these discipline groups were not present. Again, the methodology of ranking the STEM disciplines and the HASS disciplines is at the heart of this comment.

Q3.10: Does the discipline-specific approach for evaluating research quality (citation analysis or peer review for specific disciplines) continue to enable robust and comparable evaluation across all disciplines?

Agree. A discipline specific approach is needed for various disciplines to enable fair and equitable analysis of research within each individual discipline. There is no ideal approach but a combination of academic impact (citations, FWCI, Q1 rankings) and impact more broadly is a reasonable balance.

One alternative would be to enable disciplines that normally publish work in journals to have the option of having a portion of the research outputs available as to be assessed as peer review outputs. It is possible that researchers have produced high quality work in journals that may not be associated with many citations, due to the unique or innovative nature of the work produced. Having an option to include these research outputs for peer review analysis could be useful.

Q3.11: The citation analysis methodology for evaluating the quality of research is appropriate. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.

Strongly agree. For the STEM fields the use of relative citation impact is appropriate given the primary research outputs are academic journals and this should be retained as a primary measure of research performance.

The citation analysis methodology is appropriate, because it is used for assessing research performance for disciplines where the primary research output is in academic journals. This is an equitable approach for disciplines in STEM, but as has been mentioned above, this is not necessarily appropriate approach for many of the HASS disciplines. The other issue is that citations build over time. High impact work submitted towards the end of an ERA period many not yet have accrued the level of citations the quality of the work suggests or deserves.

There are some fields currently peer reviewed where the metrics of publications and citations could be reasonably applied rather than peer review and this should be considered and potentially validated by comparisons in the next ERA.

A further refinement could be to place a greater emphasis could be placed on Field-Weighted Citation Impact (FWCI) to ensure equity across disciplines.

Q3.12/13: What are the strengths/weaknesses of the citation analysis methodology? Please describe.

The citation analysis methodology has world citation benchmarks for various disciplines. It is therefore an equitable process for the specific disciplines concerned. However, it needs to be noted that even across similar disciplines, for example in STEM comparing Chemistry and Engineering, the citation rates are much higher in Chemistry than in Engineering, so comparison within a discipline is appropriate but should not be used to compare one discipline with another.

Q3.14: Can the citation analysis methodology be modified to improve the evaluation process while still adhering to the ERA Indicator Principles? Yes/No.

a. If you answered 'Yes', please describe how the methodology could be improved

Yes. Consideration could also be given to a similar metric of citation analysis, which is the more commonly applied and understood field weighted citation impact (FWCI).

Q3.15: The peer review methodology for evaluating the quality of research is appropriate. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.

Strongly Agree: *This is an appropriate approach for disciplines where the primary research outputs are not in journals. These disciplines do not traditionally enjoy the same number of citations that are inherent in STEM disciplines. It also acknowledges that that impact per se has an element of greater importance to community as end-users rather than peer-determination of impact.*

Peer-review assessment remains the most appropriate option for assessing research quality in the Economics (FOR 14), Commerce/Management (FOR 15) and Law (FOR 18) disciplines as well as across the majority of the HAS disciplines. In support of this position to switch to a citation-based assessment of research quality would therefore set Australian business schools apart from the international norms for assessing research quality in the business discipline. Using a citation-based approach for evaluating the research quality of law research would therefore run contrary to the disciplinary norms for assessing legal research.

Q3.16: What are the strengths of the peer review methodology? Please describe.

The peer review methodology enables very comprehensive and detailed analysis of the research outputs produced by the discipline groups for which this methodology serves. The approach relies on expert opinion and its use is very methodical, meticulous and systematic.

Q3.17: What are the weaknesses of the peer review methodology? Please describe.

The weakness of the peer review methodology is that it is carried out by a series of individuals and it is opinion driven. This could be seen as a strength (or a weakness), but it appears to be the latter. Furthermore, it does not enable any form of global benchmarking.

Q3.18: Can the peer review methodology be modified to improve the evaluation process while still adhering to the ERA Indicator Principles? Yes/No. If you answered 'Yes', please describe how the peer review methodology could be improved.

Yes. It would be useful to have additional assessors allocated to each discipline group to complete the peer reviews. However, this is an extremely labour-intensive exercise and would need to be undertaken with due diligence and attention to detail if the resulting peer review is to be of value to the assessment process. It also needs to be undertaken in addition to current responsibilities. Minimising the evaluation process would reduce the impact on assessors.

Q3.19: The volume and activity indicators are still relevant to ERA. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.

Strongly Agree. While research quality is, and should remain, the main driver, research volume and activity indicators show both the breadth of the research and the potential future impact and translation.

Q3.20: The publishing profile indicator is still relevant to ERA. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.

Strongly Agree. The publishing profile provides some important contextual information about a particular discipline. It should be noted that a calculation of the number of papers per member of academic staff can disadvantage some universities where teaching is given a higher priority and so the staff profile may differ when compared with more research-intensive staff or more research-intensive universities.

Q3.21: The research income indicators are still relevant to ERA. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.

Strongly Agree. The research income indicators are extremely relevant. While the winning of research grants is in itself only the starting point to research excellence, it is regarded as an indicator of prior quality of research outputs, both in terms of high-quality publications and in terms of other outputs such as HDR completions.

Q3.22: The applied measures are still relevant to ERA: a. Patents. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree.

It was the inability of applied measures to properly take into account the strengths of all disciplines across ERA that lead to the introduction of the EI process in parallel with the ERA process. For example, in the social sciences, an applied measure might be the endorsement of an intervention, uptake in relevant communities or a measure to be adopted at a state or national level. In the STEM disciplines and applied application of research may result in patents or a spin-out company, as one measure of success. However, business assists and the ability to solve key problems for industries with the advanced tools at a university's disposal are also relevant applied measures.

Q3.23/24: The five-band ERA rating scale is suitable for assessing research excellence. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.

Strongly Disagree. The 5-band indicator needs to be revised. It does not take into account any increase in volume of research across an institution. The plot Figure 1. in the ERA EI Review Document (page 16) which shows the

percentage in any given category but does not reflect any change in the volume of outputs produced by a given university. While research quality needs to be the defining criteria, the amount of that research quality also needs to be assessed and published.

The fact that 90 per cent of units evaluated are now considered to be at or above world standard would tend to indicate that the rating scale was in need of modification, particularly at the top end.

Q3.25: *The ERA low-volume threshold is appropriate. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.*

Strongly agree. The threshold should remain to ensure a reasonable level of equity across the four-digit FoR codes. The threshold of 50 publications is sufficient to reliably assess the quality of the research over a sustained period and allows evaluation of research conducted by smaller or emerging research groups. Implementing higher thresholds could lead to smaller research teams merging and congregating at one institution, resulting in the potential for group think and limiting the cross exposure with other research teams and disciplines. Encouraging opportunities for choice, independence and flexibility of options allows different influences to support creativity and development.

Q3.26: *Are there ways in which the low-volume threshold could be modified to improve the evaluation process? Please describe.*

Consideration should be given to introducing a quantum indicator, at least in the publication of the assessment outcomes if not the assessment itself, so that the magnitude of the research contribution to the field at different universities can be understood.

Q3.27: *What is the more appropriate method for universities to claim research outputs—staff census date or by-line? Please explain your answer.*

Q3.28 *What are the limitations of a census date approach? Please describe.*

Q3.29 *Would a by-line approach address these limitations? Yes/No. Please explain your answer.*

Q3.30 *What are the limitations of a by-line approach? Please describe*

The staff census date is a more reliable and transparent approach. The use of ERA staff census date seems potentially open to 'gaming' by institutions, and thus may undermine objective evaluation of research performance. However, it is unclear if this is actually a major or a minor problem, but with regular movements of staff across institutions this likely balances out. The by-line approach, whereby outputs are only accounted for by the institution(s) named on publications is also limited as it is unable to provide a snapshot of current research capacity, which is one of the most useful aspects of the activity. Thus, the use of ERA staff census date should be retained.

By-lined publications are a lag indicator of the research quality of an institution, and for some disciplines/journals, this lag can be significant. An analysis conducted by Huisman and Smits (2017), for instance, found that the length of journal review across all disciplines was longest in the business disciplines (FORs 14 and 15), with an average review length of 25 weeks, double that of the disciplines with the shortest review period (12-14 weeks: medicine, public health and the natural sciences). Notably, the figures underestimate the length of the review process in the most prestigious journals, which provide the clearest evidence of research quality.

An analysis conducted by Holden (2017) in the finance discipline (FOR 1502), for example, found that the median time to acceptance among the top five finance journals ranged from 9.9 months to 19.8 months. Half of all papers submitted to the top finance journals will therefore exceed these already lengthy review periods. In short, switching to by-lines (as opposed to using a staff census date) would provide an, often out-dated, representation of the true research performance of the staff employed at an institution. What is more, the impact of this effect would be felt most heavily on certain disciplines like business and – most concerningly for a research exercise aimed at assessing research quality – on the highest quality business journals as they tend to have the longest review periods.

Q3.31: *ERA adequately captures and evaluates interdisciplinary research. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer. a. If you disagreed with the previous statement, how could interdisciplinary research best be accommodated? Please describe.*

Disagree: It would appear that interdisciplinary or cross-disciplinary research, for example, areas such as Telehealth or Bioinformatics, are not captured within the FoR-code approach. Indicators for inter-disciplinary research would encourage and support innovations in this area.

Q3.32: My institution would meet ERA low-volume threshold in Indigenous studies at:

a. Two-digit? Yes/No. If you answered 'yes', please list which ones.

Yes, at the two digit level.

b. Four-digit? Yes/No. If you answered 'yes', please list which ones.

No.

Q3.33: In ERA, the best approach for evaluating Indigenous Studies is (choose one):

a. Using established ERA methodology i.e. the low-volume threshold would apply to the Indigenous Studies discipline and all its specific disciplines

Q3.34: What would be the advantages and/or disadvantages of your preferred approach for evaluating Indigenous studies in ERA? Please describe.

This would provide consistency in approach with other fields and disciplines.

ERA Process:

Q3.35: ERA should move to an annual collection of data from universities. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.

No. The impact on staff to produce these data would be far too onerous and the costs substantial, if on an annual basis. If it could be collected remotely and checked annually by the institution that may be an appropriate compromise.

Q3.36: What would be the advantages and/or disadvantages of an annual data collection? Please describe.

Strongly disagree. Deakin would recommend that ERA was an exercise undertaken ever six years. While the current gap between ERA's is to be five years, we would consider a six year interval would be more appropriate and would make comparison with previous rounds of ERA much more comparable, given that each previous ERA actually considered the previous six years' worth of data.

An annual collection of these data would create a major reporting burden and costs rather than streamline the process. An analysis of the cost of the ERA process to universities; i.e. the cost of providing the data, including staff time, would prove significant and in any new post-COVID-19 world it would be difficult to justify this expenditure and resourcing.

Q3.37: In future ERA rounds, should the volume of outputs submitted for each unit of evaluation be published?

a. Yes, Please explain your answer.

b. No, Please explain your answer.

The volume of research outputs is not a measure of research quality. Hence one argument would indicate that volume should not be used to compare and distinguish between universities. The contribution to a chosen field should be compared only on the quality of the research, in terms of publications, citations, design and peer review. However, the same discipline may vary substantially in size across Australian institutions, often in response to the strategic priorities of each institution. An assessment of 'above world standard' or 'well above world standard' therefore means different things depending on whether the area being assessed is small/niche or large. For this reason, the volume of outputs produced at the 4-digit FOR level should be reported for each institution alongside each institution's academic staffing levels (FTE) at the 4-digit For level. This shift would help to distinguish between high quality/niche research areas and high quality/high concentration research areas.

Q3.38: In future ERA rounds, research outputs should be published with their assignment to specific disciplines following completion of the round. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer. a. What would be the advantages? Please explain your answer. b. What would be the disadvantages? Please explain your answer.

Neither agree nor disagree. One could take the academic view of publish or perish. Deakin does not see any issues with publishing research outputs with their assignment to specific disciplines.

SECTION 4: ENGAGEMENT AND IMPACT ASSESSMENT (EI)

EI Overview

Q4.1: Considering that EI is a new assessment, to what extent is it meeting its objectives to:

- a. encourage greater collaboration between universities and research end-users, such as industry, by assessing engagement and impact? A very large amount; A large amount; A moderate amount; A small amount; Not at all. Please explain your answer.
- b. provide clarity to the Government and the Australian public about how their investments in university research translate into tangible benefits beyond academia? A very large amount; A large amount; A moderate amount; A small amount; Not at all. Please explain your answer.
- c. identify institutional processes and infrastructure that enable research engagement? A very large amount; A large amount; A moderate amount; A small amount; Not at all. Please explain your answer.
- d. promote greater support for the translation of research impact within institutions for the benefit of Australia beyond academia? A very large amount; A large amount; A moderate amount; A small amount; Not at all. Please explain your answer.
- e. identify the ways in which institutions currently translate research into impact? A very large amount; A large amount; A moderate amount; A small amount; Not at all. Please explain your answer.

A moderate amount. Universities such as Deakin are active in collaborating with industry and will continue to have a focus on applied and collaborative research with real societal impact. Highlighting a limited number of case studies that emphasise the biggest impact and engagement outcomes over an extended period of time shows a trend in terms of institutional behaviour, but consideration also needs to be given to the last interaction with the industry partner. Researchers have used a narrative approach when engaging with industry, however, the narrative tends to be specific to the industry in questions with the latest deliverables/outcomes.

A key point here is demonstrating to Government and the public that there is a definite return on investment. Hence public engagement and outreach are extremely important.

Deakin has over many years developed clear and appropriate processes and infrastructure to support research translation and show clear and definitive impact of our research.

Q4.2: *The EI objectives are appropriate for the future needs of its stakeholders. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.*

Agree. The EI objectives will be more appropriate than ever for the future needs of its stakeholders in the new post-COVID-19 world. The link between research funding and impact on industry in terms of translation of research into new technologies to aid the both state and federal economic recovery in the next three to five years will be essential.

Q4.3: *What impact has EI had on:*

- a. *the Australian university sector as a whole? Please describe.*
- b. *Individual universities. Please describe.*
- c. *researchers. Please describe.*

The introduction of EI coupled with the new reality of a post-COVID-19 world will impact the sector, individual institutions and individual researchers in terms of the future directions of their research and research strategies. While there will always be the need to innovate with blue-sky and fundamental research, the application and translation of that research into viable products to aid the economic recovery of Australia will likely be high on the agenda for most new funding opportunities.

Q4.4: *How do you, or your organisation, use EI outcomes? Please describe.*

Deakin has used EI outcomes as part of its wider library of innovation and successful engagement with industry to generate high value benefits and real world impact, to demonstrate our value to potential partners in various industry sectors, government and our communities.

Q4.5: *The EI outcomes are valuable to you or your organisation. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.*

Neither agree or disagree. We have existing and new mechanisms that create similar materials for engagement, although national visibility the time of outcomes has been useful.

Q4.6: How else could EI outcomes be used? Please describe.

The EI outcomes could provide an example for ECR's and others to follow on how impact can make a difference in their research. The administrative burden in collecting, maintaining and curating the EI information and the narrative over a much longer time scale that is required for publications and other metric driven outcomes, far outweighs and certainly negates any of the potential benefits. Mentoring of researchers in ways to develop applied research and translate it would be a far more time and cost-effective way to develop researchers and processes.

Q4.7: The current Engagement definition is appropriate. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree.

a. If you don't agree, what are your suggested amendments to the Engagement definition? Please describe.

Agree.

Q4.8: The current Impact definition is appropriate. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree.

a. If you don't agree, what are your suggested amendments to the Impact definition? Please describe.

Agree.

Q4.9: The current end-user definition (and other terms) are appropriate. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. If you don't agree, what are your suggested amendments to the end-user definition? Please describe.

The key weakness in the current research end-user definition is that it lacks a focus on inclusiveness specific to sub-populations of the community that are rarely, or sub-optimally, engaged as end-users.

Deakin suggests that specific detail be encompassed in the definition to read 'Community engagement, specifically for the mutually beneficial transfer of knowledge, methods and resources, should have a specific focus on ensuring diversity in gender representation, Indigenous status, disability, culture, language, socioeconomic disadvantage, and health literacy.' It would seem that the term 'community', with regard to being a research end-user, should be further clarified.

Q4.10: Are there other key terms that need to be formally defined? Yes/No.

There is a need to make the criteria for evaluating impact clearer, both to help researchers/institutions calibrate their understanding of low/medium/high impact and to guide the EIA panels in their assessment of research impact. The UK's REF, for example, specifies that impact should be evaluated on the basis of 'reach and significance' and the adoption of a similar assessment framework would hold merit for the EIA context.

Q4.11: Are the two-digit Field of Research codes the most appropriate method to define units of assessment for Engagement and Impact? Yes/No. Please explain your answer.

Yes. The 2-digit FoR codes are a good first sorting of research areas. However, research outputs are evaluated at the 4-digit FoR level, while research engagement/impact is only evaluated at the 2-digit FOR level. If the aim of EIA is to encourage greater collaboration between institutions and research end-users, then research engagement/impact should ultimately be given similar treatment to that of research outputs. Consequently, research engagement/impact should also be examined at the 4-digit FOR level.

Q4.12: Are there other ways to classify units of assessment in EI, for example, SEO codes? Yes/No. Please explain your answer.

Not in the STEMM subjects. In some HASS subjects an SOE code may provide additional information and may be a better indicator of interdisciplinary work and collaborations.

Q4.13: Should there be more or fewer units of assessment per university? More units of assessment; The same number as in EI 2018; Fewer units of assessment. a. How many and why? Please explain your answer.

More units of assessment: In some disciplines/at some institutions, the number of academic staff at the 2-digit FoR level can be substantial. Requiring only one impact case study submission per 2-digit FOR code consequently provides too narrow a lens through which to evaluate the magnitude of impact occurring within that FoR code. Following the

UK REF approach, the number of impact case studies should scale with the number of FTE within each 2-digit FOR code.

The average rating of all impact case studies submitted within each 2-digit For code could therefore be used to provide a global assessment of the level of impact occurring within those 2-digit For codes. This approach would provide several additional advantages:

- It ensures that impact case studies that span multiple 2-digit For codes can have the ratings associated with those case studies credited to each contributing 2-digit For code (e.g., as a weighted average).
- It provides a streamlined and more inclusive alternative to the single 'Interdisciplinary impact study' that can currently be submitted via EI.

Q4.14: The EI low-volume threshold should continue to be based on the number of research outputs submitted for ERA. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree.

Agree. In the UK's EI Exercise in 2014 (and again for the 2021 Exercise) one Impact case study was required for every eight academics submitted to the REF process. The aim was to balance the submission between fundamental and applied research, but there were a series of issues with this approach and lead to a number of people not being submitted to the process. Since the number of staff submitted also linked to the potential financial return, more staff meant the need for more impact studies, and more block grant.

Q4.15: The low volume threshold is set at the appropriate level. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Agree. The current level appears to align reasonably with sufficient levels of activity to generate a body of work that has significant impact on knowledge and translation.

Q4.16: Overall, the engagement indicator suite for the assessment of research engagement is suitable. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Explain your answer.

Deakin would like to revisit the assessment of research engagement. Lessons need to be learned from the 2018 EI and from other similar exercises from around the world e.g. UK's REF Impact and Engagement case studies.

Q4.17: The cash support from research end-users indicator using HERDC data is appropriate for the assessment of research engagement? Strongly agree; agree; neither agree nor disagree; disagree; strongly disagree. Please explain your answer.

Strongly disagree. Cash support is an important indicator of research engagement, but its absence does not necessarily mean that research engagement is not occurring. For example, certain sectors, such as NGOs, community groups, and other not-for-profits, have less cash resources available to directly fund research, yet research engagement still occurs through a broad range of activities such in-kind contributions, advisory board participation, other joint research activities.

While institutions have the opportunity to articulate these broader indicators of research engagement as part of the engagement narrative, it is important that EI evaluation panels recognise the diversity of pathways through which research engagement can occur, and that cash support is but one manifestation of research engagement. For this reason, having a clearer indication on the criteria by which engagement is being evaluated would be helpful.

Q4.18: The research commercialisation income is appropriate for the assessment of research engagement. Strongly agree; agree; neither agree nor disagree; disagree; strongly disagree. Please explain your answer.

Strongly disagree. While any return on investment in research through translation and commercialisation is the constant aim of universities engaged in a significant body of applied research, the time scales for this income can be long and extremely variable depending on the nature of the research and the impact. In addition, one big win on a single commercialisation opportunity or sale of IP could significantly skew this as a reliable indicator metric.

Q4.19/20: Are there additional/alternative metrics that would be appropriate across many or all disciplines? Yes/No. If you answered 'Yes', please outline the metrics. If you answered 'No', please explain your answer.

Yes. One possibility would be to measure the number of auditable Business Assists i.e. interventions where a piece of work has assisted an industry or industry partner solve an on-going issue (however small). Another would be to identify the number of publications that have an industry partner as a co-author.

Q4.21: Should any of the current Engagement metrics be redesigned? Yes/No. If you answered 'Yes', which ones and how?

Yes. It would be appropriate to reconsider all the Engagement metrics and evaluate how well they had worked in the 2018 EI Exercise.

Q4.22: The co-supervision of HDR students should be made an engagement indicator in future rounds of EI. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Disagree. Students working on an industry project/problem and not the co-supervisor of said PhD students, is the indicator of engagement. The aim is for students to have the best supervisory arrangements but not all industries have staff who are qualified research supervisors and can commit additional time to support the student. However, industries are generally very clear on the project or problem to be solved and provide great practical experience for the student. We do not recommend allowing industry research supervision to be about the number of industry supervisors, but to be about ensuring a great applied research opportunity and excellent training experience for the student.

Q4.23: In your opinion, are any of the ERA applied measures appropriate indicators of research engagement in EI?

a. Patents. Yes/No.

No. Patents do not necessarily imply engagement.

b. Research commercialisation income. Yes/No.

Yes. Income demonstrates engagement, but not all engagement generates income.

c. Registered designs. Yes/No.

No.

d. Plant breeder's rights. Yes/No. Please explain your answer.

N/A

e. NHMRC endorsed guidelines. Yes/No.

No.

Engagement narrative

Q4.24: The narrative approach is suitable for describing and assessing research engagement with end-users. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer. a. If you disagree, what alternative approach could be used to replace the narrative? Please explain your answer. If you are suggesting indicators, please be specific.

Agree. The validity and utility is limited by the extended length of time between engagement and outcomes, lack of defined criteria as to what constitutes a recognised and effective way of engagement and the inability to verify that such mechanisms were in place and activities occurred .

Q4.25: One engagement submission per broad discipline is sufficient for capturing the research engagement within that discipline. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Disagree. Research engagement can be a broad church so it is hard to capture with one engagement submission.

Q4.26: The engagement narrative needs to be longer. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Disagree. The current length requirements for an engagement narrative are fine.

Q4.27: Additional evidence is needed within the narrative. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer. a. If you agreed, what evidence should be provided? Please describe.

Disagree. Refer to comments above re lack of consistent, objective and verifiable evidence.

Q4.28: The narrative approach is suitable for describing and assessing impact. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer. a. If you disagree, what alternative approach could be used to replace the narrative? Please explain your answer. If you are suggesting indicators, please be specific.

Agree. However additional verifiable measures should be required and included.

Q4.29: One impact study per broad discipline is sufficient for capturing the research impact within that discipline. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Disagree. The majority of disciplines are multi-faceted and a single impact study may miss out whole areas of research. Research impacts can be very varied even within a given discipline.

Q4.30: The impact narrative needs to be longer. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Disagree. The current length requirements for the impact narrative are fine.

Q4.31: There is a need for additional evidence to be provided within the narrative. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer. a. If yes, what evidence should be provided? Please explain your answer.

Agree. A weakness of the current EI is that impact submissions do not need to provide evidence to corroborate the claimed impact. To remedy this issue, and following the UK's REF, each impact submission should be required as part of its own distinct section to provide a set of external sources (e.g., reports, factual statements from beneficiaries) that can corroborate the claimed impact.

Q4.32: In your opinion, are there quantitative indicators that could be used to measure the impact

As has been pointed out earlier, indicators such as commercialisation income, business assists, patents, spin-out companies are all measures of impact but the timescales they operate over are very different.

Q4.33: The narrative approach is suitable for describing and assessing approach to impact. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer. a. If you disagree, what alternative approach could be used to replace the narrative? Please explain your answer. If you are suggesting indicators, please be specific.

Agree. This also is a limited approach for similar reasons: Re: lack of consistent, objective and verifiable evidence.

Q4.34: One approach to impact narrative per broad discipline is sufficient for capturing the activities within that discipline. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Disagree. (see above Q4.31).

Q4.35 The approach to impact narrative needs to be longer. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Disagree. (see above Q4.31).

Q4.36 There is a need for additional evidence to be provided. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Agree. More evidence of the impact should be provided and possibly further impact narratives are needed to capture impact across a broad discipline.

Q4.37 Would there be benefit in combining engagement and approach to impact? Yes/No. Please explain your answer.

No. These are two rather different things that may or may not be tenuously related in certain cases.

Q4.38: The engagement rating scale is suitable for assessing research engagement. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Q4.39: The descriptors for the engagement rating scale are suitable. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Q4.40: *The impact rating scale is suitable for assessing impact. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.*

Q4.41: *The descriptors for the impact rating scale are suitable. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.*

Q4.42: *The approach to impact rating scale is suitable for assessing approach to impact. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.*

Q4.43: *The descriptions for the approach to impact rating scale are suitable. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.*

Answers to all the above questions (Q4.38 – Q4.43) are strongly disagree. The existing Low/Medium/High scale is simply too broad. The introduction of a 5-point scale would be a better outcome.

EI interdisciplinary research

Q4.44: *Should EI continue to include an interdisciplinary impact study in addition to the two-digit Fields of Research impact studies? Yes/No. Please explain your answer.*

If the number of impact studies are to be increased, then interdisciplinary impact studies would be a welcome addition as much of the impact described occurs at the interface between disciplines.

EI and Aboriginal and Torres Strait Islander research

Q4.45: *Should the EI low-volume threshold be applied to the unit of assessment for Aboriginal and Torres Strait Islander research in EI 2024 with the option to opt in if threshold is not met? Yes/No. Please explain your answer.*

Yes. Similar rationale as for Indigenous Studies.

Q4.46: *Should the unit of assessment for Aboriginal and Torres Strait Islander research include engagement in EI 2024? Yes/No. Please explain your answer.*

Yes. This would better recognise what is difficult and time consuming but essential work to make material contributions in this field.

SECTION 5: OVERARCHING ISSUES COMMON TO BOTH ERA AND EI

Q5.1: *How often should ERA occur? Every three years; Every five years; Other, please specify. Please explain your answer.*

The ERA and EI exercises should take place **every six years**. This is a significant period over which to measure changes in the ERA exercise, but also a more realistic time scale for an evolution from fundamental research through translation to impact. It also reduces the burden on institutions repeating this both exercises only every six years. The reason for choosing every six years, instead of five years or seven, is for a more direct and clear comparison with the data from the previous ERA exercises, which also looked over a six year period but with an overlay of three years each time.

Q5.2: *What impact would a longer assessment cycle (i.e. greater than three years) have on the value of ERA results, particularly in the intervening years? Please explain your answer.*

While one possibility might be that institutions may not be as motivated to maintain or continue to improve their performance, a longer period allows for new trends to evolve along with new areas of research. Note both research and impact tend to evolve on a timescale much longer than three years.

Q5.3: *How often should the EI assessment occur? Every three years; Every five years; Other, please specify. Please explain your answer.*

Every six years for the same reasons outlined in Q5.1.

Q5.4: What impact would a longer assessment cycle (i.e. greater than three years) have on the value of EI results, particularly in the intervening years? Please explain your answer.

A period of six years is a significant time for new impact to have potentially evolved from fundamental research undertaken in this assessment period. Any shorter time period and it is highly likely that no impact will have yet developed, except from the time before the current assessment period.

Q5.5: ERA and EI should be combined into the one assessment. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer. a. What would be the advantages and/or disadvantages. Please explain your answer.

Strongly disagree. Research excellence and research impact are very different things as defined in the ERA and EI exercise. Research excellence and impact on knowledge are assessed and captured by ERA whereas impact in real world terms is captured by EI.

Q5.6: Are there other ways to streamline the processes to reduce the cost to universities of participating in ERA and EI? Yes/No. Please explain your answer.

While it may be possible to use AI and other new technologies to help improve the collection and organisation of publications and citations etc., there is still a significant need for peer, review primarily across the HASS disciplines. This takes both time and is costly, however, it is certainly essential to have that level of expert opinion.

Q5.7: In your view, what data sources could ERA utilise? Please explain your answer.

For the STEMM subjects, the ISI Clarivate database would seem to be appropriate for publications and citations. As for the HASS and Business and Law subjects, I can't comment on the value of this database.

Q5.8: In your view, what are the most time-consuming elements of an ERA submission? Please describe.

a. Are there efficiencies that could be introduced? Yes/No. Please describe.

The sheer effort in collecting together the data every three years and assigning FoR codes was the time-consuming element. Going forward the universities should both collect and curate the ERA and EI data as they progress. As has been pointed out, the EI data in particular, has a longer time-line to impact and the need for regular curation.

Q5.9: In your view what are the most time-consuming elements of an EI submission? Please describe. a. Are there efficiencies that could be introduced? Yes/No. Please describe.

The most time-consuming element is selecting and writing the statements for the EI submission. While these can really only be done at the time of an ERA/EI exercise, if the data was being collected and curated throughout the period of assessment, it would certainly aid the production of these impact case studies.

Recommendations:

- **Recommendation 1:**

The time scale of future ERA and EI exercises should be at a six year interval. This is a sufficiently long period so the data collection and preparation of the narratives, impact cases studies etc. is not a re-occurring burden on institutions, it is a significant period over which to measure EI and offered the most direct comparison with the four previous ERA exercises.

- **Recommendation 2:**

Devise a system for expert peer review that does not encourage universities to use a citation-based metrics system to drive publication practices in HASS. Have meaningful measures of quality and impact that require expert judgements. Criteria such as originality, significance, scholarly rigour should be included.

- **Recommendation 3:**

The existing Low/Medium/High scale for ranking of the EI is too broad. The introduction of a five-point scale would be a significantly better, more accurate and robust outcome.

- **Recommendation 4:**

Remove the limitation of the 66 per cent rule to help with interdisciplinarity issues. At the same time, allow finessing of data sets so that publications clearly out of the field, are not counted in the field.

- **Recommendation 5:**

We suggest that data linkage via citation analyses, ORCID IDs, HERDC data etc. can be integrated and automated to reduce the burden of data collection on each university.