

Appendix D—Summary of Questions

Section 3—Excellence in Research for Australia

ERA policy

Value of ERA

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

- a. Continue to develop and maintain an evaluation framework that gives government, industry, business and the wider community assurance of the excellence of research conducted in Australian higher education institutions. *A very large amount; A large amount; A moderate amount; A small amount; Not at all. Please explain your answer.***

A very large amount regarding citation disciplines. The assessment of citation disciplines is data-driven and transparent. This provides a reliable, sector-wide assessment of these disciplines and with a focus on excellence that provides insight into strengths of individual institutions benchmarked against world standards.

A small amount/not at all regarding peer review disciplines. The methodology employed to assess performance in peer review disciplines is flawed (see response to 3.17) and, as such, does not provide assurance of an accurate assessment of excellence in these disciplines.

- b. Provide a national stocktake of discipline level areas of research strength and areas where there is opportunity for development in Australian higher education institutions. *A very large amount; A large amount; A moderate amount; A small amount; Not at all. Please explain your answer.***

A very large amount. The assessment of citation disciplines provides a robust assessment of these disciplines and reliably identifies both strength and opportunity for development in these areas, enabling institutions to differentiate themselves.

A small amount/not at all regarding peer review disciplines. This assessment methodology is flawed resulting in an inconsistent and unreliable assessment. Excellence is not reliably identified and the 'stocktake' is thus incomplete.

- c. Identify excellence across the full spectrum of research performance. *A very large amount; A large amount; A moderate amount; A small amount; Not at all. Please explain your answer.***

A small amount/not at all. Flaws in the peer review assessment process (see response to 3.17) do not provide an accurate and robust assessment in these discipline areas. As such, excellence is identified with confidence only in citation disciplines and not across the full spectrum of research.

- d. **Identify emerging research areas and opportunities for further development. A very large amount; A large amount; A moderate amount; A small amount; Not at all. Please explain your answer.**

A very large amount. The assessment of citation disciplines provides a robust assessment of these disciplines and reliably identifies both strength and opportunity for development in these areas, enabling institutions to differentiate themselves.

A small amount/not at all regarding peer review disciplines. This assessment methodology is flawed resulting in an inconsistent and unreliable assessment. This prevents performance from being tracked accurately over time.

- e. **Allow for comparisons of research in Australia, nationally and internationally, for all discipline areas. A very large amount; A large amount; A moderate amount; A small amount; Not at all. . Please explain your answer.**

A very large amount regarding citation disciplines. The assessment of citation disciplines is data-driven and transparent. The use of citation/RCI provides a reliable assessment that allows institutions to benchmark their performance both against Australian institutions and internationally.

A small amount/not at all regarding peer review disciplines. The methodology employed to assess performance in peer review disciplines is flawed and, as such, does not provide assurance of an accurate assessment of excellence in these disciplines.

- 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. Please explain your answer.**

Disagree. There is a misalignment between ERA's objective to identify research excellence and the requirement that the entirety of an institution's output be submitted for assessment. This requirement places an inordinate and unnecessary burden on universities in preparing the submission where the same assessment might be performed using only a sample of the top publications produced by an institution.

Also, given its focus on research quality, ERA results should contribute to the determination of research funding allocations for universities.

- a. **If you disagreed with the previous statement, what should the primary purpose of ERA be going forward? Please explain your answer.**

Recommendation: Focus on research excellence by requiring only a percentage of outputs to be submitted for assessment.

The primary focus of ERA should be on research excellence. The entirety of an institute's output need not be assessed to identify excellence; rather, excellence might be demonstrated through the submission of a percentage of outputs. This revised approach would have the concomitant effect of reducing the significant workload involved with managing large numbers of

publications for submission (see response to 5.8). Assessments in the humanities and social sciences could be further simplified by basing them on monographs and journal articles since it is difficult to imagine how book chapters would deliver an assessment different from the combined influence of books and journal articles.

Recommendation: Translate results into research funding.

In order to improve on the application of the outcomes of this national assessment, results should translate into research funding. This would create a system that rewards quality output and research excellence whilst supporting further growth and development. It would also support universities to differentiate themselves and would provide an incentive for growing those areas of specialisation in which they have identified strength.

Q3.3 What impacts has ERA had on:

a. the Australian university research sector as a whole

ERA's impact on the sector has been positive: it has clearly driven improvement. The assessment forces institutions to think strategically in identifying where their excellence lies and to invest selectively. The extent of this benefit has, however, been limited by flaws in the assessment methodology (peer review), which impedes identification of the rate of improvement in peer-review disciplines.

b. individual universities

Positive: Universities have been able to identify areas of strength (particularly in the citation disciplines, less reliably in peer review disciplines - see discussion of methodology at 3.15) and in so doing, many have been successful in differentiating themselves, developing areas of specialisation.

Negative: The submission of all outputs requires an unnecessarily significant workload due to the volume of submissions.

c. Researchers

Positive: Researchers, guided by the assessment principles, have in many cases developed an enhanced understanding of research excellence and have refocused attention on producing quality research rather than quantity of research.

Negative: There is a significant demand on the time of researchers engaged as peer reviewers (placing a secondary burden on Institutions for the period of review).

d. Other?

Please explain your answers.

Q3.4 How do you use ERA outcomes? Please describe.

ACU uses ERA to:

- 1) plan for continued development and improvement in our areas of strength;

- 2) identify those areas in which we need to refocus effort or reduce investment;
- 3) report on performance;
- 4) track performance over time.

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

Strongly agree regarding citation disciplines, which provides a robust and accurate assessment enabling ACU to benchmark its performance and identify areas for refocused effort.

Strongly disagree regarding peer review disciplines. As an institution with one of the highest proportions of peer review codes in the sector, flaws in the assessment of these disciplines have reduced the overall benefit of the assessment for ACU.

Benefit to institutions could be increased by using ERA to determine the allocation of research funding (see response to 3.2a)

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

Please refer to suggestions relating to improved assessment methodology (3.18 particularly).

ERA methodology

ERA methodology at a glance

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

Strongly agree regarding citation disciplines. The assessment of citation disciplines is data-driven and transparent. This provides a reliable, sector-wide assessment of these disciplines and with a focus on excellence that provides insight into strengths of individual institutions benchmarked against world standards.

Strongly disagree regarding peer review disciplines. The methodology employed is flawed and, as such, does not provide assurance of an accurate assessment of excellence in these disciplines.

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

As above. The citation assessment methodology is robust and reliable.

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

The peer review methodology is flawed in that it professes to rely on the expert opinion of discipline experts, but: 1) the burden placed on reviewers is inordinate and therefore cannot result in a reliable assessment, 2) reviewers rely on independent judgement, based on varying degrees of expertise, as to what constitutes high-quality publication; and 3) the size of the assessment panels

results in insufficient expertise in particular disciplines, thereby reducing the reliability of the assessment.

The low volume thresholds are too low, resulting in an assessment of areas in which an institution may be intermittently active or largely driven by an individual or two rather than a genuine research concentration. Volume thresholds should also be adjusted to reflect different rates of publication in different disciplines.

The submission requires that all outputs be submitted and adds an unnecessary burden to the process. Unless peer-review is removed the amount of material to be assessed needs to be significantly reduced.

Citation analysis methodology

Q3.10 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. Citation analysis is a robust, quantitative, transparent method for assessing the quality of research in those disciplines with a citation rate high enough to determine a meaningful RCI.

Q3.11 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?

No. The evaluation is not consistently robust across disciplines and fails to address differences between them. Two assessment categories - peer review and citation - are insufficient for addressing variations in publishing profile across what is an identified spectrum of disciplines. It is a guiding principle of ERA that the assessment process be flexible (i.e. that it is able to be applied across a range of disciplines). This is a shortcoming of the existing assessment. What is more, the methodology for peer review is flawed and, as a result, citation scores are not commensurable with the scores applied to peer review disciplines.

Q3.12 What are the strengths of the citation analysis methodology? *Please describe.*

Citation analysis methodology is transparent and data-driven, allowing for an unbiased and consistent assessment of the performance of institutions in those codes assessed in this way. Indeed, the strength of this approach is recognised by many international ranking systems. CWTS Leiden, Shanghai Academic Ranking of World Universities (ARWU), and Times Higher Education (THE) all include a significant element of citation analysis among their metrics.

Q3.13 What are the weaknesses of the citation analysis methodology? *Please describe.*

1. The citation score/RCI of a unit of assessment can be inflated by a single highly cited paper. This would be a particular concern where volume is low, but the present analysis does appear to monitor this (i.e the identification of papers with an RCI of 8 or greater). There should be consideration of removing from the assessment those papers with hundreds of authors and huge RCIs, such as the Global Burden of Disease papers).

2. Citation analysis requires a minimum number of citations in order to derive a meaningful RCI. It cannot therefore be used as the only means of assessment across the board. We maintain, however, that an element of citation analysis can have benefits for the assessment of almost all fields of research.

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

As above in answer to 3.13

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

Peer review methodology

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 disagree. The peer review assessment methodology is flawed and, as such, scores awarded to peer review codes are not comparable with the performance of citation discipline codes. That is, the evaluation is not consistently robust across disciplines and fails to address differences between them. Research in education could be assessed entirely by citation. Note that nine Australian universities appear in the world's top 100 in the Academic Ranking of World Universities. You would not know that from looking at ERA results. This misalignment of outcomes demonstrates the inadequacy of the current assessment methodology for education.

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

When used in isolation, peer review is a misleading and flawed methodology; it could, however, with appropriate changes to the way in which it is conducted, provide a useful additional measure in those disciplines where citations are not plentiful.

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

1. Highly subjective

Panel members see the publication profile of each institution in a field of research in the form of a list of the journals in which it most frequently publishes. This is intended to help inform the assessment, but panel members are not given any indication of the quality of those journals. Reviewers' evaluation of the quality of each journal is subjective and often limited by lack of knowledge of particular disciplines.

2. Panel size and composition

The above concerns around the subjectivity of the peer review of journals is compounded by the impact of the size of each assessment panel. Panels are large and with broad expertise between panel members. Members therefore cannot have the specialised knowledge required to make an accurate assessment of the top journals in a discipline area, nor of the quality of research in a discipline that may be somewhat removed from their own.

3. Demand on reviewers

Whilst 30% of a submission is currently nominated for peer review, this often comprises an inordinate number of outputs for a reviewer to assess. The burden on individuals engaged in peer review is significant and the expectation that they review each output unfeasible.

Q3.18 Can the peer review methodology be modified to improve the evaluation process while still adhering to the ERA Indicator Principles? Yes/No.

Yes.

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

The peer review methodology is significantly flawed in that it is highly subjective, employs assessment panels of unwieldy size and unfocused expertise, and places inordinate burden on reviewers. We maintain that peer review should only be maintained *if* the submission process is changed to require a reduced proportion of outputs be submitted (see response to Q3.26). This will reduce the burden on institutions and reviewers. We also maintain that the number of review panels must be increased (for example, through the inclusion of more international reviewers and more discipline experts) and the issue of subjectivity should be addressed through the inclusion of data to guide and supplement the assessment process, including citations, an elite journal list (e.g. top 5% or 10% of journals in a field of research), and impact factors.

The four recommendations below are deemed essential to address these flaws in the peer review methodology.

Recommendation 1: *Introduce a blend of citation and peer review to those disciplines currently assessed by peer review.*

Recommendation 2: *Introduce a list of 'elite journals' to support the peer review process.*

Recommendation 3: *Include Impact Factor (IF) as a metric to guide peer review.*

Recommendation 4: *Increase the number and specialisation of peer review panels, including with stronger international representation.*

These recommendations are elaborated on below.

Recommendation 1: *Introduce a blend of citation and peer review to those disciplines currently assessed only by peer review.*

As recent ERA rounds have shown, the peer review disciplines generally have been artificially undervalued, relative to the citation disciplines, with skewed outcomes as a consequence.

As such, it is proposed that the assessment of a number of peer review disciplines—particularly those in which the level of citation is high enough to be meaningful—would benefit from the introduction of an element of citation analysis or, otherwise, from a transition wholly to citation analysis.

In addition to making the assessment of these disciplines more nuanced, the introduction of citation analysis would align peer review disciplines more closely with the way international institutional rankings are constructed (for example, Times Higher, ARWU, and CWTS Leiden) all of which are strongly influenced by either Scopus or Web of Science citation metrics rather than the subjective opinions of experts.

Detail of the methodology of two International Ranking Systems is provided following as guidance.

1. *Times Higher Ranking by Subject*

The overall THE methodology is carefully recalibrated for each subject, with the weightings changed to suit the individual fields. The following table lists the weightings for education, arts and humanities, business and economics, and psychology. All but the latter are considered peer review disciplines for the purposes of ERA. Psychology is included as a reference point.

	Education	Arts & Humanities	Business & Economics	<i>Psychology</i>
Teaching: the learning environment	32.7	37.4	30.9	27.5
Research: volume, income and reputation	29.8	37.6	32.6	27.5
Citations: research influence	27.5	15	25	35
International outlook: staff, students and research	7.5	7.5	9	7.5
Industry income: innovation	2.5	2.5	2.5	2.5

Table 1. THE Subject Ranking percentage weighting of metrics by subject area

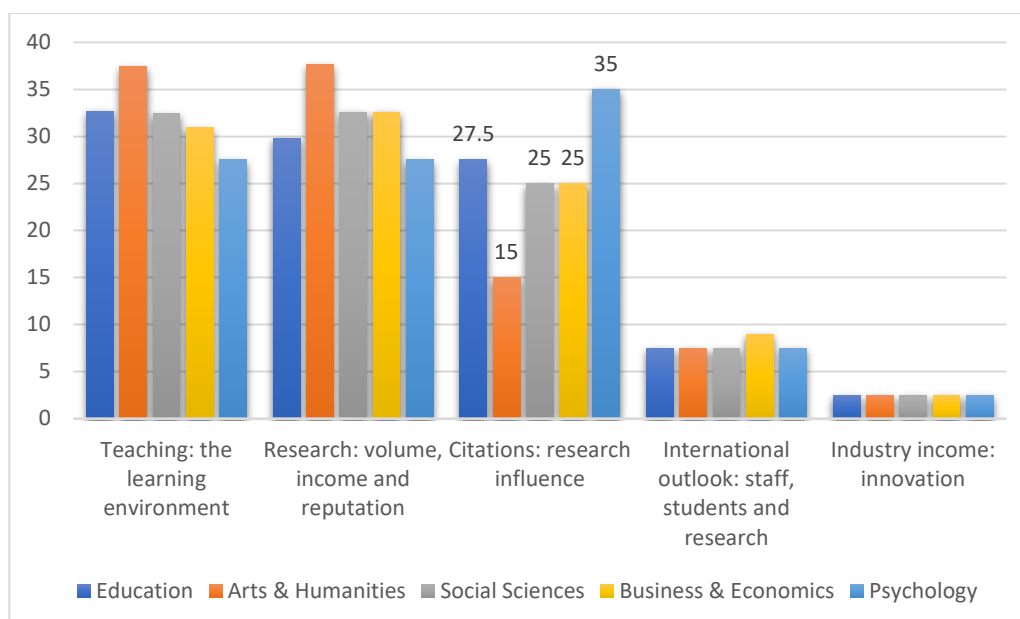


Chart 1. THE Subject Ranking percentage weighting of metrics by subject area

Looking to 'citations: research influence' by subject area, it is plain that value is seen in awarding this measure some weighting across all subject areas. Psychology is awarded the highest weighting for this metric in response to the higher relative rate of citation in this area; yet it comprises a proportion of the calculation for all subject areas, adjusted to suit the individual field.

2. ARWU Ranking by Subject

The ARWU metric and weighting system for the same general subject areas surveyed above is provided below, noting that for ARWU, psychology is considered a social science (under which all the below fall). ARWU does not rank 'arts & humanities' but sociology is included here for reference. ARWU, interestingly, does not adjust the weightings within the social sciences to the extent that THE does. What is significant, though, is that this ranking system uses entirely objective measures even for these subject areas considered peer review disciplines by ERA and that it produces a robust and reliable ranking in doing so. Indeed, one of the differences between these weightings and those for hard sciences is the omission of the Award metric in acknowledgement that this is not a reliable or useful measure for these areas. Bibliometric measures are maintained as a valuable measure across all subject areas.

	Education	Sociology	Economics	Psychology
Q1: Papers in Q1 Journal Impact Factor Quartile	48%	48%	48%	48%
CNCI: Ratio of citation of papers to world average	16%	16%	16%	16%

IC: International collaboration by proportion of co-authored papers	3%	3%	3%	3%
Top: Number of papers in top journals	32%	32%	32%	32%
Award: Staff with significant awards	0%	0%	32%	0%

Table 1. ARWU Subject Ranking percentage weighting of metrics by subject area

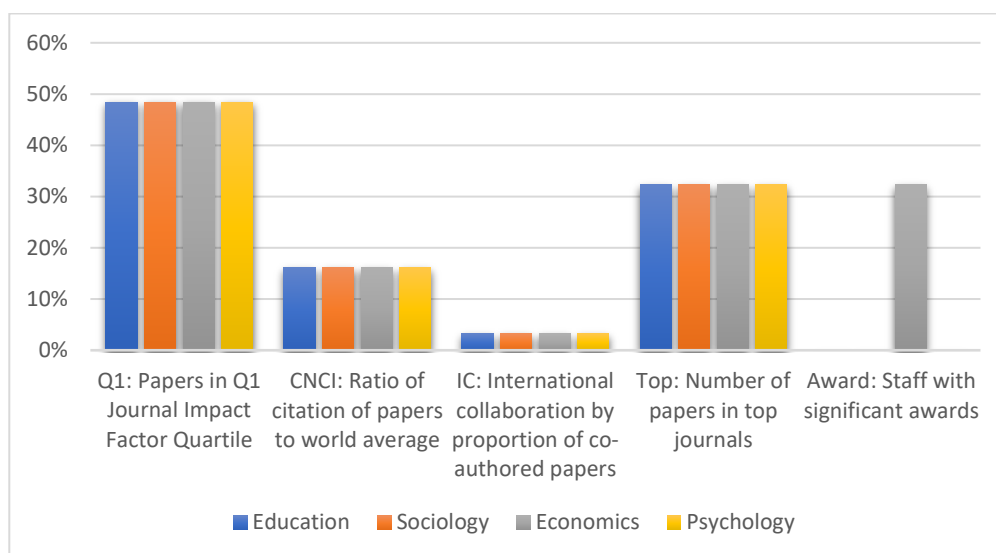


Chart 1. ARWU Subject Ranking percentage weighting of metrics by subject area

Example 1: Education

FOR39-Education has an observed rate of citation higher than some citation discipline FORs, including FOR1110-Nursing (ANZSRC 2004), yet it remains a peer review discipline. Significantly, improvement in ERA scores in education over the history of ERA is nearly non-existent. By contrast, since 2010 the number of Institutions receiving a score of 5 in 1110-Nursing has increased by 566% (from three to 20). This disparity between strength in nursing and strength in education is not reflected by Australian institutions' performance in international rankings. In 2020, for example, nine Australian institutions appeared in the global top 100 in ARWU's subject ranking for education. It is evident that peer review does not provide a reliable and robust assessment of strength in education against world standards by employing the existing peer review methodology. Nor does it provide institutions with the opportunity to improve performance, one of the guiding principles of ERA.

The assessment FOR39 would therefore benefit from the transition to citation analysis (or, at least, to a combination of peer-review and citation analysis). We would argue that especially in this case peer-review involves a large and unnecessary workload burden.

Example 2: Commerce, Management and Tourism

FOR35-Commerce, Management and Tourism (ANZSRC 2004) has a high rate of citation and many discipline associations (including the Australian Business Deans Council and certainly the ABS in the UK) are already using a mixed peer review-citation assessment model when they rank their journals

Most of the top research-led Business Schools globally (for example, NUS and most schools in the US) rely on and report citation scores as key indicators. They rely less, if at all, on subjective opinion rankings. By their very nature, the latter suffer from a "self-selection" bias towards journals the experts publish in. The assessment of FOR35 would therefore benefit from the transition to a combination of peer review and citation analysis (or, alternatively, a transition to 100% citation analysis).

Recommendation 2: *Introduce a list of 'elite journals' to support the peer review process.*

Peer reviewers currently rely on independent judgement as to what constitutes high-quality publication. A list of elite journals would be useful in the peer review disciplines and would provide a reliable and transparent assessment of the number/proportion of publications that appear in top-rated journals. This proposed approach replicates the methodology employed by some international rankings systems. For example, the ARWU Shanghai ranking identifies a very select number of top journals in a discipline in its ranking methodology. It would perhaps not make sense to be as selective as this, but ERA would benefit from an A* list or a list identifying the top 5% (or perhaps 10%) of journals. It makes no sense in our view to have every journal listed, as in the old system, if this exercise is about excellence in research

Recommendation 3: *Include Impact Factor (IF) as a metric to guide peer review.*

Impact Factor (IF) can be a useful and robust measure of the performance of a journal within its field. Whilst IF might be less useful in comparing performance across disciplines, within a 4-digit FOR code, for example, it provides a useful objective point of comparison of the quality of journals. IF would be a useful and simple means of introducing metrics into the peer review process. Rather than relying on their own knowledge of journals across a broad range of sub-disciplines, sometimes far removed from their own expertise, experts would be able to refer to IF as one indication of the strength of an institution's publication profile. This introduction would have the immediate effect of improving consistency in the peer review process, addressing to an extent this Institution's

concerns around the subjectivity of the review process and the size of review panels.

Recommendation 4: *Increase the number and specialisation of peer review panels.*

Smaller panels but with greater expertise will increase reliability of assessment in peer review disciplines.

Number of panels

The number of review panels for ERA should be increased to improve alignment of the expertise of panel members with the materials they review. Increasing the number of panels will have the concomitant effect of reducing the inordinate workload for panel members.

Presently, peer review panels are broad and the number of panels too low (eight). As an example, the UK REF employs a total of 34 sub-panels – one for each unit of assessment. With the introduction of Indigenous Studies to ANZSRC 2020, there are only 23 units of assessment (UoA) at the 2-digit level in ERA. The UK's allocation of a panel to each UoA is significant; for example, whilst sociology is its own UoA in the UK the ANZSRC code for this discipline is much broader (44-Human Society) and encompasses disciplines as distinct as anthropology, gender studies, and political science, for example.

Panel size

In concert with increasing the number of panels, panel membership should be reduced. REF sub-panel membership is comprised of Chair, Deputy-Chair, and Members (of which there are in the region of eight). Additional assessment phase members - of which there are in the region of six to 12 - support the review process. ERA Research Evaluation Committees (RECs) currently comprise in the region of 15 to 20 members (plus reviewers akin to REF assessment phase members). Taking the REF as a model, there is scope to reduce the size of each ERA peer review panel (REC).

What should be considered is that panels remain of a size that prevents their assessment from becoming idiosyncratic. One function essential to the assessment process is that it produce commensurate rankings across disciplines. That is, a score of 5 in theology should equate with a score of 5 in genetics – the excellence measured against world benchmarks should be equivalent. As such, more experts will need to be involved in the review process.

International representation

International representation on panels is important for an assessment that employs international benchmarks to allocate a score that indicates performance against world standards. Not only will this improve alignment between the pool of outputs under review and the expertise of panel members, but it creates a larger pool of possible reviewers, addressing the issues of burden on Australian reviewers and idiosyncratic review.

Contextual indicators

- 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 disagree. Given, as stated in the consultation paper, 'their presence or absence has virtually no effect on the rating given to a unit of evaluation', including these indicators increases the burden on university for little or no yield.

- 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.**

Neither agree nor disagree. Given, as stated in the consultation paper, 'their presence or absence has virtually no effect on the rating given to a unit of evaluation', this indicator has little value for citation disciplines. For peer review, publishing profile could be useful if linked to a list of journal ratings, but a list of journal ratings should include the very best (that is, the journals that were ranked A* in the old list or A* and A). That would give panel members outside of a given field a reasonable overview of how frequently the researchers in a particular FOR are publishing in the very best peer-reviewed outlets.

- 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 disagree. Research income is not an output or outcome of research but an input and therefore not a reliable indicator of the quality of research. Recommendation: Remove the research income indicator.

- Q3.22 The applied measures are still relevant to ERA:**

- a. Patents. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.**

ACU has no view on the value of this indicator; although, removing it would assist in streamlining the submission process.

- b. Research commercialisation income. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.**

Strongly disagree. Commercialisation income is, like income, a research input and not a reliable indicator of the excellence.

- c. Registered designs. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.**

ACU has no view on the value of this indicator; although, removing it would assist in streamlining the submission process.

- d. Plant breeder's rights. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.**

ACU has no view on the value of this indicator; although, removing it would assist in streamlining the submission process.

- e. **NHMRC endorsed guidelines. *Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.***

ACU has no view on the value of this indicator; although, removing it would assist in streamlining the submission process.

ERA rating scale

- Q3.23 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.***

Agree. The existing scale of 1-5 is simple and easily interpreted. There are, however, ways in which the scale might be extended to identify outstanding performance (please see response to 3.24a).

- Q3.24 Noting that 90% of units of evaluation assessed in ERA 2018 are now at or above world standard, does the rating scale need to be modified to identify excellence? *Yes/No.***

Yes.

- a. ***If you answered, 'Yes', please explain how the rating scale can be modified to identify excellence.***

The rating scale of 1-5 should be retained in order to track performance over time. In response to the variation in performance within the ERA5 bracket, the ARC should consider the addition of a 5* rating (world-leader). This rating incentivises and recognises exceptional performance without departing from the existing scale.

ERA low-volume threshold

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

Strongly disagree. The low-volume threshold is too low and results in institutions being assessed at times in codes in which they are intermittently or minimally active across the reference period. Also, it can result in an assessment being of the performance of just one or two researchers.

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

Recommendation: Increase the low-volume threshold in 2-digit codes (to minimum 150) and in 4-digit codes (to minimum 100). It should be noted that this is a proposed minimum and that in cases where there are a large number of 4-digit codes within a 2-digit code, the low volume threshold should be increased again. This will help to drive focus on areas of strength.

Also, institutions should be required to submit only a percentage of publications for assessment.

Recommendation: Require that only 20% of an institution's output be submitted for assessment.

ERA staff census date

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

Census date. This approach employs a more current dataset and avoids the significant lag time associated with the time it takes to publish. Also, this approach reduces the administrative burden involved with tracking publications and verifying by-lines.

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

Whilst there have been criticisms of the census date approach, this is the most appropriate method for claiming research outputs. It provides a clear and concise methodology whilst imposing minimal additional burden on institutions.

One limitation is that the census date approach enables an institution to include the outputs of a researcher who may have only a fractional appointment at that institution. Having people affiliated with two or more institutions, however, increases collaboration and so is something to be encouraged.

Therefore, whilst this limitation is acknowledged, it might be addressed with an alternative amendment to the methodology.

Recommendation: Introduce the requirement for institutions to report where researchers have a fractional appointment with another institution.

A further limitation of this approach relates to the definition of staff. The definition includes honorary and adjunct appointments, requiring that the outputs of these appointees also be submitted for assessment. This approach is flawed in that these appointments often involve minimal interaction with the research activity of an institution.

Recommendation: The publications of honorary and adjunct appointments should not be included in the submission data.

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

No. Concerns also exist with the by-line approach. Please see below.

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

The by-line approach introduces significant lag time between the point at which a researcher commences at a University and the point at which their publication record is reflected in the profile of the University. Secondly, assuming that the publications of researchers who have departed an institution would be included if they are published in the reference period, this approach would produce an assessment based on outdated/historical data, rather than the research environment as it stands at the point of submission. This approach also imposes a significant burden on institutions in verifying publications and tracking by-lines.

ERA interdisciplinary research and new topics

Q3.31 ERA adequately captures and evaluates interdisciplinary research. Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.

ERA does adequately capture interdisciplinary research; but it seems that it is not the intention of ERA to assess/evaluate research as interdisciplinary (i.e, whether or not an institution produces interdisciplinary research and the extent to which it does is not reported). There is, however, no indication that ERA discourages interdisciplinary research as its existence is adequately addressed by apportionment.

- a. **If you disagreed with the previous statement, how could interdisciplinary research best be accommodated? Please describe.**

ERA and Indigenous research

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.**

ACU proposes that the ARC consider increased low-volume thresholds, but, for Indigenous studies, the threshold would probably need to be set at the lower end of the scale.

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

As above. Currently, ACU is most likely to reach threshold in 4501-Aboriginal and Torres Strait Islander culture, language and history.

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**

No.

- b. **For Aboriginal and Torres Strait Islander studies by combining low-volume disciplines into single units of evaluation**

No.

- c. **For Aboriginal and Torres Strait Islander studies by combining low-volume disciplines into two units of evaluation (one unit comprising Humanities, Arts, and Social Sciences disciplines and one unit comprising Science, Technology, Engineering and Mathematics disciplines)**

Yes. Discipline clusters will need to be aggregated as such to be assessed but for peer review disciplines, it is important that the flaws in the peer review methodology are addressed (see response to 3.18).

- d. **Other. Please describe.**

No.

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

The advantage of this approach is that it would enable institutions with a lower volume in Aboriginal and Torres Strait Islander studies but where that output is of a high quality to be recognised for that excellence. A disadvantage is that it aggregates disciplines and so will not provide a granularity in its assessment

(i.e., the strength of Australia's research in Indigenous sociology will not be differentiated from Indigenous history research).

ERA process

Collection of ERA data

- 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.***

Strongly disagree. Replicating submission on an annual basis means that processes need to be performed annually instead of once per submission increasing the overall burden on institutions. This will increase an already significant workload and is an unsustainable approach for universities.

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

Disadvantage 1: Given ACUs position that only a percentage of outputs should be submitted for assessment, this approach is not feasible. Our proposed approach requires that an institution review the submission in its entirety at the end of a reference period in order to determine which publications will be included in the submission.

Disadvantage 2: The overall administrative burden would not be reduced but rather increased. Annual submission requires replication on an annual basis instead of the university preparing its submission every three years. Further, the annual data would need to be repeatedly updated and checked to remove the outputs of staff who have departed the institution and those who have since arrived.

Publication of ERA data

- Q3.37** In future ERA rounds, should the volume of outputs submitted for each unit of evaluation be included in the National Report?

a. ***Yes, Please explain your answer.***

Yes. Publishing submission data improves transparency and allows institutions to benchmark their performance.

b. ***No, Please explain your answer.***

N/A

- 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.***

Strongly disagree. Releasing this level of detail will open ERA to futile, endless, and no doubt, in some instances, ill-informed argument. It will undermine the process and create opportunity for assessments to be criticised. Certainly, there is merit to maximising transparency, but this must be done within reasonable limits. It is assumed that the checks that the ARC performs on the dataset are robust and therefore sufficient to comprise adequate review of the data submitted by universities.

a. What would be the advantages? Please explain your answer.

Publishing submission data improves transparency and allows institutions to benchmark their performance; however, it is questionable whether this level of detail is required to do so.

b. What would be the disadvantages? Please explain your answer.

Releasing this level of detail will open ERA to futile, endless, and no doubt, in some instances, ill-informed argument. It will undermine the process and create opportunity for assessments to be criticised. Certainly, there is merit to maximising transparency, but this must be done within reasonable limits. It is assumed that the checks that the ARC performs on the dataset are robust and therefore sufficient to comprise adequate review of the data submitted by universities.

Q3.39 What other data do you think the ARC should publish following an ERA round? Please describe.

None

Section 4—Engagement and Impact Assessment

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.

A moderate amount. The number of Linkage Projects funded by the ARC increased from 132 to 175 between 2017 and 2019, and the number of Linkage Project partner investigators increased from 149 to 331 during the same period (ARC research funding trend data); this data attests to the increased collaboration between universities and research end-users since the inaugural EI. Also, the importance of university and research end-user partnerships now features prominently in many university strategic plans, and the development of and investment in software packages designed to manage relationships between universities and research end-users seems to have grown.

The Government introduced the Research Support Program (RSP) at around the same time as EI. As the RSP offers incentives for universities to engage with research end-users, it may also be encouraging more collaboration.

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.

A small amount. As a selective rather than comprehensive assessment, EI has not provided clarity to the Government and the Australian public about how their investments in university research translate into tangible benefits beyond academia. The return on investment should become clearer as more research projects are appraised in future EI assessments.

- 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.**

A large amount. The process of preparing submissions for EI illuminates the activities and infrastructure that facilitate research engagement, as do the ratings awarded for 'approach to impact'. The publication of high-rating engagement and impact narratives on the ARC Data Portal enables universities to learn about the systems that other institutions have in place to advance research engagement.

- 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.**

A moderate amount. Many universities have introduced workload and performance management policies, as well as grant and award schemes that encourage researchers to translate their research into impact beyond academia. Investment in staff to develop research partnerships and support knowledge transfer activities also seems to have grown since EI 2018. Research engagement and impact is complicated in the current funding environment in that it requires resources beyond the life of a grant—resources that invariably do not exist.

- 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. The process of collecting data for EI enables universities to identify the ways in which their own institution translates research into impact. The publication of high-rating engagement and impact narratives on the ARC Data Portal is an effective way of sharing good practice; however, a comprehensive understanding of how universities translate research into impact could be achieved if all engagement and impact narratives were made available. This is because there is as much to learn from considering weak examples of engagement and impact as there is from examining strong ones.

- 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. We are not aware of anything that would render the EI objectives inappropriate for the future needs of its stakeholders.

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

- a. **the Australian university sector as a whole? Please describe.**

EI has brought into short focus that research is more than working at the bench and publishing research outcomes. In addition to promoting greater collaboration between universities and research end-users (refer to Q4.1 a.), the assessment is making universities more accountable for translating research into benefits for Australia beyond academia.

b. Individual universities. *Please describe.*

For ACU, EI has highlighted the processes and infrastructure that facilitate research engagement and impact. We have also learned from the engagement and impact narratives published on the ARC Data Portal how other universities support knowledge translation and the benefits that have resulted. Importantly, ACU has a new strategic priority that focuses on both research excellence and research impact—that is, 'world-leading research, with impact'—and has introduced policies and strategies to translate knowledge into impact. We have also increased our publicising of engaged and impactful research.

c. researchers. *Please describe.*

At ACU, the impact EI has had on researchers varies enormously—everything from sparking their interest in engagement and impact to cementing their view that knowledge translation is not their responsibility. Most researchers now understand that to achieve impact beyond academia, they need to think about more than outputs when they plan their research projects.

d. other sectors outside of academia? *Please describe.*

Sectors outside of academia are contributing cash and in-kind support to research projects to a greater extent than they were prior to the inaugural EI assessment (ARC Linkage Projects selection reports; refer to Q4.1 a.), and more research end-users are involved in research training programs (ACU HDR supervision data; Australian Council of Graduate Research commentary). Further impact could be achieved if the ARC were to promote the merits of partnership research to end-users and the Australian public.

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

ACU uses its EI outcomes to educate researchers about the value of translating knowledge into benefits for Australia beyond academia. EI is a key performance indicator in the ACU Strategic Plan 2020–2023, and we also discuss EI outcomes with some of our current and prospective research partners.

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.*

Agree. EI outcomes are valuable to ACU because they raise awareness of how research can contribute to social, economic, cultural and/or environmental benefits beyond academia; provide a new way for ACU to showcase the value of its research to end-users; point to fields of research that could be more engaged and impactful; highlight the ways in which ACU can augment its approach to impact; and inform the review of policies to incentivise knowledge translation.

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

EI outcomes could be factored into the Australian Government's funding model for universities. Funds could be allocated to universities on the basis of their performance in EI to encourage greater collaboration between universities and research end-users, and to support the translation of research into benefits for Australia beyond academia.

EI definitions

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

Agree

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

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

Agree

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

Q4.9 The current end-user definition is appropriate. *Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree.*

Disagree

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

The specific exclusions of end-users should be redefined so that legitimate end-users of research (that is, those who are excluded in the current definition) can be counted as same—see the examples in Q4.9 b..

b. Are there any end-user categories excluded in the current definition of research end-user that you think should be included? *Please explain your answer.*

Yes. End-users who hold casual or honorary appointments at higher education providers should be classified as end-users if their employment contracts at these organisations do not exceed 0.2 FTE. Also, excluding higher education providers in all jurisdictions fails to acknowledge that research which engages with and benefits the development of higher education providers in low-income countries is very different to researching with colleagues in developed countries. Importantly, many organisations that are affiliates, controlled entities or subsidiaries of higher education providers—such as medical research institutes—are legitimate end-users of research and should be counted as same.

Q4.10 Are there other key terms that need to be formally defined? *Yes/No. If you answered 'Yes', please explain your answer.*

No

El methodology

Unit of assessment

- 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 two-digit FoR codes speak to discipline areas that are recognised and understood in the Australian higher education sector. Using the two-digit FoR codes for EI has two benefits: it creates efficiencies when collecting and curating data for the assessments, and it makes it easier to compare results across the assessments (for example, in which FoRs is a university excellent and impactful and in which FoRs is a university excellent but not impactful). Using two-digit FoR codes also helps to address the wider perspective of EI by allowing universities to combine projects and share resources within broad disciplines to translate knowledge into impact in ways that would not be possible, for example, at the four-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.*

No. While Socio-Economic Objective codes and Sustainable Development Goals could be used to classify units of assessment in EI, it is less burdensome and more transparent for universities to collect and report on data using a single classification system when the same data (or some of the same data) is going to be used for both ERA and EI.

Selectiveness of EI

- 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.*

The same number as in EI 2018

- a. *How many and why? Please explain your answer.*

The same number as in EI 2018 but adjusted for any changes to the number of two-digit FoR codes that are adopted by the ARC as a result of the ANZSRC Review (assuming that it does adopt them).

EI low-volume threshold

- 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

- a. *If you disagree, how should eligibility for assessment in EI be determined? Please explain your answer.*

- 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.*

Disagree. The low-volume threshold of 150 weighted outputs at the two-digit FoR level is too low. In some FoRs, the large number of four-digit codes that sit beneath the two-digit codes and the way in which research is dispersed across the codes (that is, very thinly in some instances) can result in there being insufficient or no impact to report. Increasing the threshold to c. 250 weighted outputs at the two-digit FoR level would help to ameliorate this problem, assuming that universities that do not meet the threshold could opt in to EI should they wish to have their engagement and impact assessed. Similarly, universities that do meet the threshold should be able to request that a unit not be assessed if the research area at the university is new (as they were able to in EI 2018).

Engagement indicators

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. Please explain your answer.*

Disagree. The engagement indicator suite used for EI 2018 does not recognise in-kind support from research end-users. While such support can be difficult to quantify, the in-kind support that is referenced in successful grant applications could be included. Co-supervision of HDR students by research end-users should also be included in the engagement indicator suite, and possibly co-authorship with research end-users.

All engagement indicators should continue to be assessed holistically and within the context of the engagement narrative and indicator explanatory statement.

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.*

Neither agree nor disagree. The cash support from the research end-users indicator is appropriate for the assessment of research engagement in some but not all disciplines. This indicator favours those units of assessment that collaborate with commercially-focused industries rather than, for example, community and not-for-profit organisations. Organisations that are not commercially-focused may only be able to offer in-kind support, which can be just as valuable to a research project as cash support.

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. Research commercialisation income is appropriate for the assessment of research engagement in a small number of disciplines. Including it in the engagement indicator suite provides a paradoxical motivator for universities to sell their research outcomes rather than share them freely for the benefit of all Australians.

Q4.19 Are there additional 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. In-kind support from research partners should be included as an engagement metric where that support has been referenced in approved grant applications, and co-authorship with research end-users should be considered for inclusion. We say considered because the revised definition of a research end-user will determine the feasibility of this suggestion.

Q4.20 Are there alternative metrics that would be appropriate across many or all disciplines? Yes/No. Please specify the metrics.

No

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

No

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.

Agree. The ACOLA Review of Australia's Research Training System recommended that a greater proportion of HDR training projects should involve end-user co-supervisors, given the significant amount of new knowledge that HDR students generate and could translate into benefits beyond academia. The percentage of HDR students with an end-user co-supervisor (within a unit of assessment) would be the appropriate statistic for EI assessment panels to consider.

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

a. Patents. Yes/No. Please explain your answer.

No. Universities should be encouraged to refer to patents in their engagement narratives (in those units of assessment where they are relevant), rather than using patents as an indicator of research engagement in EI.

b. Research commercialisation income. Yes/No. Please explain your answer.

No. Including research commercialisation income in the engagement indicator suite provides a paradoxical motivator for universities to sell their research outcomes rather than share them freely for the benefit of all Australians.

c. Registered designs. Yes/No. Please explain your answer.

No. Universities should be encouraged to refer to registered designs in their engagement narratives (in those units of assessment where they are relevant), rather than using registered designs as an indicator of research engagement in EI.

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

No. Universities should be encouraged to refer to plant breeder's rights in their engagement narratives (in those units of assessment where they are relevant), rather than using plant breeder's rights as an indicator of research engagement in EI.

e. NHMRC endorsed guidelines. Yes/No. Please explain your answer.

No. Universities should be encouraged to refer to NHMRC endorsed guidelines in their engagement narratives (in those units of assessment where they are relevant), rather than using NHMRC endorsed guidelines as an indicator of research engagement in EI.

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.

Agree. The engagement narrative complements the story that the engagement indicators tell. The narrative enables universities to talk about the post-award engagement activities that were undertaken to translate knowledge into impact. How project teams engaged with end-users in a mutually beneficial way and what the purpose of the engagement was are important and cannot be assessed using the engagement indicators.

Universities and assessors would probably benefit from a more structured template that provides examples of what should be included and excluded in the narrative.

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.

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.

Agree. One engagement submission per unit of assessment highlights the engagement activity that has taken place across the broad discipline, rather than focusing on a specific case study. If universities were required to submit multiple engagement narratives (for example, at the project or even four-digit FoR level), they would be onerous for universities to prepare and for the ARC to assess.

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

Disagree. To increase the length of the engagement narrative would impose more work on universities and it may not increase the quality of the submission. Universities should be succinct and provide information in their narratives that is relevant to the assessment within the specified character limit (that is, the same character limit that applied to EI 2018 submissions).

Q4.27 Additional evidence is needed within the narrative. Strongly agree; Agree; Neither agree or disagree; Disagree; Strongly disagree. Please explain your answer.

Agree. The requirements for EI 2018 would have made it difficult for assessment panels to gauge whether the evidence referred to in the engagement narratives was accurate and sat within the reference period (especially within the end date of the reference period).

a. If you agreed, what evidence should be provided? *Please describe.*

Universities should be required to attach evidence to their narratives and demonstrate that the activities they refer to occurred within the reference period. Random audits and publishing all engagement narratives on the ARC Data Portal would also be effective ways of encouraging accountability in this regard.

Impact narrative

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.*

Agree. The narrative approach for describing impact enables universities to describe who or what has benefited from the research; how the research made a social, economic, cultural and/or environmental impact beyond academia; and the extent of the impact.

Universities and assessors would probably benefit from a more structured template that provides examples of what should be included and excluded in the narrative.

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.*

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.*

Agree. One impact study per unit of assessment provides an example of the impact that has been achieved in the broad discipline. If universities were required to submit multiple impact studies (for example, a number per FTE academic staff), they would be burdensome for universities to prepare and for the ARC to assess.

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 impact narrative requires universities to be succinct in describing the impact of a research project. The character limit that applied in EI 2018 is sufficient.

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.*

Agree. The requirements for EI 2018 would have made it difficult for assessment panels to gauge whether the evidence referred to in the impact narratives was accurate and sat within the reference period (especially within the end date of the reference period).

a. If yes, what evidence should be provided? *Please explain your answer.*

Universities should be required to attach evidence to their narratives and demonstrate that the activities they refer to occurred within the reference period. Random audits and publishing all engagement narratives on the ARC

Data Portal would also be effective ways of encouraging accountability in this regard.

Q4.32 In your opinion, are there quantitative indicators that could be used to measure the impact of research outside of academia? Yes/No. Please explain your answer.

No. Universities can include quantitative indicators, such as altmetrics (for example, citations in public policy documents, discussions on research blogs, mainstream media coverage and mentions on social networks such as Twitter), in their impact narratives for consideration by assessment panels. Where they are provided, the quantitative indicators can be considered within the context of the narrative.

a. If you answered 'yes' to the previous question, please name and describe the quantitative indicator/s, and the disciplines for which they are relevant. Please list and describe.

Approach to impact Narrative

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.

Agree. The approach to impact narrative enables universities to explain how they facilitated the impact described in the impact study. It highlights, among other things, what sort of support was provided, how the support was implemented by the researchers and how researchers engaged with research end-users. Universities and assessors would probably benefit from a more structured template that provides examples of what should be included and excluded in the narrative.

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.

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.

Agree. One approach to impact narrative per unit of assessment provides an example of the approach to impact in the broad discipline. If universities were required to submit multiple narratives of approach, they would be onerous for universities to prepare and for the ARC to assess.

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. The approach to impact narrative requires universities to be succinct in describing their approach to impact for a research project. The character limit that applied in EI 2018 is sufficient.

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. The requirements for EI 2018 would have made it difficult for assessment panels to gauge whether the evidence referred to in the approach to impact narratives was accurate and sat within the reference period (especially within the end date of the reference period).

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

No. While there is considerable overlap between engagement and approach to impact, engagement relates to the unit of assessment and approach to impact relates to the impact study. The approach to impact narrative also requires universities to provide details that the engagement narrative does not; for example, how the institution supported the impact and how that support was implemented by the researchers, evidence of reviewing impact processes and outcomes, and financial or other resources made available to facilitate the realisation of the impact. To avoid (or reduce) any overlap of activities reported in submissions for engagement and approach to impact, the research project that is the focus of the impact/approach to impact narrative could be excluded from the engagement narrative.

El rating scales

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.

Disagree. The three-point engagement rating scale used in EI 2018 is raw and does not have the gradation to differentiate between units of assessment that have a broad range of outcomes. A five-point rating scale (for example, A, B, C, D and E) would enable more nuanced ratings for universities and would align with the five-point scale used for ERA; also, it would allow for comparisons between the outcomes in the two assessments at the two-digit FoR level. If funding were to be attached to performance in EI (refer Q4.6), missing out on a rating of 'A', for example, and being awarded a 'B' would not have the same consequences as missing out on a 'high' and being awarded a 'medium' under the rating system used in EI 2018. (This assumes that performance funding would be attached to ratings above 'C' or 'medium' in the respective scales.) While a five-point rating scale would likely result in more equivocation by the assessment panels, it would be possible to manage given the experience that panel members should have in using scales that have more than three points.

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.

Disagree. The descriptors for the three-point engagement rating scale are not calibrated equally; that is, there is a small difference in the descriptors for 'high' and 'medium' ("highly effective interactions" versus "effective interactions" and "well integrated" versus "incorporated") but a vast difference between both of these and the descriptor for 'low' as "little or no". Consistent with our response to Q4.38, if a five-point scale were to be adopted, appropriate descriptors would need to be developed. Rating moderation exercises should also take place both within and across assessment panels to ensure the same standards are applied and that outcomes are comparable across units of assessment.

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.*

Disagree. The three-point engagement rating scale used in EI 2018 is raw and does not have the gradation to differentiate between units of assessment that have a broad range of outcomes. A five-point rating scale (for example, A, B, C, D and E) would enable more nuanced ratings for universities and would align with the five-point scale used for ERA; also, it would allow for comparisons between the outcomes in the two assessments at the two-digit FoR level. If funding were to be attached to performance in EI (refer Q4.6), missing out on a rating of 'A', for example, and being awarded a 'B' would not have the same consequences as missing out on a 'high' and being awarded a 'medium' under the rating system used in EI 2018. (This assumes that performance funding would be attached to ratings above 'C' or 'medium' in the respective scales.) While a five-point rating scale would likely result in more equivocation by the assessment panels, it would be possible to manage given the experience that panel members should have in using scales that have more than three points.

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.*

Disagree. The descriptors for the three-point impact rating scale are not calibrated equally; that is, there is a one word difference in the first descriptor for 'high' and 'medium' ("highly significant contribution" versus "significant contribution") and no difference on the second descriptor ("a clear link") but a vast difference between both of these and the descriptor for 'low' as "little or no contribution". Consistent with our response to Q4.40, if a five-point scale were to be adopted, appropriate descriptors would need to be developed. Rating moderation exercises should also take place both within and across assessment panels to ensure the same standards are applied and that outcomes are comparable across units of assessment.

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.*

Disagree. The three-point engagement rating scale used in EI 2018 is raw and does not have the gradation to differentiate between units of assessment that have a broad range of outcomes. A five-point rating scale (for example, A, B, C, D and E) would enable more nuanced ratings for universities and would align with the five-point scale used for ERA; also, it would allow for comparisons between the outcomes in the two assessments at the two-digit FoR level. If funding were to be attached to performance in EI (refer Q4.6), missing out on a rating of 'A', for example, and being awarded a 'B' would not have the same consequences as missing out on a 'high' and being awarded a 'medium' under the rating system used in EI 2018. (This assumes that performance funding would be attached to ratings above 'C' or 'medium' in the respective scales.) While a five-point rating scale would likely result in more equivocation by the assessment panels, it would be possible to manage given the experience that panel members should have in using scales that have more than three points.

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.*

Disagree. The descriptors for the three-point impact rating scale are not calibrated equally; that is, there is a one word difference in the first descriptor for 'high' and 'medium' ("highly effective and well integrated" versus "effective and integrated") and no difference on the second descriptor (both are "facilitated the impact") but a vast difference between both of these and the descriptor for 'low' as "not effective and integrated" and "did not facilitate". Consistent with our response to Q4.42, if a five-point scale were to be adopted, appropriate descriptors would need to be developed. Rating moderation exercises should also take place both within and across assessment panels to ensure the same standards are applied and that outcomes are comparable across units of assessment.

El 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.*

Yes. It is important that those universities that wish to have an interdisciplinary research project assessed for impact be able to do so. The submission of interdisciplinary impact studies should be voluntary (as it was in EI 2018).

El 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.*

No. Given that the two-digit code for Indigenous Studies was created this year, the EI low-volume threshold should not be applied to the unit of assessment for Aboriginal and Torres Strait Islander research in EI 2024. Instead, universities should be able to opt in if they would like to have their research in this area assessed (in the same way they were able to in EI 2018).

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. Given that engagement is the means by which impact is achieved, it should be included for Aboriginal and Torres Strait Islander research in EI 2024 (assuming that data on research engagement in this unit of assessment is available before the submission deadline). Consistent with our response to Q4.45, engagement submissions to EI 2024 for Aboriginal and Torres Strait Islander research should be on an opt-in basis.

Section 5—Overarching Issues Common to both ERA and EI

Frequency of ERA and EI

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

Every three years. Conducting ERA every three years enables performance to be tracked in greater detail (with more data points) whilst still employing a meaningful 6-year reference period. ACU uses the assessment to inform strategic planning, such as developing our research orientation and focus, the creation of new research institutes/centres, and informing our recruitment decisions.

With a five-year reporting period, institutions would not be able to benefit from a more regular assessment of progress and ability to respond to identified areas for development would be delayed unnecessarily. That is, a five-year period between ERA rounds would slow progress and provide fewer opportunities for universities to respond and develop. That being said, in employing a 3-year cycle, the ARC should consider introducing the measures recommended here to reduce the administrative burden of the submission process.

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.

Increasing the assessment will not increase the reliability of results but only reduce the number of opportunities for institutions to benchmark their performance and respond to identified areas for development, slowing progress.

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

Every three years. EI should align with ERA, following it by c. six months so that the EI submission is based on the data that has been submitted for ERA and accepted by the ARC. This is important because some of the data that is used for ERA is also used for EI, and there are efficiencies to be realised in curating it once for the two assessments. If three-year intervals are not feasible for EI, then six-year intervals would be the next most appropriate option. To conduct EI every five years would put it out of sync with ERA (assuming that ERA is conducted at three-year intervals), requiring universities to carry out an additional data collection and curation exercise. Rather than streamlining and simplifying ERA and EI, this would impose more assessment-related work on institutions.

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.

An assessment cycle that exceeds three years would diminish the value of EI results in that there would be fewer opportunities for universities to reflect on areas of strength and weakness, and to develop and implement strategies to improve research engagement and impact outcomes.

Streamlining and simplifying ERA and EI

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.

Strongly disagree. These assessments should be kept separate in order to reduce the burden of reporting, which will be felt particularly by smaller institutions if the assessments are combined. Separating ERA and EI allows for effective distribution of the reporting workload.

a. What would be the advantages and/or disadvantages. *Please explain your answer.*

Disadvantage: Combining the assessments will place an inordinate administrative burden on institutions at the time of reporting. This is a concern for smaller institutions in particular.

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.*

Yes. Introducing an element of citation analysis to codes currently assessed by peer review only will reduce the burden on reviewers and provide a more efficient means of assessment (see response to 3.17).

The collection of metadata not used for the assessment process is discouraged. We suggest that there is limited usage of the applied measures (including patents, research commercialisation income, registered designs, plant breeder's rights and NMHRC endorsed guidelines) to assess research quality and these could be removed from the ERA collection.

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

ERA would benefit from including data available from international ranking systems in its assessment. Both Leiden and Times Higher Education rankings conduct assessments in subject areas congruent with ANZSRC codes and provide data that might be used to guide the assessment process. Data from across the reference period might be provided to guide panels in their deliberations.

ACU has been utilising external data sources to streamline our internal processes to reduce the costs of our participation in ERA and EI. Because of the gaps in coverage of these external data sources, we rely on the manual submission and classification of research outputs from our researchers to ensure a comprehensive submission. The issues of using external data sources, are those of comprehensiveness, lack of curation, and the accurate allocation of fields of research. For these reasons, ACU would not endorse the harvesting of metadata by the ARC to assess the quality of ACU's research outputs in the ERA or EI assessment exercises.

Comprehensiveness

External databases' -such as Scopus or Web of Sciences- coverage of journals is a subset of the ERA journal list, and as such misses out in their coverage of disciplines that have an Australian focus, such as Education, Law and Business. These external databases also do not cover all the eligible research output types, including books, book chapters, conferences and NTRO's, which are critical for the assessment of some HASS disciplines.

This issue for ORCID will be addressed in the response to Q5.10, Q5.11 and 5Q.13 below.

Lack of Curation

The classification of research outputs by external databases, into articles, reviews, short surveys etc is not accurate. Often scholarly works will be classed as an article when in fact it is a commentary or an opinion piece. The issue is

particularly acute for reviews, which have to be individually read to assess if it's research eligible or a scholarly work.

For non-journal research outputs, it should be noted the business model for commercial publishers requires them to publish both scholarly works such as textbooks as well as eligible research outputs. Where these works appear in external databases, they need to be read to assess their research eligibility.

Accurate Fields of Research

Some external data providers classify their journals to the ERA four-digit coding scheme. This process is not accurate enough to represent the research content of the research output. This is even more problematic for the multidisciplinary journals and for the journals covering a broad disciplinary area. Unfortunately, the proposed use of machine learning (ML), is not seen to be feasible or when done, not seen to be accurate enough to not misclassify research outputs. It is assumed that to train the ML algorithm, that data will be sourced from Universities or from the submitted ERA data. The reason for this is the inherent bias contained in the data from different authors. On the same research output, an author who is a public health specialist, would judge the same research output to have a different Field of Research, than say another author who specialises in clinical sciences. This bias will mean that the ML algorithm will replicate the bias as it has been trained in that biased data. The recent change in the classification scheme for fields of research would mean that the quantity of data required to train an ML algorithm will not be available after the 2023 ERA submission has closed.

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

Ensuring accurate apportionment is an extremely time-consuming exercise. Research staff frequently allocate incorrect FoR codes to publications and classify non-research material as A-, B-, C-1. There is a significant amount of work involved in reviewing the submission and ensuring its accuracy. Even with a systematic process of review in place, given the scale of work involved in reviewing the submission, errors in the dataset can be missed.

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

Similarly time consuming is the process of balancing the very strict technical requirements to nominate exactly 30% of research outputs per publication type for assessment. This is particularly felt by smaller institutions for whom a minor change in the submission data can result in the 30% needing to be significantly reworked.

Q5.9 In your view what are the most time-consuming elements of an EI submission? Please describe.

For EI 2018, the most time-consuming elements of the submission were collecting the engagement, impact and approach to impact data for the narratives. Given that ACU now has systems in place to collect these data on an ongoing basis, the preparation of submissions for future EI assessments should be less onerous.

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

No. Assuming that EI continues to be a selective rather than comprehensive assessment, there are no efficiencies that could be introduced for EI at this time.

Utilising technological advances and pre-existing data sources

Q5.10 ORCID iDs should be mandatory for ERA. *Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.*

Strongly disagree. The lack of transparency in the ORCID system means that we would not be able to judge if the correct identifier has been obtained for the correct author. There are too many ORCID profiles which are not publicly viewable to have confidence in the match of author to ORCID.

a. What are the advantages and/or disadvantages? Please explain your answer.

Disadvantage: The lack of transparency in the ORCID system means that we would not be able to judge if the correct identifier has been obtained for the correct author. There are too many ORCID profiles which are not publicly viewable to have confidence in the match of author to ORCID.

Q5.11 The automatic harvesting of output data using ORCID iDs would streamline a university's submission process. *Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.*

Neither agree nor disagree. If the external data providers were able to provide this consistently, for all authors, then this would assist with the deduplication of our authors, especially for external authors. Unfortunately, this data is not consistently available from the external data providers. This approach would introduce errors into the submission process and create an incomplete dataset that, rather introducing efficiencies, would be more time consuming for universities to fix.

a. What are the advantages and/or disadvantages? Please explain your answer

This approach would introduce errors into the submission process and create an incomplete dataset that, rather introducing efficiencies, would be more time consuming for universities to fix.

Q5.12 DOIs should be mandatory for ERA. *Strongly agree; Agree; Neither agree nor disagree; Disagree; Strongly disagree. Please explain your answer.*

Neither agree or disagree. ACU takes an active interest in collecting DOIs as these help to improve the accuracy of our metadata for reporting purposes. These also help streamline the collection of data via the use of Application Programming Interfaces. As many research outputs do not have a DOI, we could not see how DOI's could be made mandatory. Rather we believe that the reporting of DOI's should be strongly encouraged.

a. What are the advantages or disadvantages? Please explain your answer.

Advantage: DOIs help to improve the accuracy of metadata for reporting purposes. These also help streamline the collection of data via the use of Application Programming Interfaces.

Disadvantage: As many research outputs do not have a DOI, these could not be made mandatory. Rather we believe that the reporting of DOIs where available should be strongly encouraged.

Q5.13 Are there new ways to collect data to reduce the cost and burden to universities of participating in ERA and EI whilst maintaining the robustness of the ERA and EI process? Yes/No. Please explain your answer.

Yes. ACU has been investing in new ways to collect data to reduce the cost and burden of participating in ERA and EI. We have successfully transitioned from a manual paper-based data collection process, to an e-forms based workflow system integrated with the automatic harvesting of metadata and soft copies of journal articles using API's.

a. What are the advantages and/or disadvantages? Please explain your answer.

The usage of the ORCID system as a replacement for the external data provides such as Scopus or Web of Science is not supported. The lack of curation of the data is a concern, as individual authors are responsible for the submission of outputs both scholarly and research. Institutions would need to unpick this data to align it with the ERA reporting requirements, adding to the reporting burden. Also, too many of the ORCID profiles are private and, as a result, institutions would not be able to ensure the accuracy of their ERA and EI submission.