Combining Excellence in Research for Australia (ERA) and the Higher Education Research Data Collection (HERDC).

December 2014
1. Introduction

1.1 Purpose and Scope of this paper

The Australian Government has accepted all the recommendations of the PhillipsKPA Review of the Reporting Requirements for Universities. The development of and implementation of a single higher education research data collection is a key element of those recommendations. A single data collection will reduce the reporting burden on universities by ensuring that research data collected by the Government is collected once and can be used for a range of purposes. The Department of Education (the Department) and the Australian Research Council (ARC) have been tasked with implementing the recommendation.

This paper forms a preliminary part of that task. The Department, together with the ARC, is seeking the views of universities regarding a preferred model for a single collection of data for the Higher Education Research Data Collection (HERDC) and Excellence in Research for Australia (ERA).

The paper outlines 3 models for a future single collection of research data at Australian universities. Comment is sought on which model is preferred, taking into account the advantages and disadvantages of each model, the potential work effort involved, and the value of incorporating other higher education research data collections.

To allow for institutions to prepare for the changes, the implementation of a single higher education research data collection is planned from 2016–17.

1.2 Background

The Australian Government is committed to fostering and rewarding Australia’s university research excellence, effort, and capacity. Australia has a dual funding system for university research and research training. The system comprises competitive research grants, which are distributed through peer-reviewed processes (for example, ARC and National Health and Medical Research Council (NHMRC) grants) and Research Block Grants (RBG), which are allocated according to performance-based formulae and support the indirect cost of conducting competitive grant research, institutional research priorities, and research training. This integrated system rewards and encourages high quality research, engagement with end users and research training.

Competitive research grants reward university performance directly through successful applications assessed through merit based processes. The majority of competitive research grant funding is provided by the Australian Government through the ARC ($787 million in 2013), NHMRC ($610 million in 2013) and Rural R&D Corporations ($82 million in 2013)\(^1\).

In 2014, the Australian Government will provide $1.72 billion to higher education providers as block grants to support research and research training, through a number of performance-based schemes

\(^1\) 2014 Higher Education Research Data Collection (HERDC).
under the Higher Education Support Act (HESA) 2003. The block grants under HESA are ongoing programs.

Universities also source research funds from other government, private and overseas sources.

1.2.1 Review Recommendations

Reducing the reporting burden placed on universities is an important issue for the Australian Government. In December 2013, the Australian Government responded to the PhillipsKPA Review of Reporting Requirements for Universities (the Review). Recommendation 19 of the Review was:

That [the department] consider developing a single higher education research data collection that could serve more than one purpose, noting that this would inevitably require some compromise to the specific objectives of the existing data collections and related funding programs, and hence would require policy change.

The Government agreed with this recommendation, and the Department and the ARC were tasked with:

...the development of a single higher education research data collection. There will be further consultation on its design, development and implementation. The aim will be to ensure research data is only collected once.

Currently two research data sets are collected by the Australian Government for separate purposes. The HERDC is collected to inform the allocation of the RBG by the Department. The ERA data set is collected to evaluate the quality of university research by the ARC.

1.2.2 HERDC

HERDC is a dataset that records research income earned and research publications coming from eligible higher education providers. These data are key drivers of the allocation of the RBG along with higher degree by research student data. Data is collected annually by the Department directly from universities. The HERDC specification document provides guidance for universities on the specific content and data requirements in providing research income and research publications data. The HERDC specifications also provide a framework for the collection of ERA research income.

For further information about HERDC, please visit the Department of Education webpage: https://education.gov.au/higher-education-research-data-collection.

---


3 Eligible higher education providers are defined in HERDC as those institutions identified as Table A and Table B providers in sections 16-15 and 16-20 of the Higher Education Support Act 2003.
1.2.3  ERA

ERA is an assessment of the research quality within eligible Australian higher education providers\(^4\) and gives Government, industry, business and the wider community assurance of the excellence of research being conducted with public funds. It provides a national stocktake, by research discipline areas, of research strength against international benchmarks. ERA assessments were conducted in 2010 and 2012.

ERA outcomes do not feed into the competitive grants processes. They do, however, moderate the allocation of the funding provided to universities through the Sustainable Research Excellence (SRE) research block grant scheme. The next round of ERA will be conducted in 2015. The timing of future rounds of ERA has not yet been announced. ERA is an ongoing program.


1.2.4  Key differences between the HERDC and ERA data

ERA is generally a more comprehensive research data collection than HERDC, encompassing research income, publications, researcher, applied and esteem measure data. However, in terms of research publication and income data, the information collected through ERA and HERDC is broadly the same. The key differences between the HERDC collection and the ERA collection are:

- A co-authored publication is attributed proportionally to the institutions involved in the work in HERDC whereas each institution is able to claim a whole co-authored publication in ERA;
- More detailed Category 2, 3 and 4 income information is collected in HERDC compared to ERA, and more detailed publication information is collected in ERA compared to HERDC;
- Fees for non-RTS domestic students and international HDR students are included as Category 3 income in HERDC but not counted in ERA;
- ERA data is submitted by Field of Research (FoR) whereas HERDC data is not; and
- Non-traditional research outputs (NTROs) are collected in ERA but not HERDC.

1.3  Aims of a single data collection for ERA and HERDC

The primary purpose of a single higher education research data collection is to eliminate the duplication imposed on universities by having to report research publications and income data to both the Department, through the HERDC, and the ARC, through ERA. A single data collection will provide a more efficient and streamlined process for the collection of data to inform RBG calculations and evaluate excellence in research in Australian universities. Potential efficiencies include:

\(^4\) Eligible higher education providers are listed in the [ERA 2015 Submission Guidelines](http://www.arc.gov.au/era/era_2015/era_2015.htm). This list is based on the Table A and Table B providers identified in the Higher Education Support Act 2003.
• IT systems at universities only need to cater for one data set;
• workflow arrangements for the collection and verification of data can be rationalised;
• streamlining of data collection and verification processes will reduce opportunities for error and misunderstanding, thereby enhancing the integrity, transparency and utility of the data collected;
• administrative and academic staff involved in the collection and verification of data at the universities only need to be trained to work with one set of data specifications; and
• auditing requirements in relation to the data will be streamlined.

1.4 Proposed IT system for data submission

The ARC uses a platform-neutral format for the submission of ERA data. The Department uses downloadable smart forms for the submission of HERDC data. The ERA dataset is generally more comprehensive and provides a greater level of disaggregation than the HERDC, for example, through the assignment of FoR to data elements. Therefore, the ERA IT system forms the logical starting point for a single higher education research data collection.

Throughout this paper it is assumed that the ERA system will be used to support all of the models under discussion. This builds on the existing capabilities of universities to provide disaggregated data, and aims to limit potential changes to university systems in order to implement a single higher education research data collection.

2. Options for a single higher education research data collection

This consultation paper outlines 3 options for reporting publications and income data to the ARC. Universities will need to continue to collect research publications and income data. However, under all options proposed in this paper, this data will only need to be provided once to the ARC. No data will need to be submitted by universities directly to the Department of Education, resulting in savings for universities in HERDC submission costs.\(^5\)

In providing feedback universities should consider the potential benefits (or otherwise) of the proposed options and estimate the workload involved in implementing the options.

The 3 models proposed are:

1. collect publications and income data every 3 years via ERA rounds;
2. collect ERA data annually; or
3. a hybrid model, with publications and income data collected annually, and a full set of ERA data collected every 3 years.

\(^5\) HERDC submission costs were not separately identified in the Review of Reporting Requirements for Universities final report.
2.1 Option 1—Collect publications and income data every 3 years via ERA rounds

Option 1 involves a collection of publications and income data once every 3 years. The data collection will be used to inform the RBG calculations and the evaluation of research excellence through ERA.

Both ERA and HERDC record past performance. The RBG calculations use the two most recent HERDC years available. This means the RBG are allocated using data that are 2 and 3 years old. For example, 2015 funding is allocated on the basis of performance in 2012 and 2013.

ERA, however, evaluates research quality over a longer timeframe of six years. For example, ERA 2015 will review outputs from 2008 to 2013 and income from 2011 to 2013. The next round of ERA will be conducted in 2015. The timing of future rounds of ERA has not yet been announced.

One impact of collecting publications and income data every 3 years would be a reduction in the responsiveness of RBG allocations to changes in performance within the sector. Currently, the annual HERDC data collection allows RBG allocations to use data collected within 3 years of the allocation year. If publications and income data were collected once every 3 years, annual RBG allocations would be calculated using research publications and income data ‘fixed’ at a certain point in time.

The table below sets out a hypothetical scenario based on RBG allocations being calculated using publications and income data collected annually through HERDC until 2015 (the final 2015 HERDC collection capturing 2014 data) and then subsequently through ERA rounds in 2018 and 2021. Under this scenario, RBG allocations would use data collected within 6 years of the allocation year.

<table>
<thead>
<tr>
<th>RBG Allocation Year</th>
<th>Income and Publications Data Input Years</th>
<th>Income and Publications Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>2013 &amp; 2012</td>
<td>2014 HERDC</td>
</tr>
<tr>
<td>2016</td>
<td>2014 &amp; 2013</td>
<td>2015 HERDC</td>
</tr>
<tr>
<td>2017</td>
<td>2014 &amp; 2013</td>
<td>2015 HERDC</td>
</tr>
<tr>
<td>2018</td>
<td>2014 &amp; 2013</td>
<td>2015 HERDC</td>
</tr>
<tr>
<td>2019</td>
<td>2016 &amp; 2015</td>
<td>2018 ERA</td>
</tr>
<tr>
<td>2020</td>
<td>2016 &amp; 2015</td>
<td>2018 ERA</td>
</tr>
<tr>
<td>2021</td>
<td>2016 &amp; 2015</td>
<td>2018 ERA</td>
</tr>
<tr>
<td>2022</td>
<td>2019 &amp; 2018</td>
<td>2021 ERA</td>
</tr>
<tr>
<td>2023</td>
<td>2019 &amp; 2018</td>
<td>2021 ERA</td>
</tr>
</tbody>
</table>

While this may act to limit the responsiveness of RBG allocations, the RBG formulae already have inbuilt mechanisms to limit variation in funding between institutions and years. For example, the Research Training Scheme and Joint Research Engagement (base) allocation have safety nets so that no less than 95% of an institution’s previous year’s funding amount is provided in the next year. Therefore, while reported publications and income data can fluctuate widely from year to year, the operation of the RBG formulae moderates associated funding variations, leading to adjustments in the distribution of RBG allocations between institutions of, on average, around $20 million per annum (prior to the application of annual indexation measures). In simple terms, a group of institutions will ‘lose’ a share of around $10 million in RBG funding each year to a group of
institutions that performed relatively better (as measured by annual performance data - publications, income and HDR student data). Of this variation, around 58 per cent is attributable to changes in reported research income and around 8 per cent is attributable to changes in research publications.

There is a precedent for RBG data inputs being fixed for a period. The SRE block grant allocations are currently moderated to a small extent by ERA outcomes, which are fixed for the period between ERA rounds. SRE allocations also continue to be marginally influenced by SRE Staff Hours Survey data collected in 2011.

2.1.1 Advantages

Option 1 has the following advantages:

• universities are only required to submit data once every 3 years;
• ERA collection and submission processes do not change;
• the administrative burden to universities is reduced through the elimination of the HERDC submission processes; and
• universities may be able to better plan due to the RBG allocations more being predictable over a 3 year cycle.

Disadvantages

Option 1 has the following disadvantages:

• RBG would use older data in some years to inform RBG allocations;
• changes in research performance would be less frequently reflected in RBG allocations; and
• RBG allocations may be less responsive to changes in university performance over the longer term.

2.2 Option 2—Collect ERA data annually

Option 2 involves the annual collection of ERA data rather than the current 3 year data collection cycle. These data would be used for the calculation of RBG allocations by the Department each year, and reused by the ARC for ERA evaluation rounds every 3 years. This option would require universities to submit an annual collection of all ERA data including, for example, the ERA Non-Traditional Research Outputs (NTROs) and data relating to the ERA applied measures and esteem measures.

This option involves universities identifying FoR codes annually for all data submitted, allowing for changes in the ERA evaluation year only for staff movements and reassignment of outputs.

The table below sets out a scenario based on option 2 whereby HERDC is phased out in 2015 or 2016, and the collection of publications and income data transitions to full annual ERA data collections thereafter. Depending on the logistics involved, the data collection for the 2017 RBG
allocation year will either be through HERDC or ERA. Transitional arrangements may also require the back-capture of additional data for an ERA 2018 evaluation.

<table>
<thead>
<tr>
<th>RGB Allocation Year</th>
<th>Income and Publications Data Input Years</th>
<th>Income and Publications Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>2013 &amp; 2012</td>
<td>2014 HERDC</td>
</tr>
<tr>
<td>2016</td>
<td>2014 &amp; 2013</td>
<td>2015 HERDC</td>
</tr>
<tr>
<td>2017</td>
<td>2015 &amp; 2014</td>
<td>2016 ERA (full) or 2016 HERDC</td>
</tr>
<tr>
<td>2018</td>
<td>2016 &amp; 2015</td>
<td>2017 ERA (full)</td>
</tr>
<tr>
<td>2019</td>
<td>2017 &amp; 2016</td>
<td>2018 ERA (full)</td>
</tr>
<tr>
<td>2020</td>
<td>2018 &amp; 2017</td>
<td>2019 ERA (full)</td>
</tr>
<tr>
<td>2021</td>
<td>2019 &amp; 2018</td>
<td>2020 ERA (full)</td>
</tr>
<tr>
<td>2022</td>
<td>2020 &amp; 2019</td>
<td>2021 ERA (full)</td>
</tr>
<tr>
<td>2023</td>
<td>2021 &amp; 2020</td>
<td>2022 ERA (full)</td>
</tr>
</tbody>
</table>

An expected increase in administration costs associated with providing ERA data each year is expected to be offset by no longer having to provide multiple years of data in a single ERA assessment year.

### 2.2.1 Advantages

Option 2 has the following advantages:

- the collection of data remains on an annual basis for the purpose of RGB allocations;
- changes in research performance are reflected in RBG allocations annually; and
- the administrative burden on universities is reduced through the elimination of the HERDC submission processes.

### 2.2.2 Disadvantages

Option 2 has the following disadvantages:

- universities would be required to submit the full range of ERA data annually; and
- universities would be required to assign FoR codes to research data annually, well in advance of an ERA submission.

### 2.3 Option 3—Hybrid collection model

Option 3 involves the annual collection of only income and publications data rather than the full suite of ERA data. These data would be used for the RBG allocations each year, and reused by the ARC for ERA evaluation rounds every 3 years. The additional ERA data including, researcher data, NTROs, and applied and esteem measures would be collected every 3 years.

This option involves universities identifying FoR codes for all data submitted annually, allowing for changes in the ERA evaluation year only for staff movements and reassignment of outputs.
The table below sets out a scenario based on option 3, whereby HERDC is phased out in 2015 or 2016, an abridged ERA data collection is conducted thereafter (publications and income data only) and the additional ERA data are collected in ERA assessment years. Depending on the logistics involved, the data collection for the 2017 RBG allocation year will either be through HERDC or ERA. Transitional arrangements may also require the back-capture of additional data for an ERA 2018 round and modifications to the ERA reference periods and evaluation timetable.

<table>
<thead>
<tr>
<th>RBG Allocation Year</th>
<th>Income and Publications Data Input Years</th>
<th>Income and Publications Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>2013 &amp; 2012</td>
<td>2014 HERDC</td>
</tr>
<tr>
<td>2016</td>
<td>2014 &amp; 2013</td>
<td>2015 HERDC</td>
</tr>
<tr>
<td>2017</td>
<td>2015 &amp; 2014</td>
<td>2016 ERA (abridged) or 2016 HERDC</td>
</tr>
<tr>
<td>2018</td>
<td>2016 &amp; 2015</td>
<td>2017 ERA (abridged)</td>
</tr>
<tr>
<td>2019</td>
<td>2017 &amp; 2016</td>
<td>2018 ERA (full)</td>
</tr>
<tr>
<td>2020</td>
<td>2018 &amp; 2017</td>
<td>2019 ERA (abridged)</td>
</tr>
<tr>
<td>2021</td>
<td>2019 &amp; 2018</td>
<td>2020 ERA (abridged)</td>
</tr>
<tr>
<td>2022</td>
<td>2020 &amp; 2019</td>
<td>2021 ERA (full)</td>
</tr>
<tr>
<td>2023</td>
<td>2021 &amp; 2020</td>
<td>2022 ERA (abridged)</td>
</tr>
</tbody>
</table>

No additional ERA costs are expected over a 3 year cycle as the publications and income data collected annually would not be required to be collected in an ERA assessment year.

### 2.3.1 Advantages

Option 3 has the following advantages:

- the collection of data remains on an annual basis for the purpose of RGB allocations;
- only ERA publications and income data collected in non-ERA years;
- changes in research performance are reflected in RBG allocations annually; and
- the administrative burden on universities is reduced through the elimination of the HERDC submission processes.

### 2.3.2 Disadvantages

Option 3 has the following disadvantages:

- universities would be required to assign FoR codes to publications and income data annually, well in advance of an ERA submission.

---

Q1. Please indicate which option your institution would prefer for a single research data collection and why?
Other costs and savings (labour, IT etc.) involved in transitioning to a single higher education research data collection may be experienced by the sector. For example, an IT system dedicated solely to the submission of a single return may involve some establishment costs, but could yield potential long-term savings for universities.

**Q2. What are the anticipated transitional activities, costs and savings for your institution through the creation of a single higher education research data collection?**

### 3. Options for a single higher education research data collection

While removing duplication of reporting publications and income data is the primary focus of this consultation, universities are also encouraged to identify any other higher education research data, such as the National Survey of Research Commercialisation and/or NCRIS project survey data, which might be incorporated into a single higher education research data collection.

**Q3. Are there further benefits to be gained by incorporating other higher education research data collections?**

### Submission of feedback

Please provide feedback to the Australian Research Council via email at era@arc.gov.au

Responses are due Wednesday, 4 February 2015.

Should you have any queries regarding the consultation paper, please contact Ms Sarah Howard, Director, Research Evaluation at the Australian Research Council on (02) 6287 6755, or at era@arc.gov.au.