LADRN PROJECT
The Assessment of Australian Legal Research

Authors: Lyria Bennett Moses and Catherine Renshaw
(LADRN co-chairs)
Research assistant: Nazim Rahman

Date: 14 February 2024

Table of Contents

PART (I) - INTRODUCTION 3

PART (II) – METHODOLOGY 4

PART (III) - BACKGROUND 8

(A) - CALD STATEMENT 8
(B) - ERA AND EI EXERCISES 10
(C) - INTERNATIONAL APPROACHES TO RESEARCH ASSESSMENT 13
(D) - DORA 14
(E) - LEIDEN MANIFESTO 15
(F) - HONG KONG PRINCIPLES 15
(G) - EUROPEAN COMMISSION 16
(H) - COARA 16
(I) - ACOLA 17

PART (IV) - LITERATURE REVIEW 18

(A) - MOTIVATIONS FOR QUALITY ASSESSMENT 18
(B) - OPPOSITION TO QUALITY ASSESSMENT 20
(C) - APPROACHES TO QUALITY ASSESSMENT 21
Approach 1: Criteria 21
Approach 2: Journal lists 23
Approach 3: Citation metrics 24
(D) - SUMMARY OF LITERATURE ON ASSESSMENT APPROACHES 26
(E) - IMPACT 26

PART (V) - SURVEY RESULTS 28

(A) - LEGAL RESEARCHERS 28
(1) - Audience for academic legal research 29
(2) - Perceptions about the value of different outputs 30
(3) - Perceptions of independence in setting research agendas 32
(4) - How researchers choose where to publish 32
(5) – Metrics 35
(6) - Mechanisms for evaluating quality 37
(7) - Impact 41
(9) - Peer Review 43
(B) - INSTITUTIONAL QUALITY AND IMPACT ASSESSMENT 44
(1) - Nationwide Assessment of Research 45
(2) – Value of Different Research Outputs 45
(3) – Attitudes Towards ARC Assessment 46

PART (VI) - DISCUSSION 47

(A) - QUALITY ASSESSMENT 48
(1) - Input-based controls 48
(2) - Criteria-based assessment 50
(3) - Aggregated quality scores 50
(4) - Journal lists 51
(5) - Citation metrics 52
(B) - IMPACT ASSESSMENT 52
(1) – Recognising breadth of impact types 52
(2) – Qualitative nature of impact assessment 53

PART (VII) – PRELIMINARY RECOMMENDATIONS 53

Part (I)- Introduction

Since the 1980's, academic legal researchers, their institutions, and representative bodies such as the Council of Australian Law Deans (**CALD**), have sought to produce a concrete and functional description of legal research and a consensus position on how quality of legal research should be assessed. The consistent view has been that peer review, by panels familiar with the discipline and its sub-specialities, is essential. The predominantly national character of much legal research has traditionally been regarded as ill-suited to the application of 'metrics' which draw on citation statistics in international journals. But a convincing alternative to metrics has yet to emerge. In 2010, the Australian Research Council produced a Quality Law Journal ranking list, which was notoriously distorted, illogical and unfair. Attempts to produce something better have floundered, in Australia and internationally.[1] In the last decade, 'impact' has emerged as a new frontier for measuring productivity. There is no common understanding across Australian law schools of what 'impact' in legal research is and how it can best be demonstrated.

The need to address the issue is now pressing, because of several factors, namely:

- Legal researchers are facing new pressure from within their own universities to subject their work to metrics-based measurements, particularly citation-based measurements including requirements to publish in journals with top quartile citation metrics. It is not clear how the use of metrics encourages the production of research that is high quality and/or of national importance.
- Despite official, national law journal ranking lists falling into desuetude, individual law schools continue to list and rank journals and to incentivise staff to publish in certain journals.[2] The methods used to produce journal ranking lists are not transparent and the effect on research decisions made by individual researchers is unclear. Anecdotally, the effect is pernicious, distorting decision-making about what questions are worth asking and the best place to public research. Some institutions, which do not have their own law journal ranking list, subject their academics to lists produced by other disciplines, such as the Australian Business Deans Council journal ranking list.
- There is still little understanding about how 'impact' in legal research should be assessed, other than agreement that impact is important and that it would be beneficial to have a common approach to measuring and evaluating impact.

In mid-2021, CALD resolved to advance these issues by developing and publishing a (or possibly a number of) statement(s) on Legal Research. In 2022, following the announcement that the 2023 Excellence in Research for Australia (**ERA**) process would be suspended, CALD took the opportunity to gather information both about how Australian legal researchers perceived the assessment of legal research and data about how research is assessed in institutions and nationally. CALD would consider using this as the foundation for a potential re-statement about best practice in the assessment of excellence and impact in legal research. In mid-2022, the authors, who are Co-Chairs of the national network of Associate Deans of Research (the Law Associate Deans of Research Network, '**LADRN')** and Associate Deans of Research (**ADR)** at UNSW Sydney and Western Sydney University, undertook to set

[1] Rob Van Gestel, 'Ranking, Peer Review, Bibliometrics and Alternative Ways to Improve the Quality of Doctrinal Legal Scholarship' in Hans-W Micklitz and Edward L Rubin Rob van Gestel (ed), *Rethinking Legal Scholarship* (Cambridge University PRess, 2017); see 'Metrics and Q1' below for discussion on Council of Australian Law Deans Survey.

[2] Ian Murray and Natalie Skead, 'Who Publishes Where?: Who Publishes in Australia's Top Law Journals and Which Australians Publish in Top Global Journals' (2020) 47(2) *University of Western Australia* 220; see 'Journal Lists' below for discussion on CALD Survey.

up a project to investigate two key questions. The first is whether the ways in which legal research is currently evaluated by individual institutions, and by bodies such as the Australian Research Council (**ARC**), are perceived as being sufficiently transparent, fair, and appropriate for the purpose of identifying the quality of legal research and its impact. The second is whether there are other frameworks legal researchers think might be more appropriate for evaluating the quality and impact of legal research and, if so, what these might be.

We anticipated that there might be (i) a range of measures applied by institutions and by individual researchers to assess legal research quality and impact and (ii) a correlation between the kinds of measures used to assess legal research and certain factors. For example, we anticipated that various measures to assess research might be perceived differently by researchers, depending on factors such as the researchers' stage of career and the kind of institution they worked at (for example a GO8[3] or an Innovative Research University[4]). One potential benefit of the project is that it establishes empirically what research leaders in law units[5] and researchers themselves consider to be important in assessing research quality and impact. This is critical in terms of building stakeholder buy-in for any new regime for assessment.

This paper provides a preliminary overview of some of the project's early findings. Part II describes the methodology adopted to investigate our research questions. Part III details the background to research assessment exercises in Australia and relevant international jurisdictions. Part IV contains a review of the literature associated with the assessment of legal research quality and impact. Part V contains a report on the results of our investigation. Part VI goes into analysis and discussion of the data. Part VII notes recommendations for CALD.

Part (II) – Methodology

We wanted to understand how Australian tertiary institutions firstly assessed the quality and impact of legal research; secondly, how researchers themselves experienced and perceived these assessment processes; and thirdly, how researchers viewed alternative or other means of assessing the quality and impact of legal research. We decided to investigate the research questions by:

- Carrying out a desk-based literature review of scholarship relating to assessment of legal research in Australia and in key overseas jurisdictions such as the United Kingdom, Netherlands and New Zealand. These countries, like Australia, have adopted modified versions of the UK Research Assessment Exercises (RAEs, now known as the Research Excellence Framework (REF)). Literature on legal research assessment from the US was also reviewed, given the centrality of US-based citation metrics.
- Conducting a survey of legal research academics and legal research leaders across Australian universities, to obtain data about: the types of research being pursued (by subject area and methodology); the kinds of work being produced and how these were valued within institutions (e.g. books, chapters, textbooks); the extent to which researchers consider

[3] The Group of Eight (Go8) comprises Australia's leading research-intensive universities – the University of Melbourne, the Australian National University, the University of Sydney, the University of Queensland, the University of Western Australia, the University of Adelaide, Monash University and UNSW Sydney: Group of Eight Australia, 'About the Go8' <https://go8.edu.au/about/the-go8#:~:text=The%20Group%20of%20Eight%20(Go8,Monash%20University%20and%20UNSW%20Sydney.>.

[4] Innovative Research Universities comprises of Flinders University, Griffith University, James Cook University, La Trobe University, Murdoch University, University of Canberra and Western Sydney University: Innovative Research Universities, 'Our Universities' (Web Page) <https://iru.edu.au/our-universities/>.

[5] 'Law units' is a term we use to cover different institutional structures – so a law unit may be a school, faculty, part of one of them or cutting across more than one of them.

themselves free to work on certain subjects or publish in particular venues; how researchers (and their institutions) measure the quality of legal research and/or its impact; whether institutions use journal lists as a measure of quality and if so how this affects researchers' decision-making about what to research and where to publish; the factors that influence journal and publisher ranking lists (where they are used); how book chapters and monographs are assessed; perceptions about the effectiveness and fairness of peer review, both in publishing and grant assessment; and how research impact is and should be measured.

We applied to the Ethics Committee of Western Sydney University for approval to carry out the two surveys, using the Qualtrics program for conducting surveys.[6] Surveys were distributed to (1) all legal researchers (**Researchers Survey)** and (2) Associate Deans of Legal Research (**Research Leaders**) (or their approximate equivalent in different institutions) in all 41 Australian universities. We used the LADRN network to distribute surveys to ADRs and requested that ADRs distribute surveys to individual researchers in their school or faculty.

In order to distribute the survey, we first consulted the ADRs of each institution on 5 July 2023 in a meeting of LADRN. On 20 July 2023, an email was sent to the ADRs with a generic link to the Researchers Survey. That email requested the ADRs forward the link to the legal academics in their institution. For the ADR Survey, a separate generalised link was sent to the ADRs for them only to complete. On 21 August 2023, a reminder email was sent to the ADRs requesting on distribution. Some reminders were sent individually to various people, for example, from universities with no responses.

In relation to the Researchers Survey, we aimed for sample representativeness based on a cross section of responses from different (a) genders; (b) age ranges; (c) participants from GO8 research intensive faculties, smaller schools and regional universities; and (d) academic levels, from associate lecturer to Emeritus Professor. Ultimately, from an estimated 1100 legal researchers,[7] we received 257 responses. Participants came from 35 Australian universities, although the number of participants from each university varied significantly from 26 (UNSW Sydney) to 1 (from six universities). Just over half of the participants were in continuing roles (123/52%). We had more responses from senior than junior academics: 58% were in the professoriate, although we had responses at all levels. 41% identified

[6] The study was approved by the Western Sydney University Human Research Ethics Committee on 4 July 2023. The Approval Number is H15549. This approval was also registered at UNSW Sydney.

[7] To obtain an accurate estimate of the total number of legal researchers in Australian Law Schools, only law school staff members from institutions that produce legal research were counted. It was assumed all law schools and faculties ranked in the 2018 ERA exercise produce legal research: 'Law and Legal Studies ERA Research Rankings' *University Rankings* (2018, Web page) <https://www.universityrankings.com.au/law-and-legal-studies-rankings-2/>. Thirty-three schools or faculties received rankings. In addition, 3 participants from institutions outside this ranking participated in the survey. Accordingly all legal researchers from these 3 institutions were included in the count. The number of academics in research roles as listed on the websites of each of the 36 institutions was then counted. Where it was clear a staff member was purely in a teaching or non-research role, they were not counted. For example, the following types of researchers were counted: Assistant Professors, Associate Deans, Associate Professors, Deans, Deputy Heads, Directors of Research, Emeritus/ Emerita Professors, Executive Deans, Heads of Law Schools, Lecturers who had research outputs, Professors, Research Associates, Research Fellows, Professors of Law, Postdocs, Senior Research Associates and Senior Research Fellows. However, the following types of academics were not counted: Adjuncts, Academic Coordinators, Associate Clinical Educators, Clinical Solicitors, Course leaders, Clinical Legal, Educators, Directors of Legal Centres, Directors of PLT programs, Discipline Leads, Education Focused Academics, Faculty Executives holding a purely executive role, Honoraries, Legal Profession Mentors, PhD candidates not holding another role in their faculty, Practice Professors, Research Assistants, Assistant Teaching Fellows, Senior Teaching Fellows, School Managers, Staff listed on the website but not given a title, Staff listed in 'executive' roles, Teaching Associates, Vice Chancellors and Presidents of multiple faculties (eg Vice Chancellor and Presidents of Faculty of Arts & Society, Law), Visiting Scholars/Fellows.

themselves as a senior researcher compared to 34% mid-career researchers, 23% early career researchers and 6 participants aspired to a research career but did not yet have one.

While most (52%) had no leadership role in relation to legal research, 25% were members of editorial boards, 17% had a university, faculty or school leadership role, 12% directed a Centre, 9% had a managing editor role (or similar), and 9% had another leadership role.

The gender split was 54% female, 40% male, and 5% non-binary or declining to answer. The expertise of participants was not evenly distributed across fields of research within the domain of law, with the highest numbers of researchers reporting that their primary research area was 'law in context' and 'international and comparative law'.

There are biases associated with who chose to participate in our survey. This includes the over-representation of some universities relative to others, the skew towards senior academics, and the higher participation of women. Overall statistics will reflect the biases of the sample. To mitigate against this, we analysed the responses separately for these cohorts and note where there are statistically significant differences associated with professors/non-professors, university type and gender. To do this, we use the tests recommended by Stats iQ (a tool available within Qualtrics). These tests are standard but rely on assumptions, specifically:

- T-test: For continuous, normally distributed data, Stats iQ employs Welsh's T-test under the assumption that the data within each group are normally distributed with few or no outliers. This test is ideal for comparing the means of two groups, providing insights into significant differences in the central tendency of a quantitative variable under two different conditions.

- Chi-Squared Test: For categorical data, Pearson's Chi-Squared test is utilized under the assumption of independence between observations. This non-parametric test is suitable for examining the relationship between two categorical variables, assessing whether the distribution of sample categorical data matches an expected distribution.

- Ranked T-Test: For data that are not normally distributed or for ordinal data, a rank-based T-test is applied. Qualtrics Stats iQ uses rank transformation followed by Welch's t-test.[8] This test is robust to outliers and does not assume normal distribution.

In relation to the Research Leaders Survey, we aimed to obtain a cross-section of responses based on (a) representatives of GO8 research intensive faculties, smaller schools and regional universities and (b) stand-alone law faculties and schools and faculties that are part of a larger faculty with other disciplines, for example business or social sciences. We received 26 responses. Participants came from 24 Australian universities, with 2 respondents choosing not to disclose their institution. Most participants held School/Faculty level leadership roles, for example, Research Director, Associate Dean, Dean or Head of School 25/96). About half of participants answered that 'law' was situated in their university as a stand-alone school or faculty (15/58%).[9]

[8] See discussion at https://www.qualtrics.com/support/stats-iq/analyses/statistical-test-assumptions-technical-details/#T-Tests.

[9] We attempted further research and found the law units of 31 institutions in Australia were part of a larger faculty, program or college whereas the law units of only 9 institutions were stand-alone faculties. These numbers were obtained by examining the websites of 41 institutions and ascertaining whether their law school (if the institution possessed one), was part of a larger faculty or college (e.g. Faculty of Criminology and Law, Faculty Business and Law, Arts, Faculty of Social Sciences and Law etc.). Obtaining this number, however,

We analysed qualitative data obtained in the Researchers Survey across 7 areas: characteristics of researchers (career stage, gender, institution to which the researcher belongs etc); subject areas of research; methods used in research (doctrinal, empirical etc); factors that determined researcher views about the quality of legal research; perceptions of the extent to which institutional and national assessment processes were fair and appropriate (e.g. ARC peer review panels); understandings of the meaning and modes of measurement of research impact; how research quality and impact was assessed in their own institution; the effect particular forms of assessment had on individual research decisions (e.g. in determining what kinds of research to pursue and what forms of output they considered important, and for what reasons).

We did a preliminary coding of researcher views about measurement of research excellence and impact, and then coded again on the basis of:

a) Whether the researcher was from a school/faculty in a GO8 or another university, which might indicate that factors to do with size and resourcing had an effect on how research was assessed and consequently had impact on perceptions of the utility and fairness of different forms of assessment;

b) Gender, in order to determine whether there are connections between perceptions of research excellence, and barriers to research performance and gender issues; and

c) The level of research seniority, to determine whether and if so the extent to which career stage was a factor in satisfaction or dissatisfaction with various forms of measurement.

Once an initial coding was complete, we examined themes emerging from the data. Dominant themes were perceptions about the use and utility of journal lists as a means for assessing research excellence; standards and consistency of peer review; the application of metrics, including journal citations, to Australian legal research; and differences in the way measurement regimes effected researchers based on their stage of career.

In relation to the Research Leaders Survey, we analysed the data on the basis of whether or not law was part of a standalone school or part of a broader faculty, to determine whether this had implications for the way that research excellence is measured (e.g. by use of journal lists such as the Australian Business Deans Council) and potentially for research culture and interdisciplinary research. We hypothesised that there may be a difference in perception about degrees of freedom to pursue own research goals as between research leaders (in the Research Leaders survey) and legal researchers (in the Researchers Survey).

We received a sufficient quantity and range of responses to enable us to draw the preliminary conclusions outlined below.

We presented a draft of this report at a LADRN (Australia's Law Associate Deans Research Network) meeting on 8 February 2024. We received thoughtful and helpful feedback from members. We also asked attendees whether they in principle endorsed our recommendations. We stated explicitly that we were only asking for their comment on the recommendations, not other sections of the report. One attendee was a delegate and needed to confirm with their ADR, the remainder (20) who

revealed difficulties. For example, UNSW has a Faculty of "Law and Justice" which includes Criminology and thus is not technically stand-alone (although two of the three schools within that Faculty are exclusively Law). Some other Faculties/Schools combine Law with a small number of people in a cognate discipline. Accordingly, self-perception may be more accurate than noting disciplinary affiliations or nomenclature on websites.

responded offered 'in principle' endorsement, although 2 of those expressed some reservations about Recommendation 5 (in line with their substantive comments earlier in the meeting). We had one LADRN member offer written feedback prior to the meeting.

In addition to members of LADRN who have supported this work through our meetings and through encouraging participation in the survey, we would like to particularly thank colleagues from the University of Sydney (Professor Emily Crawford), University of Melbourne (Professor Kimberlee Weatherall and Professor Rebecca Giblin), UNSW Sydney (Professor Theunis Roux, Professor Fleur Johns, Professor Kathy Bowrey), Western Sydney University (Dr Sandy Noakes), Tilburg University (Professor Rob van Gestel) and Australian National University (Professor Will Bateman) for their feedback, input and advice in carrying out this project. We take sole responsibility for any remaining errors and limitations.

Part (III)- Background

(a)- CALD Statement

In 2005, in response to the Australian Commonwealth Department of Education, Science and Training issues paper '*Research Quality Framework: Assessing the quality and impact of research in Australia* (**RQF**)', the Council of Australian Law Deans (**CALD**) published 4 principles for the evaluation of legal research (the **CALD Statement**).[10] In summary form, these were:

- Any assessment of quality and impact in relation to legal research must be informed by and sensitive to the nature and diversity of legal research, thus taking into account, fairly and transparently, the extent to which legal research is distinctive and the extent to which it is part of the mainstream of the humanities and social sciences. Peer review by discipline experts and external endorsement by relevant end users are clearly appropriate, so long as assessing panels are constituted having regard to this principle, especially to the diversity of the kinds of legal research. Implementation of the principle requires a law-specific panel, that is, a panel tasked to assess law, constituted predominantly by discipline, subdiscipline, and cognate experts.

- Measures to assess the quality and impact of legal research should be applied sensibly and in context, by informed peers, with due regard to the ways in which legal research is distinctive and, equally, the ways in which it is part of the mainstream of the humanities and social sciences. The range of criteria, including metric criteria, is not inappropriate, so long as it is applied to the discipline of law with care, and adapted, modified or supplemented as appropriate.[11] For example, the criterion of citation would need to be applied not just to other

[10] Council of Australian Law Deans, *Statement on the Nature of Legal Research* (Position Statement, October 2005) 1 <https://cald.asn.au/wp-content/uploads/2017/11/cald-statement-on-the-nature-of-legal-research-20051.pdf>.

[11] The CALD publication referenced, as an example, the sensible adoption of the metrics developed in the 2005 Allen Consulting Group report titled '*Measuring the Impact of Publicly Funded Research*': ibid 5. The Allen report suggested that research impact should be expressed as a social rate of return. This metric would be calculated by assessing benefits to society in the material, human, environmental and social dimensions after a time span of ten to twenty years had elapsed. The metrics used for assessment in each dimension would be varied. Impacts in the **material dimension** would be measured by the quality and diffusion of research as indicated by metrics such as: citation counts, patents and the numbers of presentations to industry and government knowledge users. The **human dimension** would be measured qualitatively by indicating how the

academic writings but also extended to court and tribunal judgments, transcripts of oral argument, documents of the parties constituting the written argument so far as that is on the public record, government reports. Attention would also need to be paid to the long-standing and well-recognised problems of citation measures, such as the difficulty of distinguishing positive from negative citation. A non-quantitative evaluative context is required.

- 'Impact' is an appropriate and important criterion for research assessment, with both quantitative and qualitative aspects and with particular relevance to the discipline of law. The impact of legal research is often manifest in policy development and legal change, through adoption or reliance on the research by government, government agencies, statutory commissions, courts, tribunals, and even private bodies, though often with a substantial time lag that presents a challenge for evaluation confined to a relatively recent time period. The elevation of the importance of legal research published in submissions, working papers, and reports for parliamentary, government, law reform and international agencies is appropriate and welcome, and again has qualitative as well as quantitative aspects.

- The assessment of impact, and of quality more generally, cannot be purely quantitative. It must be anchored in peer review, be evidence based and incorporate qualitative judgment.[12] The incorporation of research results into international/national policies, codes and/or practices, media presence through articles, debates, coverage, and expert advice/submissions/panel membership at government enquiries are of particular importance to the discipline of law. Peer review could assess claims to public benefit. Serious consideration should be given to establishing a standing discipline-specific advisory panel to fine tune the quantitative measures (for example, ranking or otherwise evaluating the relevant journals) and to inject the qualitative element.

The CALD Statement pointed to the diversity of approaches to legal research and listed 10 types of legal research: doctrinal, theoretical, critical/reformist, fundamental/contextual, empirical, historical, comparative, institutional, process-oriented and interdisciplinary.[13] The CALD Statement acknowledged that there was a spectrum where some legal research sat comfortably in the humanities or social sciences (and could align with their traditions and practices) whereas others focussed more

research impacts peoples' health and happiness, for example if the research is translated into health advice. The **environmental dimension** would be measured by any calculation of material benefits, for example, if it was known how much salinity damage has been prevented in Australia due to Australian publicly funded research and how much this has contributed to agricultural output. The **social dimension** would be measured by a mixture of quantifiable immediate outputs, such as the number of submissions made to Parliamentary enquiries, the completion of a number of research projects, and interim outcomes, such as impacts on government policy: Allen Consulting Group, *Measuring the Impact of Publicly Funded Research* (Department of Education, Science and Training of the Australian Government, 2005) ch 6 <https://www.adelaide.edu.au/rqf/pdf/allenreport.pdf>.

[12] In this respect, the CALD Statement referenced the RQF parameters. These parameters were developed to include quantitative and qualitative measures, for example, reduced pollution, regeneration or arrested degradation of natural resources, increased literacy and numeracy rates, positive reviews of creative publications and performances, increased cultural awareness, licenses, changes in procedures, behaviours, outlook etc, new policies, guidelines, legislation etc, citations of research in legal judgments which become case law, contracts and industry funding, number of presentations involving contact with end-users, community awareness of research, citations in government reports, Hansard etc, provision of expert advice and submissions to enquires, invitations to be a visiting researcher or researcher in residence at an end-user institutions: Quality Framework Development Advisory Group, *Research Quality Framework: Assessing the Quality and Impact of Research in Australia* (Final Report, 2006) 1, 9.

[13] Council of Australian Law Deans (n 9) 2.

on doctrine and distinct in important ways. Quality assessments thus needed to be sensitive to the nature and diversity of legal research.

(b)- ERA and EI Exercises

As noted, the CALD statement was in response to the RQF, an initiative of the Howard Liberal government. It was announced in 2004 with various models proposed over 2005-2006. The proposal was to link the assessment of research quality and impact to the distribution of government funding.[14] The RQF was abandoned following the election of the Labor government in 2007 and replaced with the Excellence in Research for Australia (ERA) process, involving separate evaluation of "excellence" and "impact and engagement". The first round of ERA in 2010 relied on a tiered journal list, which was abandoned in 2011 (although the list itself survives as a zombie and is sometimes referred to inappropriately in hiring, promotion and grant applications).[15] Subsequent rounds took place in 2012, 2015 and 2018. The round scheduled for 2023 was cancelled and a new model is yet to be determined.

In December 2015, the Australian Government announced the development of an Engagement and Impact (EI) assessment to run alongside the ERA. The EI assessment examines how universities are translating their research into economic, environmental, social, cultural and other benefits.[16] In EI 2018, five assessment panels comprising researchers and research end-users[17] were responsible for assessment. The assessments comprised of the following:[18]

- **Engagement**—Panels assessed research engagement activity based on an engagement narrative, a small suite of quantitative indicators, and an engagement indicator explanatory statement. The engagement narrative explained the interaction between researchers and research end-users outside of academia for the mutually beneficial transfer of knowledge, technologies, methods or resources.

- **Impact Studies**—Panels assessed research impact and the institution's approach to impact based on qualitative impact studies that detailed the impact, the research associated with the impact, and the approach to impact for each UoA (Unit of Assessment). Each UoA received two ratings—one for impact and one for approach to impact.

The ARC commenced a comprehensive review of ERA and EI in 2020.[19] Among the review's findings were the following:

---

[14] Research Quality Framework Development Advisory Group, *Research Quality Framework: Assessing the Quality and Impact of Research in Australia* (2006).

[15] See discussion at '
' below.

[16] Australian Research Council, 'Engagement and Impact Assessment' <https://www.arc.gov.au/evaluating-research/ei-assessment>.

[17] A research end-user is an individual, community or organisation external to academia that will directly use or directly benefit from the output, outcome or result of the research. Examples of research end-users include governments, businesses, non-governmental organisations, communities and community organisations: Australian Research Council, 'Introduction' (Web Page) <https://dataportal.arc.gov.au/EI/NationalReport/2018/pages/introduction/index.html?id=background>.

[18] Ibid.

[19] ERA EI Review Advisory Committee, *ERA EI Review 2020–2021* (Final Report, Australian Research Council, 2021) 5; Kate Williams et al, *An Evaluation of the Impact Component of the Australian Research Council's 2018 Engagement and Impact Assessment* (Report, 2020) <https://www.arc.gov.au/sites/default/files/final_arc_report.pdf>.

- The use of narratives remains the fairest and most effective way to assess a diverse range of EI strategies. Although the use of peer review and citation analysis on a discipline-specific basis reflects international best practice, the Advisory Committee considered there were ways to improve the transparency and fairness of the assessment process.[20]

- While HEIs were able to identify and articulate impact, challenges arose around understanding guidelines, articulating approach to impact and evidencing impact.[21]

- Stakeholders raised concerns regarding the differing performances of citation analysis (largely STEM disciplines) and peer review disciplines (largely HASS disciplines) in ERA. Since 2010, the number of STEM disciplines receiving above world standard ratings in ERA has increased at a faster rate than HASS. The underlying reasons were stated by the Advisory Committee to be complex, involving factors beyond the scope of ERA, such as varying profiles in research staff, income, investment and capacity across different disciplines.[22]

- Stakeholders expressed concerns that ERA peer review methodology is not comparable to the citation methodology, and that the concept of 'world standard'[23] may unintentionally be set higher for peer review disciplines than in citation analysis disciplines.[24]

- Sector feedback has also highlighted concern regarding optimisation or potential gaming in citation disciplines.[25] Opportunities for optimisation are greater in citation disciplines. By modelling citation profiles in advance, universities are able to strategically assign research outputs to certain FoR codes to achieve the best possible results. However, in peer review disciplines, it is harder for universities to predict and optimise performance in this way.

- Similarly, in the citation methodology, evaluators have a limited ability to detect and act against miscoding.[26] In peer review disciplines, evaluators can identify and discount miscoded outputs more easily.

- The assessment of impact as part of EI 2018 was considered a significant new resource burden for HEIs.[27]

As a result, a set of 22 recommendations was proposed in the Report which were subsequently adopted in full by the ARC.[28] Among the recommendations were:

- **Recommendations 6 and 7:** That the ARC revise the ERA and EI rating scale, the benchmarks for citation and peer review assessment, the definition and appropriateness of 'world standard' and the relevant guidance material for ERA assessors.
  .
- **Recommendation 19:** That the dual methodologies of citation analysis and peer review continue in ERA, that disciplines continue to be assessed using the most appropriate

[20] ERA EI Review Advisory Committee (n 18) 3.1.
[21] Williams et al (n 18) 4.1.
[22] ERA EI Review Advisory Committee (n 18) 3.2.1.
[23] A rating of 3 on the ERA five-tier rating scale indicates performance at world standard.
[24] ERA EI Review Advisory Committee (n 18) 3.2.1.
[25] Ibid.
[26] Ibid.
[27] Williams et al (n 18) 3.2.
[28] ERA EI Review Advisory Committee (n 18) 7–9; Australian Research Council, 'ERA EI Review' (2022).

methodology, and that ANZSRC 2020 FoR codes define Units of Evaluation in ERA and Units of Assessment in EI.

- **Recommendation 20:** That further steps be taken to ensure the robustness of the peer review and citation methodologies used in ERA. In particular:

    - Improve the application of the peer review methodology to ensure the appropriate application of 'world standard', increase the size, quality and diversity of the peer reviewer pool, and improve training of peer reviewers.

    - Provide Research Evaluation Committees with authority to exclude Units of Evaluation or research outputs where significant miscoding has occurred and request a recalculation of citation profiles

- **Recommendation 21:** That the ARC substantially retain the existing definitions for EI, with the following adjustments:

    - That the definition of 'research end-user' for EI be expanded to include publicly funded research organisations, with guidance emphasising the nature of the impact or intent of the activity rather than the type of organisation.

    - That the ARC develops additional guidance with examples to support the definitions of impact, engagement, and research end-user and that the ARC develops a definition of 'approach to impact.'

- **Recommendation 22**: That the ARC continue to monitor international and best practice understandings of research excellence and investigate how they may be incorporated into future rounds of ERA and EI.

The most recent development has been a pause on the ERA and EI scheme in order to consider the best way forward.[29] This is occurring through various consultation processes. At the centre of the conversation is that around a Universities Accord, announced in December 2022. The Final Report was presented to the Minister for Education on 28 December 2023 but has not yet been publicly released.[30] There is one legal academic member of the panel undertaking the review, Professor Larissa Behrendt.[31] Research quality is one topic within the review, although it has largely been overshadowed by other issues. In May 2023, Nous released a summary of the submissions made to the Accord review – 33 submissions addressed ways to maintain and improve research quality, with some critiquing the ERA methodology, pointing out the potential for gaming and discipline bias.[32] The Interim Report, issued

[29] Australian Research Council, 'ERA 2023' <https://www.arc.gov.au/evaluating-research/excellence-research-australia/era-2023>.
[30] Australian Government Department of Education, 'Australian Universities Accord', *Australian Universities Accord* (Web Page, 19 July 2023) <https://www.education.gov.au/australian-universities-accord>.
[31] Australian Government Department of Education, 'Australian Universities Accord Panel' (Web Page, 2 January 2024) <https://www.education.gov.au/australian-universities-accord/panel>.
[32] Nous Group, *Australian Universities Accord Discussion Paper Submission Analysis* (Report, Nous Group, 12 May 2023) <https://www.education.gov.au/australian-universities-accord/resources/australian-universities-accord-discussion-paper-submission-analysis>.

May 2023, referred to the need for clear measures on the usefulness of research, including metrics, but did not go into specifics.[33]

(c)- International Approaches to Research Assessment

Australian developments in legal research quality and assessment have taken place in the context of similar international efforts.[34] Particularly relevant has been the experience of the United Kingdom. In the UK, research assessment exercises have existed since 1986, when the first Research Selectivity Exercise was undertaken.[35] The current UK system is the Research Excellence Framework ("REF"), which was first conducted nationwide in 2014 and repeated in 2021.[36] REF outcomes are used to inform the allocation of around £2 billion per year of public funding for universities' research.[37]

Assessment relies on expert panels, comprised of academics and industry experts for each of the 34 subject based Units of Assessment.[38] Panels conduct assessment in respect of three criteria. The first is the quality of research outputs measured in terms of originality, significance, and rigour.[39] This accounts for 60% of the assessment.[40] The second is the impact of research measured in terms of the effect on, change, or benefit to the economy, society, public policy or services, culture, health, the environment or quality of life, beyond academia.[41] This accounts for 25% of the assessment.[42]

The third is research environment, which is an institution's research strategy, research facilities, opportunities for collaboration and environment, including research income and research degrees awarded.[43] This accounts for 15% of the assessment.[44] Research environment is assessed in terms of its vitality and sustainability.[45] It considers how effectively research is supported at the whole-

---

[33] Australian Government Department of Education, *Australian Universities Accord Interim Report* (Interim Report, 19 July 2023) <https://www.education.gov.au/australian-universities-accord/resources/accord-interim-report>.

[34] For example, New Zealand have adopted the Performance Based Research Fund (PBRF). For the purposes of the PBRF, research is defined as a process of investigation or inquiry leading to new, recovered, or reinterpreted knowledge or understanding which is effectively shared and capable of rigorous assessment by the appropriate experts. Research excellence is assessed in terms of originality, rigour, reach, and significance, with reference to the quality standards appropriate to the subject area and to the unique nature of New Zealand's research. Research is assessed by expert peer review panels who consider the research activity of each eligible staff member presented in an Evidence Portfolio (EP). See Tertiary Education Commission, 'Performance Based Research Fund: Guidelines for the Quality Evaluation 2026 Assessment Process' 15, 18, 21 <https://www.tec.govt.nz/assets/Publications-and-others/Guidelines-for-the-Quality-Evaluation-2026-assessment-process.pdf>.

[35] Mehmet Pinar and Timothy J Horne, 'Assessing Research Excellence: Evaluating the Research Excellence Framework' (2021) 31(2) *Research Evaluation* 173, 174. This was followed by Research Assessment Exercises in 1996, 2001 and 2000.

[36] Research Excellence Framework, *Summary Report Across the Four Main Panels* (Report, 2021) [13]. The next exercise is planned for 2029. Research Excellence Framework, 'REF 2029' (Web Page) <https://www.ref.ac.uk/>.

[37] Research Excellence Framework, 'Your Simple Guide to REF' (Guide, 2021) <https://www.exeter.ac.uk/media/universityofexeter/ref2021/Simple_Guide_to_REF2021.pdf>.

[38] Research Excellence Framework, 'Summary Report Across the Four Main Panels' (n 35) [1].

[39] Birmingham City University, 'How REF Is Measured' (2021) <https://www.bcu.ac.uk/research/ref-2021/introduction-to-ref/how-ref-is-measured>.

[40] Research Excellence Framework, 'Your Simple Guide to REF' (n 36).

[41] Ibid; Birmingham City University (n 38). Impact is assessed from case studies which document the underpinning research and the impact resulting from the application and use of that research.

[42] Research Excellence Framework, 'Your Simple Guide to REF' (n 36); Birmingham City University (n 38).

[43] Research Excellence Framework, 'Your Simple Guide to REF' (n 36).

[44] Ibid.

[45] Birmingham City University (n 38).

institution level, the university's strategy and resources that support research, and the development of researchers.[46] The number of research doctoral degrees awarded and research income secured is also relevant.[47] The assessment of research environment is both backward looking, narrating progress made since REF2014, and forward looking, identifying the university's research strategy for the next five years.[48]

A major review of REF 2014 resulted in significant changes to the framework.[49] Among these was the decision to include all staff who have a significant responsibility for research in the assessment; and an expansion of the definition of impact to emphasise public engagement and impact on teaching.[50]

Several studies have analysed the issues and challenges with the UK attempt to assess the impact of research, including legal research.[51] Some express concern over evidence of bias towards more research-intensive universities in the assessment of 'research environment',[52] while others consider the extent to which the judgment of assessors may be affected by implicit bias and a 'halo effect' where assessors allocate higher scores to departments with long-standing records of high quality research.[53] The 2014 exercise was criticised on the basis of the high cost of preparing submissions and undertaking the exercises. [54]

National assessments of research excellence and impact are driven in part by an internationally competitive environment that ranks institutions, with consequences for research funding, international student applications and staff talent. Whether or not there is a formal national assessment process, universities will be measured by rankings bodies. Concerns about poor methodologies for measuring research has led several international bodies to suggest improvements.

(d)- DORA

The San Francisco Declaration on Research Assessment (DORA) was developed in 2012 to improve the ways in which researchers and scholarly research are evaluated.[55] DORA sought to transform the assessment of academic researchers by promoting the adoption of non-bibliometric metrics by funders, institutions, and publishers.[56] DORA emphasises the need to eliminate the use of journal-

[46] Ibid.
[47] Ibid.
[48] Ibid.
[49] Lord Nicholas Stern, *Building on Success and Learning from Experience: An Independent Review of the Research Excellence Framework* (Report, Government of the United Kingdom, July 2016).
[50] Research Excellence Framework, 'Your Simple Guide to REF' (n 36).
[51] See, eg, Mehmet Pinar and E Unlu, 'Evaluating the Potential Effect of the Increased Importance of the Impact Component in the Research Excellence Framework of the UK' (2020) 46 140; Lawrence McNamara, 'Understanding Research Impact in Law: The Research Excellence Framework and Engagement with UK Governments' (2019) 29(3) *King's Law Journal* 437: dealing with 'impact' assessment in the context of legal research using data from REF 2014 and 2021.
[52] Mehmet Pinar and E Unlu, 'Determinants of Quality of Research Environment: An Assessment of the Environment Submissions in the UK's Research Excellence Framework in 2014, Research Evaluation' (2020) 29 *Research Evaluation* 231.
[53] Horne (n 34).
[54] Horne (n 34). For the direct administrative costs of the 2021 REF exercise see *Research Excellence Framework 2021: REF Director's Report* (Report, REF, October 2023) [673] <https://archive.ref.ac.uk/media/1918/ref-directors-report.pdf>.
[55] DORA, 'Home', *DORA* (Web Page, 1 August 2023) <https://sfdora.org/>.
[56] Australian Council of Learned Academies, *Research Assessment in Australia: Evidence for Modernisation* (Report, 2023) 6.1.1.

based metrics, such as Journal Impact Factors, and the importance of exploring new indicators of significance and impact.[57]

In terms of research assessment, DORA recommends considering the value and impact of all research outputs (including datasets and software) in addition to research publications and considering a broad range of impact measures including qualitative indicators of research impact, such as influence on policy and practice.[58] DORA has been signed by 25 Australian organisations including the National Health and Medical Research Council, the Australian Academy of Science and the University of Melbourne.[59]

(e)- Leiden Manifesto

In the decade since DORA, international statements on the assessment of research excellence and impact have mushroomed. The Leiden Manifesto (**LM)** for Research Metrics was published in 2015, and sought to codify best practice principles in metrics-based research assessment.[60] The principles sought to use a diverse range of metrics based on the research field and nature of research, to reduce potential misuse of metrics and ensure greater transparency in the use of metrics.

Motivations for codification of the LM arose from a growing worry that "impact-factor obsession" was leading to inadequate judgement of scientific material that should be worthy of fair evaluation.[61] In 2017, the European Association for Research Libraries published a substantial review on the Leiden Manifesto.[62] It concluded that the Manifesto was a "solid foundation" on which academic libraries could base their assessment of metrics.[63] Elsevier, a global leader in research publishing and information analytics, subsequently endorsed the LM in its guide to the development of improved research evaluation.[64]

(f) - Hong Kong Principles

International progress continued in 2019 when the 6th World Conference on Research Integrity formulated and endorsed the Hong Kong Principles for assessing researchers.[65] The Hong Kong Principles were chosen with a view to explicitly recognising and rewarding research integrity. The

---

[57] Ibid.
[58] Ibid appendix 8.
[59] Ibid 6.1.1.
[60] Diana Hicks et al, 'The Leiden Manifesto for Research Metrics' (2015) 520 *Nature* 429; 'Leiden Manifesto for Research Metrics' (Web Page) <http://www.leidenmanifesto.org/>. The ten principles were as follows: 1) quantitative evaluation should support qualitative expert assessment, 2) performance should be measured against the research missions of the institution group or researcher, 3) protect excellence in locally relevant research, 4) keep data collection and analytical processes open transparent and simple, 5) allow those evaluated to verify data and analysis, 6) account for variation by field in publication and citation practices, 7) base assessment of individual researchers on a qualitative judgement of their portfolio, 8) avoid misplaced concreteness and false precision, 9) recognise the systemic effects of assessment and indicators, 10) scrutinise indicators regularly and update them.
[61] Hicks et al (n 59) 429.
[62] S.K Coombs and I Peters, 'The Leiden Manifesto under Review: What Libraries Can Learn from It' (2017) 33(4) *Digital Library Perspectives* 324.
[63] Ibid 336.
[64] Elsevier, 'Our Commitments to Responsible Research Evaluation', *www.elsevier.com* <https://www.elsevier.com/en-au/insights/icsr/responsible-research-evaluation>.
[65] The World Conferences on Research Integrity Foundation, 'Hong Kong Principles' <https://www.wcrif.org/guidance/hong-kong-principles>.

principles emphasized the need to 1) foster responsible research practices, 2) enable transparent reporting, 3) promote open science and 4) value various types of research and recognise all contributions to research and scholarly activity.[66]

(g) - European Commission

In 2021, the European Commission published a report which built on DORA, the LM and the Hong Kong Principles.[67] The report proposes a set of principles to evaluate research and researchers based on their intrinsic merits and performance, rather than on the number of publications and where these are published.[68] It seeks to promote qualitative judgement with peer-review, supported by a more responsible use of quantitative indicators.[69]

(h) - CoARA

This emphasis on qualitative evaluation and the centrality of peer review is also articulated in the 2022 Agreement of the Coalition for Advancing Research Assessment (CoARA). CoARA involved over 350 organisations from more than 40 countries and builds on international progress made through DORA, the LM and the Hong Kong Principles.[70]

The Agreement articulates shared principles such as support for diversity, inclusiveness and collaboration in research.[71] It emphasises the need to be respectful of the autonomy of organisations and their respective missions or strategies, including by appreciating that assessment practices can vary according to the context, type, and purpose of the evaluation.[72] The Agreement proposes 10 commitments,[73] some of which include: the importance of qualitative evaluation and peer review, the responsible use of quantitative indicators, the rejection of inappropriate uses of Journal Impact Factor

[66] Australian Council of Learned Academies (n 55) appendix 7; The World Conferences on Research Integrity Foundation (n 64).
[67] Directorate-General for Research and Innovation, *Towards a Reform of the Research Assessment System* (Scoping Report, European Commission, 2021) 8 <https://data.europa.eu/doi/10.2777/707440>.
[68] Ibid 8–10.
[69] Ibid 8.
[70] Australian Council of Learned Academies (n 55) 6.1.2; Coalition for Advancing Research Assessment, 'CoARA' <https://coara.eu/>.
[71] Coalition for Advancing Research Assessment, 'Agreement on Reforming Research Assessment' 2–3 <https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf>.
[72] Ibid.
[73] The commitments are to 1) recognise the diversity of contributions to, and careers in, research in accordance with the needs and nature of the research, 2) base research assessment primarily on qualitative evaluation for which peer review is central, supported by responsible use of quantitative indicators, 3) abandon inappropriate uses in research assessment of journal- and publication based metrics, in particular inappropriate uses of Journal Impact Factor (JIF) and h-index, 4) avoid the use of rankings of research organisations in research assessment, 5) commit resources to reforming research assessment as is needed to achieve the organisational changes committed to, 6) review and develop research assessment criteria, tools and processes, 7) raise awareness of research assessment reform and provide transparent communication, guidance, and training on assessment criteria and processes as well as their use, 8) exchange practices and experiences to enable mutual learning within and beyond the Coalition, 9) communicate progress made on adherence to the Principles and implementation of the Commitments and 10) evaluate practices, criteria and tools based on solid evidence and the state-of-the-art in research on research, and make data openly available for evidence gathering and research: Ibid 3–9.

(JIF) and h-index measurements, and the avoidance of the use of rankings for research organisations in research assessment.[74] The agreement stipulates a 1 and 5 year timeframe for reform.[75]

CoARA is notable in its insistence on the inclusion of diverse outputs, practices and activities; and flexibility in evaluation practices.[76] At the time of writing, no Australian organisation had signed the agreement. However, there are over 500 European organisation signatories, and a massive international collaborative effort underway to implement it.[77] These recent international movements indicate a global shift away from over-reliance on metrics in quality and impact assessment.

(i) - ACOLA

These international developments have also had an impact in Australia. In November 2023, the Australian Council of Learned Academies (**ACOLA**) published a report on the topic of Modernising Research Assessment.[78] Looking across disciplines, it expresses concern about current mechanisms of quality assessment, including overreliance on citation metrics, journal rankings and grant funding. It also points out the disparate impact of research assessment activities on researchers from disadvantaged backgrounds. There are a number of statements and recommendations in that report that are particularly pertinent to the assessment of legal scholarship:

> As the landscape of research assessment evolves, it is essential to strike a balance between quantitative metrics and qualitative evaluation and reflect the multidimensional nature of research impact. By addressing these concerns and fostering a more holistic approach to research assessment, Australia can move away from a focus on quantity (for example, number of publications, citations, grants received) towards valuing research quality, excellence and impact.[79]
>
> There has been a decades-long shift towards bibliometric assessment of research outputs in the global academic community. While there is some merit in this approach, it has also led to a complex and international pattern of dependence across universities, research publishers, research funders and global ranking agencies that has had an outsized impact on the type and nature of research that is undertaken in universities. For example, bibliometric assessment of research prioritises journal articles over other types of output, creating a bias against interdisciplinary work and research in many humanities and social science fields.[80]
>
> Journal-based metrics and rankings can also be a cause of concern for researchers working on Australian-focused subject matter. This is because many smaller country-based journals of the type that are critical for the publication of much Australian and Indigenous studies research are not included in journal ranking lists and therefore carry less weight.[81]
>
> Many respondents [in their national survey] raised concerns about the apparent inability of Australia's research assessment system to appropriately recognise and support new and innovative ideas. Pressure to publish – particularly with high-ranking international journals, prestige book presses or influential conference proceedings – and to do so through established

[74] Ibid.
[75] Ibid 11.
[76] Australian Council of Learned Academies (n 55) 6.1.2.
[77] Ibid.
[78] Australian Council of Learned Academies (n 55).
[79] Ibid 17.
[80] Ibid 33 (citations omitted).
[81] Ibid 34.

areas of inquiry further compounds this problem. The narrow focus of prioritised research assessment metrics and the decline in fundamental research funding negatively impact the research community's ability to generate new knowledge and foster innovation over the long term.[82]

Making changes to the culture of research and the ways in which it is assessed will help foster diverse workforces that are innovative and collaborative, while supporting dynamic and original research and disincentivising unwanted approaches such as gaming the system. No single metric can appropriately assess research or researchers. Several tools will be needed, including both quantitative indicators and metrics, and qualitative measures and peer review.[83]

Part (IV) - Literature review

Our project is an example of what Susan Bartie described as legal "meta-scholarship" in that its purpose is to understand how scholarly work in law ought to be assessed, rather than contributing to legal scholarship as such.[84] Accordingly, in this Section, we review other scholarship in that category, focussing in particular on works with a similar purpose to ours and limited, of necessity, to that published in English. The jurisdictional focus in our methodology is explained above, but we do not analyse each jurisdiction separately as similar themes emerge in different places. This section looks firstly at motivations for quality assessment of legal scholarship, secondly outlines some of the opposition to quality assessment, then thirdly turns to the different methods that might be deployed and what might be said in favour or against each. It then discusses the different axis along which legal scholarship is increasingly measured, namely impact.

In this review, we focus on explanations in the literature for the importance of assessing the quality and impact of *legal* scholarship. It is worth noting, however, that the existence of interdisciplinarity and cross-disciplinary research that incorporates legal components and the need to compare legal scholars with scholars from other disciplines (e.g. in the context of competitive funding) means that the broader and narrower justifications for quality assessment cannot be completely separated.[85]

(a) - Motivations for quality assessment

There are a variety of reasons why different actors might wish to assess and compare the quality of different legal scholarly outputs, different legal scholars (taking their works as a whole) or different institutions (taking their works and/or scholars as a whole). Often, the primary reason relates to a broader exercise involving *all* disciplines. For example, where the relative performances in the research of an institution is being measured, law units could not meaningfully opt out. As van Gestel notes, "[l]aw as a discipline cannot afford to do nothing because then, sooner or later, others will impose their rankings and quality management systems."[86]

[82] Ibid 35.
[83] Ibid 51.
[84] Susan Bartie, 'The Impact of Legal Meta-Scholarship: Love Thy Navel' (2009) 18(3) *Griffith Law Review* 727.
[85] Willem van Boom and Rob van Gestel, 'Evaluating the Quality of Dutch Academic Legal Publications: Results from a Survey' (2017) 13(3) *Utrecht Law Review* 9.
[86] Rob van Gestel, 'Sense and Non-Sense of a European Ranking of Law Schools and Law Journals' (2015) 35(1) *Legal Studies* 165, 179.

A theory of evaluation has been said to be essential for any disciplinary endeavour.[87] In the sciences, evaluation is based around the ability of a work to accurately describe observable phenomena in the natural world, in particular identifying causal relationships and accurately predicting future events.[88] Legal scholarship is said to be different as it is associated with a range of purposes, most of which are not primarily concerned with description. One description of "standard" legal scholarship is that it comprises prescriptions addressed to public decision-makers such as judges, legislators, and administrators.[89] Hutchinson's term for this kind of research is 'reform-oriented research' and describes it as research which intensively evaluates the adequacy of existing rules and which recommends changes to any rules found wanting.'[90] Our data indicates that this is a common type of research carried out by researchers in academic institutions in Australia (40% of participating researchers).[91] But there is no uniformity of purpose in legal research as is the case for many scientific disciplines.

Pragmatically, there are a variety of contexts in which academics are compared, often by people outside the sub-discipline or even the discipline. These include hiring and promotion processes, where of necessity one is being judged by people whose own expertise may not squarely align with that of the candidate. Another important context is assessment of applications for competitive funding where applicants' worthiness must be compared across fields of research. In such cases, the main alternative to assessing quality might be to focus on *quantity* alone, which is undesirable unless one wants to reward and thus incentivise *more* rather than *better* legal scholarship.[92] An even less desirable alternative is to rely on gut feeling, which, beyond problems of subjectivity, has been shown to reinforce race and gender prejudices about what merit looks like or the ideological preferences of of a group.[93]

Institutionally, scholars have described cultural forces pushing for assessment. This might be described as 'new public management', 'neo-liberalism' or 'audit culture'.[94] A more positive frame that has been described is the drive for greater transparency and accountability.[95] This might be particularly important given a public funding model for many Australian universities – if taxpayers are funding research, the "bang for their buck" needs to be explained.[96] Again, compared to alternatives such as quantity and gut feel, the scholarship is clear on a need to focus on quality.

---

[87] Edward L Rubin, 'On Beyond Truth: A Theory for Evaluating Legal Scholarship' (1992) 80(4) *California Law Review* 889.
[88] Ibid 902.
[89] Edward L Rubin, 'The Practice and Discourse of Legal Scholarship' (1998) 86(8) *Michigan Law Review* 1835, 1850.
[90] Terry Hutchinson, 'Setting the Scene: Research and Writing in Context' in *Researching and Writing in Law* (Thomson Reuters (Professional) Australia Pty Limited, 3rd ed, 2010) 7.
[91] Catherine Renshaw and Lyria Bennett Moses, 'Council of Australian Law Dean's Survey'.
[92] Linda Butler, 'Explaining Australia's Increased Share of ISI Publications—the Effects of a Funding Formula Based on Publication Counts' (2003) 32(1) *Research Policy* 143, 143–155.
[93] Rubin (n 87); Janet Chan, Fleur Johns and Lyria Bennett Moses, 'Academic Metrics and Positioning Strategies' in Btihaj Ajana (ed), *Metric Culture: Ontologies of Self-Tracking Practices* (Emerald Group Publishing, 2018) 177.
[94] Kathy Bowrey, 'Audit Culture: Why Law Journals Are Ranked and What Impact This Has on the Discipline of Law Today' (2013) 23(2) *Legal Education Review* 291; Margit Osterloh and Bruno S Frey, *Research Governance in Academia: Are There Alternatives to Academic Rankings?*, CESIFO Working Paper 2797, September 2009; Kimberlee Weatherall and Rebecca Giblin, *Inoculating Law Schools against Bad Metrics* (Working Paper, 25 January 2021).
[95] Gestel, 'Sense and Non-Sense of a European Ranking of Law Schools and Law Journals' (n 86).
[96] Christopher Arup, 'Research Assessment and Legal Scholarship' (2008) 18(1 and 2) *Legal Education Review* 31.

If one is to assess quality, there are good reasons to do it in a transparent and nationally or regionally uniform way.[97] When one is looking at more than one paper or scholar and comparisons are being made across institutions or even at a global scale, quality assessment often turns to quantitative proxies. No single person has the time or expertise to make subjective comparisons of scholarship at scale and, even if they did, they would have their own biases. That leads to a turn to 'objective' metrics. As one author writes, "[t]he very existence (and persistence) of such biased indicators and rankings seems to be a consequence of the unwritten rule that *any number beats no number*".[98] In other words, if one wants to, for example, reassure Australian taxpayers that their money on legal scholarship is well spent due to the relative quality of Australian legal scholarship compared to world standard, one may well look to (even imperfect) rankings and scores.

(b) - Opposition to quality assessment

Not everyone agrees that assessing the quality of legal research is a good idea. As Svantesson and White write, "[t]here is something fundamentally absurd about the idea of ranking research".[99] For a start, it is hard to meaningfully compare different types of legal research with different methodologies and purposes.[100] Doctrinal research, in particular, often struggles when measured against social science methodological standards.[101] This is despite its importance in increasing societies' understandings their own laws and in leading to academic synthesis and innovation for the legal system.

One of the more significant concerns about assessment is goal displacement. Any proxy for quality can be gamed and has thus been said to drive behaviour of legal scholars in non-productive ways (or ways that would be unproductive from the perspective of at least some stakeholders).[102] This is sometimes referred to as Goodhart's law: as soon as measure becomes a target, it ceases to be a good measure.[103] There are a variety of potentially negative consequences of measurement referred to in the literature including pressure to do something other than curiosity-driven research,[104] decreasing heterogeneity in research,[105] a reduced focus on professional audiences that might benefit,[106] and publication choices driven by metrics rather than audience alignment.[107]

In addition to the problem of gaming, scholars describe other negative impacts of a regime of quality assessment. For example, it may lead to heightened anxiety among legal scholars with a resulting

---

[97] Rob van Gestel and Jan Vranken, 'Assessing Legal Research: Sense and Nonsense of Peer Review versus Bibliometrics and the Need for a European Approach' (2011) 12(3) *German Law Journal* 901.
[98] Yves Gingras, *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*, ed Blaise Cronin and Cassidy Sugimoto (The MIT Press, 2014).
[99] Dan Jerker B Svantesson and Paul White, 'Entering an Era of Research Ranking - Will Innovation and Diversity Survive?' (2009) 21(3) *Bond Law Review* 173.
[100] Gestel, 'Sense and Non-Sense of a European Ranking of Law Schools and Law Journals' (n 86).
[101] Rob Van Gestel, 'Ranking, Peer Review, Bibliometrics and Alternative Ways to Improve the Quality of Doctrinal Legal Scholarship' in Rob van Gestel, Edward Rubin and Hans-W Micklitz (eds), *Rethinking Legal Scholarship* (Cambridge University Press, 2017) 352.
[102] Margit Osterloh and Bruno S Frey, 'Ranking Games' (2015) 39(1) *Evaluation Review* 102.
[103] 'Goodhart's Law', *Oxford Reference* <https://www.oxfordreference.com/display/10.1093/oi/authority.20110803095859655>.
[104] Gestel, 'Ranking, Peer Review, Bibliometrics and Alternative Ways to Improve the Quality of Doctrinal Legal Scholarship' (n 101) 379.
[105] Osterloh and Frey (n 102).
[106] Boom and Gestel (n 84); Gestel, 'Sense and Non-Sense of a European Ranking of Law Schools and Law Journals' (n 86) 919.
[107] Gestel, 'Sense and Non-Sense of a European Ranking of Law Schools and Law Journals' (n 86) 919.

negative impact on mental health and personal wellbeing.[108] It may do so without any positive effect on quality – after all, most academics are motivated by curiosity and/or social impact which means quality assessment exercises can be demotivating rather than drivers for improvement.[109] The concern has also been expressed that conservatism in the context of an assessment regime will promote 'mainstream' conformist, rather than potentially more valuable 'controversial', scholarship.[110] These mirror the observations of ACOLA in its broader disciplinary context, set out above.[111]

(c) - Approaches to Quality Assessment

There are a variety of different approaches that might be taken to quality assessment, assuming one wished to (or was forced to) engage in such an exercise. Before looking at the literature internationally, it is worth highlighting the extensive work that Professor Kathy Bowrey did for CALD, based on a 2012 commission to write a report on "Assessing Research Performance in the Discipline of Law".[112]

This report provides a holistic picture of the way that research was counted and assessed at that time, as well as pointing out some of the problems. These included confusion about the treatment of non-peer-reviewed and partly peer-reviewed journals, mixed practices in peer review more broadly, the distorting effect of reliance on obsolete journal ranking lists, differing interpretations on the scope of various FoR codes and the need to evolve such codes, and the fact that scholars in different sub fields and in different institutions might be differently placed in terms of access to well-regarded journals.

Bowrey also commented strongly on the negative impacts of research assessment, stating: "It is perverse that the current research assessment climate contributes to the fragmentation and fracturing of the discipline of law as a whole, rendering our capacity to grow, to judge and assess all legal research fairly a less and less attainable goal."[113] A follow up report in 2016 delved deeper into the methodologies underlying law journal rankings, pointing out the flaws in each and the capacity of flawed approaches to cause harm and distort scholarship.[114]

In the remainder of this section, we describe each of the main forms of quality assessment – criteria, journal lists and citation metrics – and set out some of the main critiques of each. The concerns raised by Bowrey in her reports continue to be relevant to each.

Approach 1: Criteria

One way to assess the quality of legal scholarship is for a person or group of people to read it and then assess it according to particular criteria. This is similar to the approach taken in peer review, where assessors may be asked about originality, clarity and so forth. Just as there is no consistency in peer

[108] Bowrey, 'Audit Culture: Why Law Journals Are Ranked and What Impact This Has on the Discipline of Law Today' (n 94).
[109] Ibid; UNSW Law Faculty, *What Makes You Tick?: Report on a Survey of the Factors That Condition High Quality Research* (UNSW Law Faculty, 2012) ('*What Makes You Tick?*').
[110] Marita Carnelley, 'In Search of the Perceived Quality and Impact of Accredited South African Law Journals: Exploring the Possibility of a Ranking System. A Baseline Study: 2009 – 2014' (2018) 21 *Potchefstroom Electronic Law Journal* 1.
[111] See '(i) - ACOLA' above.
[112] Kathy Bowrey, *Assessing Research Performance in the Discipline of Law: The Australian Experience with Research Metrics, 2006-2011* (Report to CALD, 2012).
[113] Ibid 120.
[114] Kathy Bowrey, *A Report into Methodologies Underpinning Australian Law Journal Rankings Prepared for the Council of Australian Law Deans* (2016).

review criteria,[115] different lists of criteria have been suggested for quality assessment more broadly. Proposed criteria from the literature include:

- normative clarity, persuasiveness, significance, applicability (in that a work contains an identifiable original insight that can be used by others);[116]
- originality, thoroughness, and profundity;[117]
- general research criteria such as the presence of an explicit research question, relevance, alignment between conclusions and the research question, novelty, balanced (non-ideological) approach, conclusions aligned to evidence;[118] and
- methodological rigour.[119]

Some of these criteria are controversial or have been said to risk favouring certain types of legal scholarship over others. For example, the originality of empirical scholarship based on new data may be easier to demonstrate than the originality of a particular doctrinal analysis.[120] Assessing thoroughness can also be culturally and jurisdictionally dependent, with US law journals placing more emphasis on extensive citations as demonstrating thoroughness, while Europeans (and Australians) prefer a more succinct style.[121]

Methodological rigour has been said to be a challenging criterion for doctrinal scholarship and may risk narrowing the range of acceptable methods,[122] although others have argued for the benefits of doctrinal scholars articulating theoretical and philosophical foundations of their work and describing their process.[123] The idea of ranking methodologies and approaches[124] would likely be more controversial and problematic than assessing methodological rigour from a more neutral stance.[125]

The starkest problem for this approach to assessment, whichever criteria one chooses, is one of subjectivity and scale. Peer reviews are not always reliable, although they remain a crucial "input control".[126] Re-reviewing everything is too much work, particularly given peer review will already have occurred, but sampling methods are also potentially controversial.[127] Expert panels can be subjective, biased and self-perpetuating because academics might prefer scholarship similar to that which they

[115] Rob van Gestel, 'Sense and Non-Sense of a European Ranking of Law Schools and Law Journals' (2015) 35(1) *Legal Studies* 165; Jadranka Stojanovski et al, 'Peer Review in Law Journals' (2021) 6 *Frontiers in Research Metrics and Analytics*; Gestel and Vranken (n 97).
[116] Rubin (n 87) (where the evaluator is in the same subdiscipline, that is shares ideological and methodological commitments of the author; otherwise these should be tempered by asking whether a piece generates doubt or anxiety about the evaluator's own approach).
[117] Gestel and Vranken (n 97).
[118] Rob van Gestel, 'Quality, Methodology, and Politics in Doctrinal Legal Scholarship' [2022] *Law and Method*.
[119] Reza Dibadj, 'Translantic Publication Fashions In Search of Quality and Methodology in Law Journal Articles' in Rob van Gestel, Edward Rubin and Hans-W Micklitz (eds), *Rethinking Legal Scholarship* (Cambridge University Press, 2017).
[120] Mathias M Siems, 'Legal Originality' (2008) 28(1) *Oxford Journal of Legal Studies* 147.
[121] Gestel and Vranken (n 97).
[122] Philip C Kissam, 'The Evaluation of Legal Scholarship' (1988) 63(2) *Washington Law Review* 221.
[123] Theunis Roux, 'Judging the Quality of Legal Research: A Qualified Response to the Demand for Greater Methodological Rigour' (2014) 24 *Legal Education Review* 173.
[124] JB Ruhl, 'The Hierarchy of Legal Scholarship', *Jurisdynamics* (2006) <http://jurisdynamics.blogspot.com/2006/09/hierarchy-of-legal-scholarship.html>.
[125] Siems (n 120).
[126] Gestel, 'Ranking, Peer Review, Bibliometrics and Alternative Ways to Improve the Quality of Doctrinal Legal Scholarship' (n 101).
[127] Arup (n 96).

produce.[128] That suggests larger, more diverse panels are warranted, but this has significant workload implications.[129] If criteria-based assessment is favoured, it might be preferable to rely on the "input control" of peer review and consider all that has gone through that process to be of the acceptable standard for scholarly discourse.[130] Alternatively, one can supplement peer review with mechanisms for self-reflection and feedback within academic units alongside consideration of select outputs in the context of prizes.[131]

Approach 2: Journal lists

Another approach is to assume a correlation between the quality of outputs and the venue in which they are published. This means creating lists with tiers of journals and publishers, and then, where relevant, scaling through a formula that fairly compares scholars or institutions.

There are a variety of studies, in Australia and elsewhere, that seek to rank outputs, scholars or institutions based on journal lists. In Australia, this includes studies by Murray and Skead[132] and Smyth.[133] In the US, this includes comparison projects for 'productivity' based on numbers of publications in "top" law journals sometimes factoring in length, co-authorship, home journal bias and other factors.[134]

Analysis of law journals in Australia shows that coverage across legal subject matter in highly-ranked journals is uneven with some specialisations well served in journal choices while others are not.[135] Currently, there tends to be a focus in high ranked journals on five areas of law: constitutional, crime, labour law, evidence and administrative law.[136]

In addition to subject-matter bias, there are challenges in the creation of the list itself. Lists that already exist are problematic. The construction of the ERA 2010 list was somewhat arbitrary and political and is in any event out of date.[137] Most lists are biased in favour of particular jurisdictions, with the dominant US lists biased in favour of US journals.[138] For example, the influence of the US journal rankings in formulating ERA law journal rankings resulted in US publications being ranked more highly than several prestigious Australian law journals.[139] This in turn resulted in articles on important jurisdiction-specific legal issues becoming less publishable, and academics directing their research to areas and topics that are more likely to be of interest to international audiences.[140] As a result, journal

[128] Carnelley (n 110).
[129] Arup (n 96).
[130] Osterloh and Frey (n 102).
[131] Ibid.
[132] Murray and Skead (n 2).
[133] Russell Smyth, 'Who Publishes in Australia's Top Law Journals?' (2012) 35(1) *University of New South Wales Law Journal* 201.
[134] Brian Leiter, 'Measuring the Academic Distinction of Law Faculties' (2000) 29(1) *Journal of Legal Studies* 451; Robert E Steinbuch, 'On the Leiter Side: Developing A Universal Assessment Tool for Measuring Scholarly Output by Law Professors and Ranking Law Schools' (2011) 45 *Loyola of Los Angeles Law Review* 87.
[135] Bowrey, 'A Report into Methodologies Underpinning Australian Law Journal Rankings Prepared for the Council of Australian Law Deans' (n 114).
[136] Benedict Sheehy and John Dumay, 'Examining Legal Scholarship in Australia: A Case Study' (2021) 49(1) *International Journal of Legal Information* 32, 48.
[137] Bowrey, 'Audit Culture: Why Law Journals Are Ranked and What Impact This Has on the Discipline of Law Today' (n 94).
[138] Gestel and Vranken (n 97).
[139] Terry Hutchinson and Nigel Duncan, 'Defining and Describing What We Do: Doctrinal Legal Research' (2012) 17(1) *Deakin Law Review* 83, 87.
[140] Ibid.

ranking as a nation-wide exercise has been rejected as an approach in both Australia and South Africa.[141] Suggested 'objective' methods, such as rejection rates, have also been proven to be flawed.[142]

One suggestion from the literature is to create a list by fiat, nominating (even arbitrarily) which journals and publishers are in which tier.[143] This might then become a self-fulfilling prophesy in that, going forward, and removing the possibility of retrospective assessment based on the lists, the behaviour of authors and reviewers would reinforce that 'truth', at least locally.[144] This has been said to leverage rather than disparage the self-fulfilling prophecy nature of journal rankings.[145] Therefore, at some point in the future, a journal list that is flawed on inception will correlate more closely with quality.

Approach 3: Citation metrics

A popular way of assessing the quality of work in scientific disciplines is reliance on bibliometrics and, in particular, citation metrics. Such citation metrics might operate at the level of a piece of scholarship or author (e.g. citation counts, h indices) or at the level of journals (e.g. SCImago journal rank, CiteScore) with outputs scored based on publication venue.

Citation at the individual level has been said to be a useful proxy because it indicates that the work has been read and considered of sufficient importance to refer to it.[146] However, numerous problems have also been identified. Counting citations does not distinguish between positive, negative and neutral mentions.[147] Even then the issue is not clear cut since a substantial number of negative citations might at times reflect the fact the work is controversial and influential.[148] When counting citations, one has to rely on particular databases, and there are limitations of database completeness, particularly for legal scholarship.[149] Citations may be biased by extraneous characteristics such as author gender.[150]

Self-citation and citation 'cartels' are problems, including the tendency of articles in US student-edited law reviews to cite other articles in such venues in preference to peer reviewed articles.[151] Citation practices can vary across subject matter and citation context. For example, in the High Court, citation of commentary on constitutional law has been more extensive than that of commentary on most other

[141] Council of Australian Law Deans, 'CALD Standing Committee on Research and Scholarship: Recommendation 1.6' <https://cald.asn.au/wp-content/uploads/2017/11/CALD-Statement-giving-effect-to-Recommendation-1.pdf>; Carnelley (n 110); Duncan (n 139) 87.
[142] Arup (n 96).
[143] Ibid 45.
[144] Arup (n 96).
[145] Gestel, 'Sense and Non-Sense of a European Ranking of Law Schools and Law Journals' (n 115).
[146] Ian Ramsay and Geofrey P Stapledon, 'A Citation Analysis of Australian Law Journals' (1997) 21(2) *Melbourne University Law Review* 676, 677; Russell Smyth, 'Academic Writing and the Courts: A Quantitative Study of the Influence of Legal and Non-Legal Periodicals in the High Court' (1998) 17(2) *University of Tasmania Law Review* 164, 169.
[147] Ronald H Coase, 'The Problem of Social Cost: The Citations' (1996) 71 *Chicago-Kent Law Review* 809, 810; Jeffrey L Harrison and Amy R Masburn, 'Citations, Justifications, and the Troubled State of Legal Scholarship: An Empirical Study' (2015) 3(1) *Texas A&M Law Review* 45.
[148] Russell Smyth, 'Academic Writing and the Courts: A Quantitative Study of the Influence of Legal and Non-Legal Periodicals in the High Court' (n 146) 170.
[149] Harrison and Masburn (n 147) 74–76.
[150] Marianne A Ferber, 'Citations: Are They an Objective Measure of Scholarly Merit' (1986) 11 *The University of Chicago Press* 381.
[151] Oren Perez et al, 'The Network of Law Reviews: Citation Cartels, Scientific Communities, and Journal Rankings' (2019) 82(2) *Modern Law Review* 240.

areas of the law.[152] However, in the six state Supreme Courts there appeared no correlation between citation count and subject matter.[153] Citation counts can also vary across publication type with Supreme Courts consistently citing legal textbooks more often than journal articles, though this difference is less prevalent in the High Court.[154]

There are feedback loops that have nothing to do with quality, such as where articles with many citations attract more citations (known as the Matthew effect).[155] There are other odd correlates that have nothing to do with quality, such as length of the title and area of law.[156] Despite such limitations, some argue that citation metrics adjusted appropriately at an individual level or averaged over a faculty to assess an institution can be useful.[157]

The problems increase, however, when seeking to rank publication outputs such as journals by reference to citation metrics. As set out above, there are now a variety of statements including DORA and the LM, that likewise critique flawed metrics on a broader disciplinary stage. Scholarship also identifies particular problems for legal journals. Attempts to rank journals based on citation metrics for Australian legal scholarship include analyses by Warren and Ramsay and Stapleton.[158] These tend to favour public lawyers.[159] There are also examples from other jurisdictions, which likewise tend to favour scholars working in particular fields.[160] There are differences in methodologies here, including choices as to how to manage the reality of different numbers of articles and pages per journal, how to compare older and newer journals, and whether to count self-citation. Even assuming these differences can be managed through agreed scaling, there are other problems with ranking law journals including the difficulty of comparing journals from differently-sized jurisdictions,[161] general and specialist journals[162] as well as peer reviewed journals and non-peer reviewed (typically US) journals.

The most significant problem with citation metrics as identified in the literature, both at the individual and journal level, is that it skews scholarship in harmful ways.[163] Australian legal scholars might get more citations by publishing analyses of US law in US law journals, which tend to have better citation scores. This can occur not only, at the journal level, but also at the individual level given the culture of US legal academia with large numbers, heavy citation practices and a tendency to cite articles from US

[152] Russell Smyth, 'Other than "Accepted Sources of Law"?: A Quantitative Study of Secondary Source Citations in the High Court' (1999) 22(1) *UNSW Law Journal* 19, 42. The data was collected at intervals in the years 1960, 1970, 1980, 1990 and 1996.
[153] Russell Smyth, 'What Do Intermediate Appellate Courts Cite? A Quantitative Study of Citation Practice of Australian State Supreme Courts' (1999) 21(1) *Adelaide Law Review* 51. For summary of literature on citation practices of the judiciary, including Smyth's publications see Rachel Klesch, Guzyal Hill, and David Price, 'The Academy and the Courts: Citation Practices' (2023) 42(1) *University of Queensland Law Journal* 103, 112.
[154] Russell Smyth, 'Citing Outside the Law Reports: Citations of Secondary Authorities on the Australian State Supreme Courts over the Twentieth Century' (2009) 18(3) *Griffith Law Review* 692, 712.
[155] Robert K Merton, 'The Matthew Effect' (1968) 159(3810) *Science* 56.
[156] Ian and Ayres and Fredrick E Vars, 'Determinants of Citations to Articles in Elite Law Reviews' (2000) 29(1-Part 2) *Journal of Legal Studies* 427.
[157] Gary M Lucas Jr, 'Measuring Scholarly Impact: A Guide for Law School Administrators and Legal Scholars' (2017) 165 *University of Pennsylvania Law Review Online* 165.
[158] Dennis Warren, 'Australian Law Journals: An Analysis of Citation Patterns' (1996) 27(4) *Australian Academic and Research Libraries* 261; Ramsay and Stapledon (n 146).
[159] Murray and Skead (n 2).
[160] Solomon Bopape, 'Measuring the Scholarly and Judicial Impact of Accredited Legal Journals in South Africa' (2017) 27(1) *African Journal of Library, Archives and Information Science* 53.
[161] Lyria Bennett Moses, 'Legal Scholarship and Q1 Journals' (presented to LADRN at July 2022 meeting).
[162] Gestel, 'Sense and Non-Sense of a European Ranking of Law Schools and Law Journals' (n 115).
[163] UNSW Law Faculty (n 109).

journals. The literature indicates that even such authors would struggle. They would confront the fact that non-US authors have a very low probability of being published in a top fifty US law review; in particular, the average proportion of overseas-affiliated authors on a per journal basis was just 3.1%.[164] Some Australian scholars would also be disadvantaged by US law journals' preference for theory over doctrine and for constitutional law over other areas.[165] But the larger problem is that legal scholars would, collectively, produce less work directed at understanding and improving Australian law and the Australian legal system, which is surely not in the national interest.

(d) - Summary of literature on assessment approaches

Assuming one wishes to do quality assessment, the literature is generally pessimistic as to the existence of valid means by which this might be done. This is not just a law problem - most of the indicators currently used to rank universities have no scientific validity.[166] But it is particularly a law problem given the diversity of views on what constitutes "quality" both within but particularly between different jurisdictions.[167] Thus, any method for quality assessment adopted in Australia, assuming agreement could be obtained, would not necessarily align with the metrics that others would use to evaluate us (as a jurisdiction, as institutions, as scholars, or at the level of outputs).

(e) - Impact

Quality is not the only dimension on which one might seek to evaluate legal scholarship; impact, particularly impact beyond academia, has been rising in importance.[168] As outlined in the Background section, impact has been assessed at a national level in both the UK and Australia. There is, however, significantly less literature on how this might be measured, particularly in the legal domain where one is often looking at impact on policy rather than patents and commercialisation. Impact is work, often requiring outputs to be rewritten in different formats and for different audiences in addition to other forms of engagement, and thus should be explicitly recognised if it is to occur at all.[169]

There are aspects of quality assessment that may discourage high-impact activities. Criticism has been directed towards the ERA for its discouragement of publication of student or practitioner textbooks (which are not recognised as research outputs).[170] This reduces incentives for academics to engage in an important, high-impact role. It must be remembered that legal academics are also members of a profession whose influence extends to the profession of law as well as to society as a whole.[171]

Another context in which quality assessment can run counter to impact goals is reliance on journal lists. In this context, it is interesting to note Smyth's comparison of the journals that were most cited over a long period as 'authority' by judges, with the final ranking of law journals as part of the

[164] Stephen Thomson, 'Letterhead Bias and the Demographics of Elite Journal Publications' (2019) 33(1) *Harvard Journal of Law & Technology* 203.
[165] Dibadj (n 119).
[166] Gingras (n 98).
[167] Dibadj (n 119).
[168] See, eg, Morgan Jones Molly et al, *Assessing Research Impact: An International Review of the Excellence in Innovation for Australia Trial* (RAND Europe, 2013).
[169] McNamara (n 50).
[170] Terry Hutchinson, 'Legal Research in the Fourth Industrial Revolution' (2017) 43(2) *Monash University Law Review* 567, 574.
[171] Susan Barker, 'Exploring the Development of a Standard System of Citation Metrics for Legal Academics' (2018) 43(2) *Canadian Law Library Review* 10, 11.

Excellence in Research for Australia (ERA) exercise.[172] The most cited law journal, the *Australian Law Journal*, was only a C journal on the ERA list. Other journals highly cited by judges, such as the *Criminal Law Journal* and *Criminal Law Review*, were ranked B journals on the ERA list. While there was some overlap between the lists, it is interesting that journals might have a high impact, in the important sense of influencing the development of the law, while being listed as low 'quality' for academic purposes.

This might raise questions about the meaning of 'quality' in the context of legal research, given the prominence of influencing the development of law among its purposes, but it also links with the challenge of high impact activities being discouraged by quality-based evaluation criteria. Further, because judges may rely on academic work without citation and because judicial citations where they do occur) may not be counted, academics are less likely to be funded for judicially relevant work.[173]

None of this is to suggest that judicial citation is the only indicator of impact or that it is easy to measure. The mere fact of citation (without analysing the influence on the outcome and the ratio decidendi) does not imply impact any more than academic citation (alone) implies quality. The reason a judge may cite an academic work are many, for example, convenience where there are useful summaries of the law, to assist in interpretation of earlier cases, or to cite social science evidence to support legislative fact.[174] Courts often do not cite academic work (even when they use it) and, where they do, may be subject to gender and other biases.[175]

The reluctance to cite academic work by judges has been previously rooted in the 'living author' rule, whereby living authors were not cited as an authority.[176] the 'living author' rule has now become outdated and the High Court has been more willing to cite academic work in recent years.[177] However, the six state supreme courts cited far fewer legal journals than the High Court, although there has been an increasing reliance on secondary sources over time.[178] Citations have remained even lower in the NSW District court and territory supreme courts.[179] This likely relates to the role of the High Court in dealing with more difficult questions of law and policy.[180]

There are broader challenges in measuring impact that are highlighted in the literature. There are both different definitions of impact and different types of impact, making comparisons amongst attempts

[172] Russell Smyth, 'Citing Outside the Law Reports: Citations of Secondary Authorities on the Australian State Supreme Courts over the Twentieth Century' (n 154) 720. In the High Court, the Australian Law Journal and Law Quarterly Review were the most heavily cited periodicals with 50 citations each in the measured period. In comparison, the *Sydney University Law Review* was cited 16 times, the *Melbourne University Law Review* 12 times, and the *Monash University Law Review* was cited 3 times.
[173] Rachel Klesch, Guzyal Hill, and David Price (n 153) 108.
[174] Russell Smyth, 'Academic Writing and the Courts: A Quantitative Study of the Influence of Legal and Non-Legal Periodicals in the High Court' (n 146) 167.
[175] Katy Barnett, 'Citation as a Measure of "Impact": Female Legal Academics at a Disadvantage?' (2019) 44(4) *Alternative Law Journal* 267; Susan Kiefel AC, 'The Academy and the Courts: What Do They Mean to Each Other Today?' (2020) 44(1) *Melbourne University Law Review* 447, 452.
[176] Susan Kiefel AC (n 175) 451–2.
[177] Russell Smyth, 'Academic Writing and the Courts: A Quantitative Study of the Influence of Legal and Non-Legal Periodicals in the High Court' (n 146); Susan Kiefel AC (n 175); Rachel Klesch, Guzyal Hill, and David Price (n 153) 104.
[178] Russell Smyth, 'Citing Outside the Law Reports: Citations of Secondary Authorities on the Australian State Supreme Courts over the Twentieth Century' (n 154) 706.
[179] Ibid; Russell Smyth, 'What Do Trial Judges Cite? Evidence from the New South Wales District Court' (2018) 41(1) *UNSW Law Journal*; Rachel Klesch, Guzyal Hill, and David Price (n 153) 123.
[180] Russell Smyth, 'Citing Outside the Law Reports: Citations of Secondary Authorities on the Australian State Supreme Courts over the Twentieth Century' (n 154) 707.

to measure impact difficult.[181] There are issues of contribution and attribution, and accounting for these in a more rigorous way when measuring impact can incur significant transaction costs.[182] While the impact of scientific work is sometimes measured by reference to academic citation, as Sheehy and Dumay have observed, 'law is not a citation-based discipline. In other words, the value and impact of legal scholarship cannot be assessed exclusively or primarily by reference to scholarly citations.'[183] Reehag simply answers 'no' to the question of whether publication and citation counts are reliable indicators of research productivity or impact.[184]

The legal academy has been said to be particularly vulnerable to traditional research metrics, a situation compounded by the lack of citations by the courts.[185] Qualitative methods, such as describing and contextualising impact in case studies, creates other problems including the fact that "being able to tell a good story" might be more important than the impact itself in achieving a high rating.[186] There is also subjectivity in what counts as a positive policy change, and this can be political.[187] Finally, there are issues with timing – policy change can be a slow burn, so is very much a lagging indicator and often has no particular moment to which impact can be attributed. That allows for manipulation for *when* impact is reported.[188]

Nevertheless, an attempt has been made to develop a metric to measure the influence of legal research. Susan Barker describes her development of a metric for legal researchers called the b-index.[189] However, the b-index has not been widely adopted.

Part (V) - Survey results

(a) - Legal Researchers

This section sets out our quantitative analysis of the survey results using the Qualtrics Stats iQ tool and our qualitative analysis of the free-text comments. For the latter, we used NVIVO to code for emerging themes, which were then cross-referenced to institutional characteristics (for example, stand-alone law faculty; GO8) and researcher characteristics (for example, gender; stage of career). Where we found statistically significant and interesting correlations, we detail the relevant statistical test that was

[181] McNamara (n 50).
[182] Molly et al (n 168).
[183] Dumay (n 136) 36.
[184] Sean Rehaag, 'Are Publication and Citation Counts Reliable Indicators of Research Productivity or Impact?' (2018) 43(2) *Canadian Law Library Review* 26.
[185] Rachel Klesch, Guzyal Hill, and David Price (n 153) 108.
[186] Ksenia Sawczak, 'Assessing Impact Assessment – What Can Be Learnt from Australia's Engagement and Impact Assessment', *London School of Economics and Political Science* (2019) <https://blogs.lse.ac.uk/impactofsocialsciences/2019/05/16/assessing-impact-assessment-what-can-be-learnt-from-australias-engagement-and-impact-assessment/>.
[187] Bowrey, 'Audit Culture: Why Law Journals Are Ranked and What Impact This Has on the Discipline of Law Today' (n 94).
[188] Molly et al (n 168).
[189] The b-index includes both quantitative and qualitative reporting of academic, judicial, and social impact. Judicial influence takes up 30 percent of the total, academic influence counts for 50 percent, and social impact at 20 percent**. Judicial impact** is measured by adding and averaging scores in descending order for a citation at each respective court. A score of 5 is assigned to the Supreme Court which descends to a score of 0.2 for a trial court citation. **Academic influence** is measured by totalling the number of citations the article has in academic literature. **Social impact** is measured using open source altimetrics and peer-review of an impact report that highlights a faculty member's activities in the public sphere: Barker (n 171).

used as recommended by Qualtrics given the data being analysed, the p-value and the effect size. The smaller the p-value,[190] the more statistically significant our finding that there is a difference between the two cohorts. The p-value says nothing about the size of the effect. In our discussion of the three statistical tests used in the study—T-test, Chi-Squared Test and Ranked T-Test – it is important to recognize that the interpretation of effect sizes varies by the type of test and the context of the study. Given the distinct nature of each test, we incorporate specific effect size benchmarks (small, medium, large) relevant to each test to enhance our understanding of the practical significance of the findings. Percentages are rounded to the nearest whole number.

(1) - Audience for academic legal research

The question of audience is important because it frames expectations about the appropriateness of different standards of excellence and forms of measurement. For example, if Australian legal scholars are primarily producing work for an international interdisciplinary audience, then standard quality assessment techniques applied to the social sciences might be an acceptable way of measuring quality and academic impact. In the survey, respondents could choose more than one category for their audience. 62% of respondents indicated that they wrote at least some of the time for an international legal scholarly audience, whereas the numbers for other categories were international interdisciplinary audience – 56%, Australian legal scholarly audience – 55%, Australian interdisciplinary audience – 42% and Australian legal professional audience – 31%. Looking across categories, 44% of researchers both viewed their work as relevant to an Australian scholarly legal audience and not relevant to an international legal scholarly audience. This is important because those researchers are not well served by citation metrics given the relative size of the total Australian legal scholarly audience compared to the US.

Respondents were also given the opportunity to explain why they believed their research was relevant to Australian or international legal scholarly, professional or interdisciplinary audiences. There were one hundred responses. Of those who commented that their work is relevant to *both* Australian and international audiences, two main reasons are identified. The first is that issues that there are issues important or of current concern in Australia which are also important in other countries. This makes Australian legal research on these issues relevant internationally. Such issues include, for example, access to housing, racism, climate change, access to justice, mental health law reform, regulation of the digital economy.

The second is the transnational and multidisciplinary nature of many legal issues, which means that even where work is not explicitly comparative, it is often aimed at, or considered relevant to, an international as well as an Australian audience. For example, one respondent states that '[p]rivacy and data protection straddles law, policy, computer science, education so needs to be applicable, or tailored to these audiences'[191] and another states that their work is at 'the intersection of health, science, ethics and law.'[192]

In terms of research on legal education, this is viewed as relevant to both Australian and international legal scholarly audiences. Researchers in (i) the field of international law, (ii) whose work is theoretical, or (iii) whose work examines developments in common law countries or draw on, for example, the disciplines of political science, social or political philosophy, international relations or feminist theory, all comment that their work is relevant to both Australian and international legal scholarly and

[190] Typically, the p-value should be lower than 0.05 before we can conclude the results are statistically significant.
[191] Catherine Renshaw and Lyria Bennett Moses (n 91) Q11-response 94.
[192] Ibid Q11-response 52.

interdisciplinary audiences.[193] A legal theorist comments that even if their work is mainly read by legal theorists in Australia and internationally, it might 'occasionally be read by political theorists, anthropologists or sociologists.'[194]

As noted above, just under a third of respondents claim their work is relevant to an 'Australian legal professional audience.' In this light, several scholars observed 'they have a diverse range of outputs for different audiences.'[195] For example, one respondent notes that they 'write/publish practitioner as well as academic works. So, my work is of broad interest, although much of this is unrecognised by the academy.'[196] Typical responses on this point are: 'I focus on domestic legal systems with a focus on reforming the law. This is relevant to legal academics, the legal profession and government (e.g. policy makers, regulators)'[197] and 'key ideas in my discipline affect both practice and theory, and both local and international contexts. A joy of legal scholarship is this diversity.'[198] However one respondent states that:

> 'Theoretical and interdisciplinary work is only engaged by legal scholars and not the profession. I would only be writing with other scholars in mind and not always with the legal profession. The legal profession holds their own distinct views about the law, and how it is to be used to shape their current work.'[199]

The following comments are representative of the views of many respondents who provided written comments: 'unless you aim to reach professionals little reform will occur;'[200] 'it is important for law development/reform that it be open to the widest audience;'[201] and 'it is part of our work as academics to also conduct research that is useful for the profession, not just the academy.'[202]

(2) - Perceptions about the value of different outputs

The question of how different outputs are valued within the discipline (books, journal articles, book chapters, Non-Traditional Research Outputs (NTROs etc)) is also important because of its implications for the fairness of different evaluative regimes. Book chapters in edited collections, for example, might be highly valued by legal researchers for their potential academic impact in bringing together leading authors writing about the state of the field in a particular sub-discipline of law. Yet such a chapter may be subjected only to review by the editors and not to double-blind peer review, and its citation count may be low because of limited accessibility.

We asked participants to rank which outputs they most valued writing. Participants ranked the following outputs in order from most valued to least:[203]

---

[193] For example, one respondent stated: 'My research is more theoretical based which means that my research can apply cross-jurisdictionally, although is likely most of use to a western audience.': Ibid Q11-response 70.
[194] Ibid Q11-response 46.
[195] Ibid Q11-response 63.
[196] Ibid Q11-response 31.
[197] Ibid Q11-respondent 36.
[198] Ibid Q11-respondent 41.
[199] Ibid Q11-respondent 27.
[200] Ibid Q11-response 2 - early career researcher.
[201] Ibid Q11-response 75.
[202] Ibid Q 11-response 55.
[203] To obtain a rank for the perceived importance of ranking published works, the following method was used. For each publication type, an aggregate score was obtained. The score was calculated by multiplying the number of participants who voted for a specific value and then summing the total score. For example, since 70

1) Journal Article (228);
2) Monographs (311);
3) Chapters in edited collections (417);
4) Non-traditional research outputs (645);
5) Textbooks (699);
6) Case notes (705); and
7) Casebooks (859).

This was similar to the answers given in the survey completed by LADRN members. Non-G08 participants tended to rank textbooks lower than G08 participants (Ranked T test P-Value 0.00572; effect size 0.0407 which is small).

This question generated a diverse range of qualitative responses. Many respondents note that monographs constitute the most sustained form of legal research and that for this reason they should be recognised as the most significant form of scholarly output.[204] Several respondents note that the accessibility of articles makes them highly valuable;[205] and that journal articles go through a critical review processes that most book chapters, for example, do not.[206] One respondent suggests that case notes symbolise the 'over-reliance' of legal scholars on 'secondary research'.[207] In relation to book chapters, one respondent notes the variability in quality and length, commenting that much depends on the collection itself.[208]

Another respondent notes 'I would have like to have seen online op-eds, opinion editorials, and online academic online forums, journals, and academic blogs included in this list.'[209] Whether textbooks should be recognised as research outputs generated several responses: 'Textbooks have a much greater impact on shaping the next generation of thinkers than often pay wall protected academic papers'[210] and: 'I am especially concerned that textbooks are low on that list given what I see as the diminishing quality of legal education.'[211] One respondent states that encyclopaedias deserved more recognition.[212]

Several respondents observe that traditionally 'lower-ranked' publications (NTROs, case notes, textbooks and casebooks) 'are very valuable aspects of legal scholarship and should be recognised as research outputs.'[213] Several respondents note the specific importance of NTROs.[214] For example, one respondent states:

---

participants attributed a value of '1' to Monographs, a score of 70 was given. Repeating this step for each respective value an aggregate score of 311 is obtained for monographs (1x70 + 2x23 + 3x14 + 4x15 + 5x7 +6x5 + 7x4= 311). This was repeated for each publication type. Since a score of 1 was given to the most valued and 7 to the least valued, the lowest aggregate score is the most highly valued output.

[204] Ibid Q12, response 4, response 7, response 8, response 10, response 11, response 15, response 20, response 34, response 37, response 41, response 44, response 53, response 54, response 60, response 78, response 81, response 82, response 83, response 93.

[205] Ibid Q12-response 2, response 24, response 25, response 49, response 51, response 58, response 72.

[206] Ibid Q12-response 39.

[207] Ibid Q12-response 6.

[208] Ibid Q12-response 88.

[209] Ibid Q12-response 95.

[210] Ibid Q12-response 86.

[211] Ibid.

[212] Ibid Q12-response 1.

[213] Ibid Q12-response 14, response 15.

[214] Ibid Q12-response 85.

> ‘I have long been frustrated that the impact of academic contributions to policy and consultations is not recognised. The impact is instant, significant and accessible to far broader audience than that for the other categories.’[215]

Another respondent states that:

> ‘NTRO submission and the collaborations and network that goes into them is the foundation of all of the other forms of scholarly and professional writing - the direct line to impact is what matters with these, and they can and do make that happen. These then drive more scholarly research and other scholarly contributions.’[216]

One respondent observes that: ‘Monographs and journal articles are the bread and butter of an academic, yet in law reform, often it is the submissions to domestic and legal inquiries that yield the most impact.’[217] Several responses suggest that different forms of output are relevant to different audiences, and should perhaps be weighted and recognised differently according to their purpose:

> ‘I aim my work to an interdisciplinary policy, legislative reform audience in general and the most useful outputs are journal articles and non trad outputs - websites, reports policy documents. I think of text books and case books as primarily aimed at an internal law audience who are 'doing law' rather than trying to change it - primarily students and practitioners.’[218]

The following comment is a useful summary of the tenor of many comments:

> ‘These [monographs, NTROs, journal articles, chapters, textbooks] are all important. Textbooks and casebooks make the most immediate impact on how the law affects people in our society. Articles and chapters have a scholarly audience and end up feeding into the structures of the law as perceived by the academy. There are significant times when any of these is critical.’[219]

(3) - Perceptions of independence in setting research agendas

Most participants believe themselves to be relatively independent of their institution in setting a research agenda. While most respond to opportunities where they arise (85% agree), less than 10% agree that they focus on areas because their institution encourages them to, although there is a statistically significant greater proportion agreeing from non-GO8 universities (18.4% compared to 8.0%; Chi-Squared test P-Value 0.0264; effect size 0.183 which is small). A higher proportion (31%) agreed that they focussed on areas of research that might benefit them career-wise (this goes up to 35% if professors are ignored) and 15% agreed that they were guided by what their school or faculty found important. Nevertheless, 97% agreed that they were guided by curiosity or personal interest. Interestingly for the debate about research alignment with teaching, only 41% agreed that they focussed their research on areas aligned with their teaching (21% disagree).

(4) - How researchers choose where to publish

[215] Ibid Q12-response 32, response 43.
[216] Ibid Q12-response 70.
[217] Ibid Q12-response 96.
[218] Ibid Q12-response 33.
[219] Ibid Q12-response 100.

When asked to rank how they choose publishers or journals, on average the most important factor was the participants' own opinion of quality (score of 342).[220] This was followed by 2) reputation for quality (452),[221] 3) receiving a personal invitation (453), 4) inclusion on an institutional list (455), 5) ranking or metrics (464),[222] and 6) an association with a particular conference (690). Reference to an institutional list was ranked higher for non-GO8s compared to G08s (T-Test P-Value <0.0001; effect size 1.06 which is large; 1.55 difference between average ranking). This also aligns with a greater reliance on institutional lists when explaining the quality of research or assessing the quality of others' research in non-G08 settings (Chi-Squared Test P-Value 0.0003; effect size 0.28 which is small), as shown in the table below:

Weight given	Little value	Some value	Great value
GO8	45%	43%	12%
NotGO8	25%	40%	36%

There were a few interesting correlations with gender. Women are more likely than men to respond to invitations as they arise (Chi-Squared Test P-Value 0.0000215; effect size 0.268 which is small) and choose research topics that allowed them to collaborate with people they enjoyed working with (Chi-Squared Test P-value 0.0298; effect size 0.177 which is small).

The survey provided respondents with an opportunity to explain their ranking. This question generated 100 responses, with the majority of respondents expressing discontent with the way decisions about research quality are influenced by their institution's mandates around recognition of output, including journal lists, and the link between institutionally sanctioned output and workload allocations. This response, for example, is typical:

> 'These days I am unfortunately driven by rankings and lists and how the quality of the journal will affect my research allocation in subsequent years. This means I don't publish in journals that I would like to publish in because it is unranked or ranked low'.[223]

One respondent states: 'Publishing in ERA2010 A or A* list guarantees that research workload will be accrued;'[224] another states that: 'Very restricted and prescriptive Workload allocation determines research choice.'[225] In short, many respondents convey that the institutional response to producing outputs in forums that are not approved by institutions (such as journals that are not recognised on the institution's journal list) results in a higher teaching load, regardless of the quality of the work, its impact, or it suitability to the publication where it appears. For example, one respondent states:

> 'I resent the very narrow range that are promoted by the University metric system and would prefer to be able to get workload credit for publishing in the good quality specialist journals that relate to my interest areas. We get no workload credit for any books other than a monograph, nothing for chapters, submissions etc.'[226]

[220] For explanation of the methodology used to obtain the rank, see explanation provided at footnote 203 above.
[221] This was unsurprisingly more highly ranked by professors (including emeriti) – T-Test P-Value <0.00001; effect size 0.573 which is medium; difference between rank score averages 0.847.
[222] Non-professors/emeriti relied on this to a greater extent – T-Test P-Value 0.000415; effect size 0.501 which is medium; differences between rank score averages 0.784.
[223] Catherine Renshaw and Lyria Bennett Moses (n 91) Q23-response 55.
[224] Ibid Q23-response 21.
[225] Ibid Q15-response 58.
[226] Ibid Q15-response 6.

The following responses are from Senior Researchers, at GO8 institutions, who seem more able to determine their own publishing strategy:

> 'Metrics do not matter to me; I want my research to be read by scholars in my field. But I acknowledge that I am senior and continuing, so understand that others do not feel this way.'[227]

> 'This list didn't really speak to me, as a senior scholar with tenure I am less bothered by institutional pressures. I think about accessibility - ie open access, and audience of the journal - who am I trying to reach with this piece? and also turn around - some journals are very slow and I would avoid them.'[228]

> 'I have a clear mission in my career which has stretched back over a decade. I want to achieve a particular change in laws and society and have built my work with this in mind. I develop a strategy on publishing which supports it. I do not care about all the rubbish about offers, or journal rankings, or the like, but what will enable me to achieve my goal. [S]o that does happen to involve in very good publications and books - as that is important. [B]ut the fact I have very high publishing, speak at the UN or do work with our governments and others is only because it is relevant to the plan I have.'[229]

Some senior researchers are open about the fact that seniority gives them a freedom they would not have if they were a junior researcher: 'I think this is reflective of where I am now. However, I'm not sure I would have answered Q15 in the same way, or with the same confidence, at various times during my career.'[230] The following comment comes from a senior researcher at a non-GO8 university:

> 'I have never paid a lot of attention to rankings because they were always terribly flawed in my area. I care that people read what I write. I have a particular preference now for journals that are on Austlii because they are accessible.'[231]

Several responses note that specialist journals, which often were not recognised as 'high impact' on metric measurements and which did not often feature on law school journal lists, had significant impact and relevance in their field.[232]

Many respondents state that they pay attention to who their target audience is and publish in places where their audience is most likely to be reached.[233] Some respondents note: 'I choose to publish in [places] where it is most likely to have the broadest reach to policy-makers internationally.'[234] Two respondents commented that whether a journal was open access influences their journal choice.[235] One respondent mentions that they are guided by a sense of how efficient the review and editing process will be.[236] Another notes the tendency for academics to publish in law journals associated with academics own institutions, which creates an uneven playing field;[237] this was a concern of other

[227] Ibid Q15-response 24.
[228] Ibid Q15, response 33.
[229] Ibid Q15-response 10.
[230] Ibid Q15-response 95.
[231] Ibid Q15-response 15.
[232] Ibid Q15-response 35, response 40, response 65, Q 33-response 55.
[233] Ibid Q15-response 63, response 65, response 73, response 69.
[234] Ibid Q15-response 43.
[235] Ibid Q15-response 57, response 76.
[236] Ibid Q15-response 69.
[237] Ibid.

respondents as well, with one noting that journal lists 'inevitably result in G8 domination and past research has shown that internal faculty staff are preferred over other contributors.'[238]

(5) – Metrics

Based on the responses, there is reason to be concerned about the use of input metrics, such as grant funding, in evaluating research quality. 58% of respondents agreed that time spent on obtaining funding costs more than it brings and 45% agreed that their faculty gives too much weight to grants in hiring and promotion. Only 26% agreed that preparing and submitting funding applications improves the quality of their research.

In terms of assessing quality by measuring output metrics such as journal citation counts, most participants agreed that a subjective, human assessment should prevail over the use of quantitative metrics (70% agree; 21% neither agree nor disagree; 9% disagree). This was also reflected in the fact that only 11% of participants attached 'great value' and 45% 'some value' to metrics such as Q1 factors in assessing legal research. Descriptive factors were of far more importance in assessing quality, such as whether the research is likely to become a primary reference point (63% attached 'great value'; 31% 'some value'), the presence of a clear research question (70% attached 'great value'; 27% 'some value'), originality (81% attached 'great value'; 18% 'some value'), and empirical rigour (70% attached 'great value'; 27% 'some value').

The survey provided respondents with the opportunity to identify any other relevant factors and the importance of these factors. Many respondents emphasize that they 'simply read the article. If it is well-researched and written, I would rate it highly.'[239] Other factors considered relevant to quality assessment by respondents were whether the work has won or been shortlisted for any prizes;[240] whether the journal in which the article is published has a process of double blind peer review;[241] whether the article is published in 'a special issue with the world's leading authors;[242] whether courts cite the article or parts of it;[243] and the impact factor of the research, particularly its social impact.[244]

Most respondents were moved to comment on the question of whether there is or should be a link between a journal's perceived quality (because, for example, of its 'ranking' in a list, its citation metrics or its association with a prestigious university) and an assessment of the quality of an article published in that journal. A majority of those who responded report that they 'seldom judge the journal overall'[245] as a factor relevant to their assessment of the quality of work published in it; and that what matters is 'the actual piece of research itself and how it is written and presented.'[246] For those few respondents who view the journal as a relevant factor in assessing the quality of a research article, considerations about the 'importance of the journal in creating a body of influential work in the field' are relevant.[247] This response from a senior researcher is typical of qualitative responses:

> 'I judge the quality of research published in a journal by reading it and assessing its quality using many of the criteria in question 22. The underlying assumption here seems to be that

[238] Ibid Q20-respondent 3.
[239] Ibid Q23-response 8.
[240] Ibid Q23-respondent 5.
[241] Ibid Q23-response 49.
[242] Ibid Q23-response 63.
[243] Ibid Q23-response 68.
[244] Ibid Q23-response 56.
[245] Ibid Q23-response 9.
[246] Ibid Q23-response 7.
[247] Ibid.

> publications can be assessed for quality without being read. This is really problematic. I may be one of the few academics who reads at least a sample of publications for applicants for jobs, etc. There is really no other way to assess quality. Publication in what is considered a top journal is not a guarantee of quality. Most journals publish quite a wide range of articles in terms of quality for whatever reason. I am not in favour of substituting poor proxies for quality for expert academic judgment of research--why would we want to do this?'[248]

Overall, respondents are opposed to the use of metrics. The use of metrics in assessing legal research is put colourfully by one senior researcher: 'Metrics are merely the smoke behind which lies a bin fire of prejudice and gaming.'[249] Another senior researcher at a GO8 university states: 'Quantitative metrics are notoriously problematic as a measure of research quality (see DORA, etc).[250] Another foresees that the introduction of metrics usually used for social sciences will mean 'the end of well researched black letter law articles that are for an Australian audience as so few people are part of this group.'[251] One respondent makes the point that '[e]mphasising metrics and points is an outdated style of management and leadership.'[252] A senior researcher at a GO8 states 'rankings create perverse incentives and narrow research into a gaming exercise, often encouraged by major publishing corporations. They can direct people away from actual public impact and distort research projects.'[253]

One of the many issues identified with metrics was the lag between article publication and citation scores. This respondent responds to the question about what factors they consider relevant to assessment of research quality:

> 'Don't care about the forum. Fun story. Long ago published in a new journal because it looks like a fun place to publish. Got all manner of shade from the then Research director. Now that journal is a Q1 and my institutions lists that pub as one of my Q1s.'[254]

Several respondents identify the commonly recognised problem with applying metrics to legal research in small jurisdictions like Australia:

> 'It is very difficult for Legal Academics to publish in high ranking journals because no Australian Law Journals are well ranked internationally but, due to the very domestic nature of a lot of legal research (including my own), Australian law journals are often the best of a selection of poor options. This makes it difficult for law academics to compare favourably with their peers in almost every other field of research.'[255]

Regardless of these problems, many institutions appear to have 'KPIs set around publishing in Q1 journals'.[256] The effect of publishing in highly ranked journals is not always negative:

> 'On the one hand I think journal articles should be based on their own merit and that is not what the research impact factor and Q journal rankings do as those are focused on the overall quality of the journal not the particular article - but I have noticed that my interdisciplinary publications in social policy/sociology/political science have attracted quite a lot more

[248] Ibid Q23-response 31.
[249] Ibid Q19-respondent 80.
[250] Ibid Q19-response 35.
[251] Ibid Q19-response 53.
[252] Ibid Q20-response 20.
[253] Ibid Q19-response 33.
[254] Ibid Q23-response 67.
[255] Ibid Q15-response 78.
[256] Ibid Q15-response 92.

> citations quickly than many of my law journal articles. I think it helps to publish in high research impact factor journals - certainly helped me get my Google Scholar cites above 525 and h-index to 13. Not that those statistical increases have been rewarded at my institution.'[257]

The following comment sums up the plea for peer review instead of the application of metrics:

> 'Academics are the experts in the quality of journals. Metrics do not work in law as it is jurisdiction-specific, and for other reasons metrics are problematic and inequitable … It is completely unnecessary and problematic to have lists and rankings of journals replace academic expertise on quality. We judge quality of research outputs all of the time, it's part of our jobs, and we should continue to do so for purposes of appointment, promotion, assessing grant applications, etc.'[258]

(6) - Mechanisms for evaluating quality

Respondents were also asked whether they agreed, disagreed, or neither agreed nor disagreed, to three statements related to the assessment of research by their institution. The statements were as follows:

- A subjective, human assessment of the substance of publications should prevail over the use of quantitative metrics such as article citation scores.
- The introduction of a national ranking of law journals in Australia would be a welcome development.
- Scoring of article quality (holistically or on particular dimensions) could be usefully added to the journal peer review process and made available to the author.

The first of these was supported by a majority of respondents – 71% agreed, 21% neither agreed nor disagreed, and only 9% disagreed.

There were mixed views on the second and third suggestions. On the second – a national ranking of journals - 38% agreed, 23% neither agreed nor disagreed, and 39% disagreed. On the third - additional scoring of quality by peer reviewers - 37% agreed, 36% neither agreed nor disagreed, and 28% disagreed. There were interesting differences though in these results depending on whether the university was or was not in the G08. For non-G08 universities, a statistically significant higher percentage agreed with a national ranking of journals (44%; Chi-Squared P-Value 0.00249; effect size 0.237 which is small) and the suggestion of additional scoring by reviewers (44%; Chi-Squared P-Value 0.00267; effect size 0.235 which is small). There were also interesting differences depending on position, with 45% of those not a professor or emeritus agreeing with a national ranking of journals (Chi-Squared P-Value 0.000642; effect size 0.218 which is small) and 41% of those not a professor or emeritus agreeing with the suggestion of additional scoring by peer reviewers (Chi-Squared P-Value 0.0111; effect size 0.206 which is small). However, even among these cohorts, neither idea attracts majority support.

The qualitative responses, here and elsewhere in the survey, can help explain these results.

One Senior Researcher at a GO8 university states:

[257] Ibid Q19-response 38.
[258] Ibid Q15-response 43.

> ‘the second and third options are ghastly. These approaches tacitly disparage academic judgment by a community of legal scholars. It adopts a kind of "STEM" approach to legal scholarship. What is far more important with legal scholarship is what happens to a paper afterwards - once it is "out there". Does it provoke debate and discussion? What do other scholars do with it? and so on. Peer review can be useful (reviewers can be amazingly generous with suggestions for improvement, and advance new ideas and insights), but the ultimate value of a piece should not be associate with where it was published (etc): that is to judge a piece ex ante. What is far, far more important is the ex post evaluation by scholarly peers. These kinds of questions buy into an approach to the value of scholarship that should be resisted at all cost.’[259]

Overall, this comment encapsulates many of the comments:

‘The 1st and 3rd suggestions in Q 19 sound lovely but I think they would be very difficult and expensive to put into practice. The 2nd suggestion in Q 19 is problematic for the same reason the old ERA journal rankings were. They are challenging to create and become self-fulfilling prophesies as submission/rejection rates increase in high ranked journals and decrease in low ranked journals due to very basic rules of supply and demand.’[260]Many of those who objected to the third option used their comments to emphasise their view or explain their critique (interestingly, only one person who endorsed the idea commented). Comments include that this is ‘a terrible idea’[261] that would make academic life ‘more onerous’[262] and ‘would not add anything to the referee reports’[263] and there were ‘concerns about the third point given how stretched reviewers are and their qualification and capacity to make consistent decisions in this respect.’[264]

Several respondents took the opportunity to comment (here or elsewhere in the survey) on the issue of journal lists, which is related to the second suggestion, but not limited to it.[265] At one end of the scale there were responses such as ‘Lists retained by institutions are the scourge of the discipline (and higher ed generally).’[266] At the other end of the scale, particularly from early career researchers seeking guidance about where to place their publications. For example, one early career researcher advocates for the adoption of ‘a ranking similar to the Australian Business Law Deans Journal Quality List. The absence of such a ranking means that decision-makers frequently just rely on their subjective opinions about the quality of what they read.’[267] Another respondent at a GO8 university, who aspires to a research career but does not yet have one, states: ‘I find it difficult to know what are 'good' or 'bad' journals and where I should direct my papers for publication.’[268] From the same GO8 university, a mid-career researcher states: ‘I think a national ranking of law journals would be very valuable, but should be given relatively low importance in research assessment. It is more useful as a reputational survey.’[269] Another respondent from a non-GO8 university states: ‘we need to try to minimise subjectivity; an

[259] Ibid Q20-response 22.
[260] Ibid Q20-response 68.
[261] Ibid Q19-response 56.
[262] Ibid Q19-response 54.
[263] Ibid Q19-response 59.
[264] Ibid Q19-response 60.
[265] See discussion at ‘
’ above.
[266] Ibid Q15-response 22.
[267] Ibid Q35-response 156.
[268] Ibid Q15-response 9.
[269] Ibid Q 15-response 11.

objective journal quality list can be of assistance in this regard.'[270] An early career researcher at a GO8 university states: 'Old ERA ranks (A* etc) still influence me'.[271]

There are many more comments, however, that are opposed to the use of journal lists. The following comments are representative of many responses: 'Journal lists are nonsense and do not measure true research impact, especially in law. Lawyers value quality ideas, regardless of where it is published';[272] and 'My institution is moving to more qualitative forms of assessment for research quality which is welcome. Please let's not waste any more time on national quality journal lists. They are utterly counterproductive and discourage critical reflexivity about journal article choice'.[273] One respondent comments 'it is extremely difficult to produce a robust and reliable ranking of law journals (as evidenced by the shortcomings of previous attempts)'[274] and another states 'there is no direct relationship between the quality of the article and the ranking of the journal'.[275]

Many respondents shared the negative impact of journal lists on their research and career:

> 'I feel enslaved to my universities ranking list even though I disagree with it. This is also because our workload is tied to research output. Without demonstrating outputs reaching min targets (measured against internal ranking list), our teaching allocation can increase and you can suddenly end up teaching new or more units. This can trap you into not being able to do research and you end up with an 80% teaching load for years until you can publish your way out of it. This would not be bad if you love teaching BUT the university promotes on research. Another distortion in the industry.'[276]

The problem with lists is felt acutely by some junior researchers:

> 'It is disappointing when our institutions (university or outside groups) put pressure on us to publish according to lists/ranking created by outside institutions (often commercial organisation) about where to publish and how they rank "A" journals to top journals with completely fail to align with what experts in the field consider to be top journals.'[277]

Senior researchers, particularly at G08 universities, appear less susceptible to the imposition of lists:

> 'Did not understand the question quite frankly. I am for publishing where the research will reach those who I seek to impact. I ignore all the crap and get to the purpose of the research - to make a difference. so if a professional mag will hit the target- I'd use it. if it was a policy doc given to gov directly- I'd use it. all the rest is just fluff and BS. I even had a professor tell me to use the old ERA journal list with the A* etc which we threw out a decade ago. that professor publishes good stuff, but I have directly impacted on the law as and the last fed budget. for me impact is more important than a silly rejected journal list. I do find it really strange that the ERA list was held up by that professor.'[278]

[270] Ibid Q20-response 6.
[271] Ibid Q15-response 97.
[272] Ibid Q23-response 37.
[273] Ibid Q19-response 79.
[274] Ibid Q19-response 35.
[275] Ibid Q19-response 87.
[276] Ibid Q23-response 35 (response from non-GO8 university).
[277] Ibid Q15-response 81.
[278] Ibid Q12-response 10.

In the context of *existing* journal lists, there were several comments about the respective weighting of domestic and internationally oriented research, with diverging perspectives: 'the current ranking is too focused on publishing in Australia. This is a globalised world, particularly in the digital economy, and other jurisdictions need to be accounted for in this process a lot more'[279] compared with 'perceptions of domestic v international publishers appear to disadvantage legal scholars focused principally on Australian law.'[280]

Several respondents cautiously welcome the idea of a national list, saying that 'it can be useful for demonstrating that you're publishing in high impact journals, as it's the kind of thing that grant assessors, etc, look for.'[281] One respondent is 'very much a sceptic of institutional assessment of research beyond more or less obvious criteria (e.g., published in dubious outputs or on dubious websites etc.)' but nevertheless thinks that 'law journal ranking in AU could be useful nonetheless if done by CALD because everything that is out there just does not serve us at all in law.'[282] A mid-career researcher at a regional university states: 'I strongly dislike journal rankings however a national ranking system might mitigate against the arbitrariness of my own institution and the whim of management in retrospectively editing the list as they see fit.'[283] A senior researcher at a different regional university agrees:

> 'While a national law ranking is not perfect, it would help us "know" and advise staff. Currently, we seem to tell our universities that law is "different" but don't present an alternative option to scopus, weighted citations etc. The Business Deans list seems to help business academics - whether you agree with the list or not.'[284]

One respondent states:

> 'law journal rankings would only be useful if university promotion panels pay attention to them. I doubt that they will. The lists then become burdensome. The research quality paradigm belongs in another age, utterly ignorant of the contemporary means of dissemination of knowledge. I think it should be scrapped altogether.'[285]

Another respondent states:

> 'I think rankings are pretty much inevitable. There are proxy rankings that used even if there is no formal ranking. A formal ranking (made through an open process) would probably clarify things, especially for early career researchers.'[286]

This final comment sums up the attitude of many respondents:

> 'I think quality assessment is inherently challenging. Ideally rankings would not come into the picture, but many institutions (such as my own) already have journal quality lists and they are somewhat random.'[287]

---

[279] Ibid Q20-response 1.
[280] Ibid Q20-respondense 4.
[281] Ibid Q20-response 58.
[282] Ibid Q19-response 69.
[283] Ibid Q19-response 75.
[284] Ibid Q19-response 77.
[285] Ibid Q20-response 19.
[286] Ibid Q19-response 62.
[287] Ibid Q19-response 84.

In a separate question, we asked about journal-specific factors for assessing research. Most of the factors we proposed received mixed results, with the exception of reputation among peers in the relevant research area, where 74% agreed this was important (a statistically significantly higher proportion for G08 respondents: 89% v 66%, Chi-Squared Test P-Value 0.001; effect size 0.255 which is small). A similar percentage agreed with the same factor being relevant for assessing monographs from a particular publisher, although there was no statistically significant difference for G08 v non-G08 in the case of monograph publishers.

(7) - Impact

There are some important differences between what respondents believed constituted impact and what kinds of impact are recognised by their institution. For example, 92% agreed that changes in people's understandings or beliefs constituted impact, whereas only 33% stated that their institution recognised this as impact (this was higher for respondents from G08 institutions; 47% v 26% of participants' reports on own institution). A majority (59%) felt that inclusion in course reading lists was impact (higher for respondents from G08 institutions; Chi-Squared Test P-Value 0.0180, effect size 0.2 which is small), but only 30% felt this was recognised as such by their institution (higher for G08; 50% v 21% of participants' reports on own institution).

There are also areas where participants and institutions seemed more in alignment, for example citation in a policy report or court judgment, citation by other academics, and research that results in change. Within this, a minority of individuals (and reportedly institutions) felt that 'change' was only 'impact' where it was positive. The role of legal textbooks and treatises as both research outputs and impacts was recognised by a majority of respondents.

Respondents addressed this question along two dimensions in their qualitative responses. One set of responses was directed towards the 'scholarly' impact of research (evidenced by, for example, citations)[288] and another set of responses was directed towards 'real world impact' (evidenced by, for example, 'the extent to which the research positively contributes to the resolution or reduction of legal problems,'[289] the prevention of legal proposals or reforms;[290] the 'development of the law',[291] or 'policy and legal change',[292] demonstrated through 'qualitative case studies'[293] or 'impact narratives'[294]).
One respondent makes the point about objectives: 'Define the objective of the research and then test if that objective was met. Different forms of research have different objectives.'[295] Some responses note that 'media or community engagement' is evidence that research has reached its target audience.[296] Another respondent notes that:

> 'As a professional discipline we need to accommodate diverse forms of 'impact' on a spectrum including scholarly, practice, community engagement, policy, law reform (statute/case) etc. The role of texts in a discipline founded on the 'textbook tradition' needs to be acknowledged

[288] Ibid Q33-response 13.
[289] Ibid Q 33-response 5, response 92.
[290] Ibid Q 33-response 47.
[291] Ibid Q 33-response 43.
[292] Ibid Q 33-response 19.
[293] Ibid Q 33-response 16, response 23.
[294] Ibid Q33-reponse 30, response 49, response 73.
[295] Ibid Q33-response 94.
[296] Ibid Q 33-response 10, response 45.

> as legitimate impact. We need to be forward looking, accommodating scholarly blogs, open access works and law reform submissions as the very fabric of legal scholarship.'[297]

Several respondents note the time lag between producing research and seeing its effect.[298] A broad approach was:

> 'The core question should be to what extent does the research contribute to discourse and debate in the community as a whole or a part of that community? The ways in which this could be measure can be left open. This encourages rather than constrains research for the purposes of impact.'[299]

Another respondent invoked the Sustainable Development Goals:

> 'Evidence-based and clearly referenced short narrative for individuals. Comparing universities might be best done through an existing framework like the SDG international league tables - too much extra work otherwise on very contested issues like indicators, metrics etc - SDGs are internationally agreed and based on decades of negotiations.'[300]

One respondent states:

> 'the impact is often made manifest through being recognised as a leading expert by others; through citations and invitations to collaborate or present with academics and/or legal professional institutions.'[301]

Another respondent notes:

> 'Comparing is ridiculous across different fields and methods. You are not comparing like with like. Subject matter specialists are needed to assess carefully. A general assessment may suffice though for institutional purposes.'[302]

There were several calls for assessment of impact to acknowledge the challenges faced by producing impactful research in non-GO8 institutions:

> 'I think recognising a broad range of research impact is important and this will vary from person to person, some people have success with ARC grants, others are media stars, others hit the top ranked law journals - and some do all of it (amazingly!) - all of it should be valued. I think there needs to be some recognition of research impact opportunity between institutions - as anything not G08 often has far worse working conditions in terms of teaching loads and service loads. I have worked within both G08 and non-G08 environments, and G08's are extreme privilege land compared to the other places I've been employed. Every other week in G08 land there was some elite personage dropping in to dispense helpful wisdom, elite contacts, 'on trend' framing devices for the top journals etc.'[303]

[297] Ibid Q 33-response 26.
[298] Ibid Q 33-response 35, response 40.
[299] Ibid Q 33-response 60.
[300] Ibid Q 33-response 65.
[301] Ibid Q 33-response 71.
[302] Ibid Q 33-response 96.
[303] Ibid Q 33-response 51.

Another respondent raises the possibility of using a national criterion-referenced assessment rubric may work: 'A national approach sounds great as long as it is fair to all law schools, especially smaller/regional law schools that may have higher teaching workloads than other law schools.'[304]

(9) - Peer Review

Throughout all responses, respondents raise the issue of peer review. Most agree that peer review is an important quality indicator. 84% of respondents agreed that scholarly legal journals should always have peer review using external referees. While this may seem standard, it is worth noting that most US scholarly legal journals would fail this test, including ones with good metrics and rankings. In terms of the manner of review, 71% agreed that double blind peer review should be preferred while only 12% agreed that triple blind review should be preferred. Peer review was considered important by fewer but still a majority of respondents beyond journal articles – 65% agreed it was important for scholarly books and 59% for scholarly book chapters.

However, numerous problems are identified. One is that peer review is not transparent.[305] In one respondent's example, an article was published after a negative review that also identified plagiarism.[306] Another respondent comments: 'I have very low confidence in Australian peer review. It is a closed shop.'[307]

One respondent states:

> 'Peer review is deeply problematic, particularly the more senior you are and the more expertise you have in an area. The simple reality is that anyone who reviews my work is very unlikely to have the same level as expertise as me - not because I am a genius but simply because I am senior and work in an area that almost no other legal academics research, either in Australia or overseas. I have wasted so much time responding to reviewers who do not have the humility or insight to realise they are not sufficiently qualified to be reviewing the article. There are also many reviewers (in my observation as an author and editor) who, as one of my colleagues used to say, review by saying 'it is not the article I would write'. I think it would be good to have reviewing guidelines that the legal academy agrees on in some way.'[308]

Another respondent states:

> 'Any 'quality' assessment is riddled with problems. There are clearly low quality works. I have been a peer reviewer for some that I have rejected, and then found they were published without revision anyway. And I have had some reviewers who have clearly not been qualified to review my work, though most are generally good and constructive. Excellent journals sometimes publish poor work and less well known journals publish excellent work. Bibliometrics are distorted by jurisdiction and self-citation (or citation groups). The only defensible way to assess quality is via peer review, but this carries the danger of differences of approach being misunderstood.'[309]

Some respondents linked the poor quality of peer review to the problems with rankings and qualitative metrics:

---

[304] Ibid Q 33-response 81.
[305] Ibid Q20-response 21.
[306] Ibid Q20-response 13.
[307] Ibid Q20-response 86.
[308] Ibid Q20-response 12.
[309] Ibid Q19-response 34.

> 'ranking is pernicious and should be avoided. It's very discouraging for niche or smaller areas of endeavour. qualitative metrics are very difficult in law, particularly in Australia where not all local journals use DOIs. given the problems with peer reviewing (frankly its random and often quite badly done, thanks I suspect in part to the problem of getting peer reviewers) adding a score does not seem helpful at all.'[310]

One respondent states:

> 'As both an editor and author I regard peer review as imperfect but necessary. It can be problematic because many reviewers seem to 'block' articles that could be published with revisions by incorporating referee feedback. Referees often want an article to be what they would write rather than assessing the article based on whether (1) it fits the journal's objectives, (2) adds to a body of knowledge or critiques it, and (3) is internally coherent and supported by evidence. It is also difficult recruiting referees because insufficient weight (or no weight) is given to this crucial aspect of research within performance management.'[311]

(b) - Institutional quality and impact assessment

Survey 2 was addressed to Associate Deans of Research (or 'Research Leaders'). The aim was to obtain information about how research was assessed at an institutional level. The survey showed some commonalities, but also differences, in the criteria against which research is judged across institutions. This is shown in the table below.

Table 2: Responses on How Research is Assessed in Law Schools

Question When assessing the quality of legal research, my school or faculty attaches the following value to:	**Great Value**	**Some Value**	**Little Value**	**Blank**
Whether the research is likely to become a primary reference point for future research in the field	35% (9)	46% (12)	12% (3)	8% (2)
The clarity and elegance of the expression	12% (3)	46% (12)	31% (8)	12% (3)
The articulation of a clear research methodology	23% (6)	12% (3)	58% (15)	8% (2)
Whether the research contributes to the development or building of theory	19% (5)	54% (14)	15% (4)	12% (3)
The presence of a clear research question	42% (11)	46% (12)	4% (1)	8% (2)
The depth of referencing	19% (5)	65% (17)	4% (1)	12% (3)
The precision of referencing	19% (5)	54% (14)	19% (5)	8% (2)
Originality	62% (16)	31% (8)	0% (0)	8% (2)

[310] Ibid Q19-response 42.
[311] Ibid Q19-response 67.

Logical consistency, coherence and persuasiveness of the argument	58% (15)	31% (8)	4% (1)	8% (2)
Empirical rigour (if relevant) of the material in the publication	44% (12)	38% (10)	8% (2)	8% (2)

(1) - Nationwide Assessment of Research

Participants were asked to express the extent of their agreement with propositions about nationwide assessments of research. The responses indicate a general disinclination towards a nationwide assessment of research. Only 6 respondents (23%) agreed that a nationwide research assessment exercise, for example the ERA, is a suitable away to compare the quality of research groups. 46% (12) of respondents considered that research assessment exercises often end up being about the informal reputations of researchers and research groups than about the actual quality of their publications and 42% (11) participants neither agreed nor disagreed with this statement. However, the majority of respondents (17/65%) agreed that a nationwide research assessment exercise, if conducted, should involve a subjective, human assessment of the substance of selected publications of each research group.

(2) – Value of Different Research Outputs

We asked participants to rank the value their school attaches to different publication types. Ranking the aggregate scores for each publication type, participants value (in order from most to least valued): [312]

1) Contributions in Journals (37);
2) Books (43);
3) Chapters in books (79);
4) Edited Collections (132);
5) Textbooks (138)
6) Handbooks (153);
7) Case Notes (176) and
8) Entries in reference works like encyclopaedias (170).

A few participants offered written comments regarding different types of outputs. One response considered that books needed to be better recognised, as they provide the basis for greater reputation to the university and researcher.[313] Another considered that more recognition should be given to edited work which not only is an intellectual contribution to the scholarship, but also provides unique opportunity to develop research collaboration that may lead to greater impact.[314]

It would appear there is an increasing trend across institutions to encourage staff to publish in highly ranked journals. We asked participants to rank different approaches of assessing the quality of journals. Of the four factors we gave them, in aggregate, participants ranked them in the following order: [315]

[312] For explanation of the methodology used to obtain the rank, see footnote 203 above.
[313] Catherine Renshaw and Lyria Bennett-Moses, 'Research Leader Questionnaire' Q11-response 28.
[314] Ibid Q11-response 18.
[315] For explanation methodology used to obtain the rank, see footnote 203 above. However, in this case, the highest aggregate score indicates the most highly ranked factor. This is because a score of 4 indicated the highest rank and a score of 1 indicated the lowest rank.

1) The expertise of the editorial board (95);
2) How the journal is ranked by my institution's journal list (68);
3) The reputation the journal has according to quantitative indicators (45) and
4) The reputation the journal has among peers in my field of research (47).

There were no observable correlations with whether participants worked in GO8 or non-GO8 institutions. A number of participants who offered written comments on this issue noted that publication choice is constrained by journal rankings. Participants suggested varying levels of pressure on academics to publish in highly ranked journals. On the less restrictive end, one participant noted "staff are free to pursue their research interests in whatever way they like although there are clearly coming pressures in terms of the increased use and reliance on metrics to judge research performance."[316] On the other end, responses noted institutions "discourage[d] publication in low-quality journals and incentivise[d] publication with higher-quality journals and book publishers"[317] or that publication choice was "clearly limited" by journal rankings.[318] Two participants noted an increasing attitude towards encouraging staff to publish in Q1 (top quartile) Scimago ranked journals.[319] However, one response suggested the Q1 ranking is not relevant or highly valued for Australian related legal research.[320]

(3) – Attitudes Towards ARC Assessment

We asked participants a number of questions in respect to their attitude towards the Australian Research Council's ("ARC") assessment of research. Responses reflected some dissatisfaction with such assessment. 11 respondents (42%) did not consider that ARC panels performed well in evaluating the quality of research proposals in law while 7 respondents (27%) neither agreed nor disagreed on this question. Only 4 participants (15%) agreed that ARC assessors pay sufficient attention to the special characteristics of legal research. Similarly, 5 participants (19%) believed ARC Panels performed well in evaluating the quality of research proposals in law. A majority (20/77%) agreed that ARC grant and fellowship procedures are "too time consuming and bureaucratic".

There appears to be general dissatisfaction towards ARC assessments of legal research, and the reasons for this are expanded on in written comments. A general theme from the responses was that ARC assessments are not well equipped to deal specifically with legal research. Moreover, the variety in the types of legal research was suggested to be not well understood. For example, one participant noted it was very difficult to get grants primarily for doctrinal research.[321] Others suggested not as many law applications are approved as "we might expect and hope" as in other disciplines[322] and while others remarked not enough funding appears to be provided for law and humanities research.[323] One participant indicated the problem goes beyond merely the ARC and stems from Commonwealth policy generally, which drives funding towards STEM research areas, particularly engineering.[324] A preferencing of GO8 institutions was also mentioned by one participant, particularly given that ROPE considerations do not factor in different amounts of time available for research outside the GO8.[325]

[316] Catherine Renshaw and Lyria Bennett-Moses (n 311) Q5-response 14.
[317] Ibid Q5-response 28.
[318] Ibid Q5-response 30.
[319] Ibid Q5-response 12, response 15.
[320] Ibid Q5-response 10.
[321] Catherine Renshaw and Lyria Bennett-Moses (n 311) Q4-response 9.
[322] Ibid Q4-response 17.
[323] Ibid Q4-response 12.
[324] Ibid Q4-response 15.
[325] Ibid Q4-response 30.

Two responses expressed concern that ARC assessment procedures are opaque with insufficient information on how assessments are carried out.[326] Less critical responses still noted that the quality and usefulness of feedback given by the ARC could be improved,[327] which perhaps reflects the need for a deeper understanding of the nature of legal research by institutional bodies like the ARC.

It is worth noting that one respondent in the Researchers Survey, although not specifically asked, also expressed significant dissatisfaction with the ERA and EIA and suggested the exercises only ended up showing that institutions with the most resources performed the best.[328]

Part (VI) - Discussion

When considering how we might understand and measure both the quality and impact of legal scholarship, purpose should be at the centre. The first relevant purpose relates to legal scholarship itself. As can be seen in both the meta-literature on legal scholarship and the survey results, legal scholarship serves a variety of purposes. Law reform is an important purpose, but less than half of survey respondents identified with it, so it doesn't fully capture the goals of legal research. It sits alongside enhancing understanding of doctrine, development of theory, and interdisciplinary work that might serve a range of purposes.

This diversity of purposes is a strength of the discipline of law, allowing researchers to address a range of research questions that are important to different audiences. That is not to say legal research should adopt an "anything goes" approach – any particular research project will need to be justified by reference to originality, methodological alignment, and significance – but it would not be productive to judge legal scholarship by reference to the extent to which it furthers a single purpose. This point is recognised in the CALD Statement's reference to the diversity of legal research. Another point to note here is that where a legal scholar is working within the frame of another discipline, for example, sociology, the quality of the output can and should be judged by the standards of that discipline. This is also recognised in the CALD Statement, although that Statement could be changed to clearly list this as an exception to the necessity of a law-specific panel in assessing such work.

The second kind of purpose that must be borne in mind is the purpose of the assessment itself. There are a variety of circumstances in which it might be important to know something about the quality or impact of a particular output or a body of work of a scholar, or it might be important to understand the performance of an institution, such as a university or sub-unit. The former is important in situations such as hiring, promotion, grant funding, or the award of a prize. These are often particularly important for early career academics seeking to establish themselves as scholars.

One goal of *transparency* in evaluation criteria is better meeting the needs of this group to understand what they should be striving for and avoiding the stress associated with conflicting advice. The latter is done for ranking institutions, as in the earlier ERA scoring process or the various ranking schemes that score research and/or impact. In the former case, evaluation of quality can be done by engaging directly with the work, ideally through the lens of clear, transparent criteria. Evaluation of impact can be done on a basis of engaging with a narrative around the work, understanding the changes in understanding, practice or policy that the work has inspired. This can extend to the latter through sampling, either random or selected, but where the volume is great, it can become onerous and there

[326] Ibid Q4-response 14, response 28.
[327] Ibid Q4- response 18.
[328] Catherine Renshaw and Lyria Bennett Moses (n 91) Q33-response 88.

is a temptation at least for quality assessment to turn to proxies such as citation counts, journal citation metrics, or quality journal lists.

The two kinds of purposes – (1) purposes of scholarship and (2) purposes of assessment – are not necessarily aligned. This is highlighted in the ACOLA report. Assessment can be generally demotivating, can discourage more "risky" innovative projects, can be "gamed" by those individuals and institutions with the resources to do so, and can create particular challenges for disadvantaged groups. Despite a desire within some institutions to measure and rank academic performance, the market-driven desire to score well in university rankings, and the importance of being accountable for the work we do, there can be a significant cost to our primary goals if we pursue measurement without recognising its costs. Turning now to mechanisms for assessing quality and impact, there are five approaches for quality assessment and two approaches for impact assessment that we consider below.

(a)- Quality assessment

(1) - Input-based controls

Input-based controls, such as peer review, set a threshold for quality of published scholarship. If something has been published following a rigorous double-blind peer review process, then at least some number of scholars have engaged with the work and judged it as meeting the quality threshold for publication. This, however, represents a kind of ideal. As the qualitative analysis demonstrates, many legal scholars have experience with articles being published after a negative review and of receiving reviews of mixed quality. The US approach is also distinct and, while student-editors of US law reviews often seek opinions from faculty when assessing articles, the reviews are rarely blind and the process varies widely.

The topic of peer review was discussed extensively at the LADRN meeting of 8 February. The views expressed paralleled our findings on the uneven quality of peer review, including the fact that some peer reviewers make inappropriate or discourteous comments. Explanations included the difficulty of finding peer reviewers, the lack of diversity of peer reviewers, the lack of workload – or any – recognition for peer review, and reliance on junior scholars for peer review (often without sufficient guidance and mentoring being available at their home institution). The importance of transparency was also emphasised, and beyond the contexts we had considered in our draft report. For example, one LADRN member mentioned that some editors do not advise authors as to the recommendation (to publish or not) made by peer reviewers. Terms such as "revise and resubmit" can have different meanings – from having changes reviewed by the same or replacement reviewers to having the entire submission evaluated from scratch by a new reviewer, and this is not always clear to a submitting author or the reviewer who makes the recommendation. It is not always clear whether peer reviewers are necessarily internal or external to a journal's editorial board or how reviewers of senior and junior scholars are weighed.

One way forward for input-based controls would be for CALD to articulate standards for peer review. These standards could specify that peer review should be double-blind, that each paper is reviewed positively a minimum number of times before acceptance (say, 2, for example), that peer reviewers should be selected based on their familiarity with the subject matter and their history of high-quality reviews, that the criteria for peer review be transparent (to both reviewers *and* submitting authors) and possibly even that journals formally acknowledge peer reviewing excellence to encourage polite, constructive and engaged reviews (and edit out inappropriate content). Journals might also be encouraged to participate in programs whereby reviewers can authenticate completion of a number of reviews while remaining anonymous – this provides a mechanism whereby the important work of peer reviewers can be recognised as part of an academic workload. There are other innovative

reviewing practices that were mentioned at the LADRN meeting, including 'open' peer review where peer reviews are published alongside papers as a comment and circulation of reviews among reviewers with opportunity for agreed feedback. As some LADRN members pointed out, there is a balance here – being too prescriptive might reduce positive innovation and could lead to journals, especially international journals, ignoring CALD's standards entirely.

Process guidelines could be supplement with guidance for reviewers, including relevant criteria for assessment (see 'criteria-based assessment' below). While CALD's influence is likely strongest for Australian, as opposed to foreign or international journals, this would allow journals to opt-in to a national standard of quality peer-review and thus allow those who publish in those journals to identify their outputs with those quality standards. This creates a quality floor for published work depending on the extent of adoption: it does not grade or rank publishable legal scholarship or provide any information on its ultimate impact.

(2) - Criteria-based assessment

The criteria used by journals, peer reviewers, referees, grant assessors, hiring committees, promotion committees and other assessors of quality and impact is not uniform and is often not even transparent. Part of the challenge here is both (1) the diversity of purposes of legal scholarship so that some criteria might be more relevant to some purposes than to others, and (2) the fact that some assessors may be outside the discipline entirely and thus unable to judge work directly against discipline-relevant criteria and thus need to rely on proxies. However, CALD may wish to consider a list of "core" criteria and "where relevant" or "optional" criteria against which the quality of legal research might be judged by those who are positioned to do the evaluation. In the case of hiring and promotion committees, which often include those outside the field and discipline, this could be used by those providing references to frame their assessment of a candidate's research.

If CALD wished to consider this option, further work would be required to formulate the relevant criteria. The literature review above suggests some useful starting points as do the existing Australian Legal Research Awards Winners (**ALRA**) criteria. This would also provide a useful tool to assist new legal scholars in understanding the terms on which their work will be judged, and in a way that cuts across institutions thus potentially enhancing mobility.

Criteria could also be used, is already being done within ALRA, for the award of prizes for excellence in legal scholarship. Journals could be encouraged to institute annual prizes, although the feasibility of this might depend on the make-up of editorial boards. For example, some journal editorial boards are made up of students who would necessarily rely on peer review reports across which it is often difficult to compare.

(3) - Aggregated quality scores

One challenge with criteria-based assessment is that one opinion is necessarily subjective. Awards such as ALRA rely on multiple assessors on each committee for this reason. This raises the question of whether one could reduce the impact of any individual assessors' bias through aggregating the views of larger numbers of assessors.

There is one way this may be done. Suppose that peer reviewers, after deciding that a piece was publishable (with or without minor revisions) also gave that piece a score, say ranging from 5 (publishable standard) to 10 (excellent). These scores would be captured in a centralised system, with data entered by journal editors (with whom the scores would be shared at the time of review). These scores would need to be scaled, so that all reviewers' scores (once a statistically useful threshold was reached) had the same mean and standard deviation (thus reducing differences between 'harsh' and 'easy' scorers). Once a statistically meaningful number of scores for a particular academic had been accumulated, they could be advised of their 'average' score, say once each year.

Such a scheme would rely on several assumptions and have critical limitations. It would rely on the quality of peer reviewers, which is currently mixed, and thus may only be applicable to journals that opt-in to a CALD peer review standard. It would entail significant costs and likely specialist software with complex authentication systems that gave different people access to particular data access and entry rights. It would be limited to Australian law journals because CALD's influence is so limited and, even there, only to those journals opting in to use the system. And without significant participation across Australian law journals, the scores would become less useful or meaningful even for those researchers whose research is mainly in such journals. Finally, it was unpopular when the suggestion was raised in our survey. We thus do not recommend this as a way forward. However, it usefully illustrates the difficulty of constructing a meaningful quality score for an individual researcher.

(4) - Journal lists

The use of journal lists was similarly unpopular in our survey of Australian legal researchers and many wrote that they do not judge an article by reference the journal in which it is published. As some survey participants commented, journal rankings are an attempt to judge the quality of legal scholarship without reading it. They are thus suspect and are often a poor proxy measure. The construction of journal lists is not an objective, neutral process, particularly in the context of legal scholarship. This is due to a number of factors including the lack of a historically recognised "top" journal or journals in law; the existence of journals with particular foci and purposes (so that, in previous lists, some subfields had no choice of top journal); the inevitability of conflicts of interest for those constructing the list who might favour journals associated with their own institution or sub-field; and the changing pool and quality of journals over time.

There were also examples from the survey of academics making publishing choices as a result of the use of journal lists that did not align with the best audience for their research. This is an example of a secondary purpose (assessment) undermining the primary purpose (the good that research can do). The effect is even more severe for early career academics, particularly at non-GO8 institutions, where narrow institutional publishing requirements can combine with workload consequences for researchers who fall outside those prescriptions.

We would thus not recommend going back to the attempt to formulate a national journal quality list. However, we note that many institutions continue to use their own journal lists or rely on inappropriate lists such as the old ERA list or the Business Deans list. This is in many ways even worse than a national list. It reduces mobility as researchers will be judged by different lists within different institutions, biases are even more acute given lists would be constructed and adjusted over time by a smaller group with the ultimate decision likely by one senior academic (such as a Dean, Head of School or Associate Dean of Research), and it creates significant and unfair pressures on those subject to them, particularly early career academics at non-GO8 institutions.

In that sense, one advantage of a national list would be that it is a less-worse problem than the current proliferation of institutional lists, particularly at non-GO8 institutions. It would also offer some benefits, including advice to early career researchers who are unsure where to place their articles. Unlike institutional lists, it could also become truer over time, as academics sought to place their best work in the top-ranked journals. Of course, this would only work if the original list allowed for 'excellence' across different kinds of legal research by ensuring that there were journals accepting work from diverse subfields in each tier. While theoretically interesting, the politics of creating the initial list, with institutions having incentives to back their own journal, would be difficult to navigate.

A solution, albeit likely difficult in the context of a body such as CALD that comprises members who currently rely on such lists, is to have a CALD statement that explicitly discourages their use. This could be supplemented by a list of recognised 'quality' journals without distinguishing among them. As LADRN members pointed out, this would involve considerable work. It would be important to have a transparent and fair procedure. This would involve:

- Agreed criteria for inclusion in the list;
- A process for legal researchers to nominate to the list, followed by a process for determining whether inclusion was warranted according to the agreed criteria;
- A process for ADRs (or other research leaders) to appeal decisions to include or exclude particular journals (which could also be done on behalf of members of their school or faculty);
- Ongoing consultation through CALD and LADRN as to the usefulness of the list and the appropriateness of procedures relating to it.

(5) - Citation metrics

Citation metrics, whether for individual outputs or journals, can seem, as numbers often do, to be objective. However, such a number fails to capture reality. For example, citation metrics favour those who publish in the US (necessarily not on topics specific to Australian law). This is because (a) the citation universe, being the number of scholars who encounter the work, is larger and (b) there is a greater tendency in US law journals to cite *everything* relevant to a particular article. There is also arbitrariness in which law journals are included in databases such as Scopus, so that even where a journal *is* included, the journals that might cite it may not be.

These issues are not resolved, as the CALD statement contemplates, merely by adding more material which can be assessed, such as court judgments and government reports. Even if this was possible from a data-collection perspective, there would still be a bias against those publishing, appropriately given the topic, to a primarily Australian audience. This is *not* just a law issue as the ACOLA report highlights when referring to those publishing on Australian indigenous issues or the Australian environment or context. These problems are additional to those that apply across disciplines to individual and journal citation metrics, as summarised in the literature review. It is particularly interesting that ACOLA, which includes STEM "citation" disciplines, has also expressed concern around misuse and overuse of citation metrics and their potential to undermine the purposes of universities and research.

CALD could issue a strong statement against the misuse of citation metrics, particularly those such as Q1 journal rankings that inherently discourage research focussed on the understanding and development specifically of Australian law. As one LADRN member noted, there is some urgency here given the imminent conclusion of the Australian Universities Accord process.

(b) - Impact assessment

This is less well developed for legal scholarship than quality assessment despite the fact that law reform and other positive change remains *an* important purpose of legal scholarship. Productive relationships with government, industry, not-for-profits, the judiciary and the profession are both a frequent ingredient for impact and an outcome of legal scholars engaging in potentially impactful projects. These endeavours take time, and unless impact is acknowledged in performance assessment and workload, only those in a position to ignore assessments of their output, such as those willing to work additional hours or senior scholars in certain institutions, will have the capacity to do this type of work. As one LADRN member observed in written feedback, there are also concerns that 'demand-driven' research for partners might detract from scholarly, researcher-driven projects.

Unlike quality, achievement of impact can in that sense often be "unfair" – the world might be a better place if particular recommendations were adopted, but they may not be because of extraneous factors. Further, impact often takes time and recommendations made in the course of research might be adopted many years later when political or other winds change. Impact assessment needs to be sensitive to these factors, as well as the fact that basic or theoretical research is important despite its impacts on legislation, common law, policy or practice being non-existent or indirect.

(1) – Recognising breadth of impact types

An important finding in our survey is the gap between what individual researchers consider to be impact and what institutions recognise as impact. In our view, impact should be understood broadly recognising that the impact can come through engagement with the profession, through casebooks

and teaching, and through changing people's minds. Given the gap between institutional and individual assessments of impact, there would be benefits in CALD making a revised statement on the breadth of potential impacts of legal scholarship using examples from our survey. This could include explicit recognition of the important role played by policy submissions (even when not commissioned), casebooks, treatises/textbooks and legal encyclopedias.

(2) – Qualitative nature of impact assessment

Given the above point about the variety of kinds of impact that should be recognised, the only way to do the assessment is qualitatively, based on narrative. There is no single quantitative metric that can be used as a proxy, although quantitative measures will play a role for some, for example, where a policy change has been shown to have saved the government a known figure or avoided incarceration of a known number of people. Narratives will allow researchers to capture influence over the law itself by capturing the impact of scholarship and ideas on judicial decisions where, even in the absence of citation, other evidence supports such an impact.. It will allow academics to talk about the number of students who have been influenced in their thinking based on the use of their work in casebooks or course reading lists.

One question that arises is whether it is worth assessing impact in the context of a particular framework, such as the Sustainable Development Goals. While that will be useful for many kinds of impact, particularly law and policy change, it will not capture the diverse array of things that participants in our survey identified as important, and from which people, individually or collectively, derive benefit. However, there are advantages in linking impact with Sustainable Development Goals where appropriate, particularly as some institutions and ranking schemes are already measuring contribution to these.

Part (VII) – Preliminary Recommendations

1) CALD should issue a revised Statement on the Nature of Legal Research, to provide a new generation of legal researchers a clear and concise conceptual framework for describing the work they do. A revised Statement would signal the continuing relevance of much that was in the original Statement. The Restatement should recognise methodological diversity and the various paradigms within which legal research may sit (for example, the humanities, or social sciences, or doctrinal research) and state that appropriate means of assessing the quality of legal research will vary depending on the kinds of legal research being undertaken. The opportunity could also be taken to update the references in this Statement, both in terms of compendiums of scholarship and in terms of references to the RQF. Some particular components of the Statement might be rethought, in particular:

    a. The need for a "law-specific panel". Given what is said earlier about the diversity of legal scholarship, and what our survey reveals about the extent of interdisciplinary scholarship, this might be overstated. Further, as this was referring to the ARC, given the multi-disciplinary nature of all panels, and reliance by panels on expert peer reviewers, it does not seem realistic.

    b. The failure to recognise non-peer reviewed journals is arguably historic, although it is still worth making the point that there are many highly regarded non-peer reviewed law journals, particularly in the United States.

c. Discussion of the use of metrics in quality assessment would be superseded by our second recommendation.

d. Discussion of assessment of impact would be superseded by our fourth recommendation.

2) CALD should issue a Statement outlining Best Practice for Assessing the Quality of Legal Research:

   a. Emphasising the importance of peer review and the need for clear, consistent and transparent standards for peer review and peer review processes
   b. Specifying the kinds of criteria against which quality of legal scholarship can be measured, noting that some of these are broadly applicable and others applicable to particular types of legal research; and
   c. Stating that (1) citation metrics as the basis for quantitative comparison or ranking of outputs or scholars, and (2) journal ranking lists as the basis for measuring the quality of their contents are inappropriate ways of assessing the quality of legal research.

   The Statement would replace the document currently titled "CALD Standing Committee on Research and Scholarship: Recommendation 1.6". The Statement should be shared with Australian law journals, which should be encouraged to state explicitly whether they adopt the CALD standards for peer review. Journals should also be encouraged to recognise excellence in peer review.

   The Statement can itself be short and simple, and in a format that can be shared with the Australian Research Council, university research leadership, promotions and appointments panels, international equivalents to CALD (in the hope of broader disciplinary reform), and international rankings bodies (who might be encouraged, although this will be difficult, to reduce reliance on citation metrics). It can cross-reference this Report as the basis for the recommendations rather than providing a lengthy discussion within the document itself.

3) Given the centrality of peer review to the integrity of research evaluation and the problems identified in this report, CALD should consider producing a statement on best practice for peer review. For example, peer review should be double blind and publication should require at least two supportive reviews by appropriate experts. If CALD endorses the core recommendations in this Report, LADRN could undertake broader consultation (including with law journals themselves) to ensure the practicality of the final text. The requirements should be drafted to preserve the ability of journals to innovate through mechanisms such as open peer review and review sharing. Similarly, this Report provides a starting point for identifying core and "where relevant" criteria for evaluating quality, but further consultation would be required on a final formulation.

4) CALD should issue a Statement on Assessing the Impact of Legal Research, outlining the various forms that evidence of impact can take, and emphasising the necessity of narrative accounts of impact and the importance of peer review. The existing Statement on the Nature of Legal Research provides a basis for this, particularly its emphasis on the importance of textbooks and casebooks, but the increasing prevalence of impact assessment suggests that reiteration and updating would be beneficial. A revised Statement could also reiterate the importance, in the national interest, of engagement among legal academia, the legal profession, policy audiences, and the public which is strengthened through engagement and more diverse dissemination than may be usual for scholarship in other fields. This includes practitioner-oriented journals,

submissions, and reports as well as diverse NTROs. Accordingly, these impact-oriented activities should be recognised as an important component of many legal researchers' work, while also noting that it is not the sole purpose of legal scholarship and is not relevant to all legal scholarship or all legal scholars.[329] As with Recommendation 2, the Statement should itself be brief and in a format that can be easily shared.

5) CALD should consider producing an unranked Australian Quality Law Journals List, with journals listed alphabetically within subject areas (including a general subject area). This could note those journals adopting the CALD peer review standards. Before creating such a list, there would need to be agreed criteria and processes (such as those described above).

6) CALD should continue to publicise the Australian Legal Research Awards and consider encouraging more recognition of excellent legal scholarship through awards and prizes through Australian law journals, disciplinary associations (national and international), scholarly conferences, and universities.

Noting that these recommendations will involve additional work, we propose one or more working groups of LADRN and CALD members to work on various items. We note that some may be particularly urgent given the imminent conclusion of the Australian Universities Accord process. We also call on CALD members to address some of the concerns raised in the report that relate to the operation of institutions and recognition of workload, including the work involved in evaluating quality and impact (as peer reviewers or otherwise).

[329] This links back to the Statement on the Nature of Legal Research, and the diversity of purposes across legal scholarship.

References

Allen Consulting Group, *Measuring the Impact of Publicly Funded Research* (Department of Education, Science and Training of the Australian Government, 2005) <https://www.adelaide.edu.au/rqf/pdf/allenreport.pdf>

Arup, Christopher, 'Research Assessment and Legal Scholarship' (2008) 18(1 and 2) *Legal Education Review* 31

Australian Council of Learned Academies, *Research Assessment in Australia: Evidence for Modernisation* (Report, 2023) <https://acola.org/wp-content/uploads/2023/11/ACOLA_ResearchAssessment_FINAL.pdf>

Australian Government Department of Education, *Australian Universities Accord Interim Report* (Interim Report, 19 July 2023) <https://www.education.gov.au/australian-universities-accord/resources/accord-interim-report>

Ayres, Ian and and Fredrick E Vars, 'Determinants of Citations to Articles in Elite Law Reviews' (2000) 29(1-Part 2) *Journal of Legal Studies* 427

Barker, Susan, 'Exploring the Development of a Standard System of Citation Metrics for Legal Academics' (2018) 43(2) *Canadian Law Library Review* 10

Barnett, Katy, 'Citation as a Measure of "Impact": Female Legal Academics at a Disadvantage?' (2019) 44(4) *Alternative Law Journal* 267

Bartie, Susan, 'The Impact of Legal Meta-Scholarship: Love Thy Navel' (2009) 18(3) *Griffith Law Review* 727

Boom, Willem van and Rob van Gestel, 'Evaluating the Quality of Dutch Academic Legal Publications: Results from a Survey' (2017) 13(3) *Utrecht Law Review* 9

Bopape, Solomon, 'Measuring the Scholarly and Judicial Impact of Accredited Legal Journals in South Africa' (2017) 27(1) *African Journal of Library, Archives and Information Science* 53

Bowrey, Kathy, *A Report into Methodologies Underpinning Australian Law Journal Rankings Prepared for the Council of Australian Law Deans* (Report, University of New South Wales Faculty of Law Research Series, 2016)

Bowrey, Kathy, *Assessing Research Performance in the Discipline of Law: The Australian Experience with Research Metrics, 2006-2011* (Report to the Council of Australia Law Deans, 2012)

Bowrey, Kathy, 'Audit Culture: Why Law Journals Are Ranked and What Impact This Has on the Discipline of Law Today' (2013) 23(2) *Legal Education Review* 291

Butler, Linda, 'Explaining Australia's Increased Share of ISI Publications—the Effects of a Funding Formula Based on Publication Counts' (2003) 32(1) *Research Policy* 143

Carnelley, Marita, 'In Search of the Perceived Quality and Impact of Accredited South African Law Journals: Exploring the Possibility of a Ranking System. A Baseline Study: 2009 – 2014' (2018) 21 *Potchefstroom Electronic Law Journal* 1

Chan, Janet, Fleur Johns and Lyria Bennett Moses, 'Academic Metrics and Positioning Strategies' in Btihaj Ajana (ed), *Metric Culture: Ontologies of Self-Tracking Practices* (Emerald Group Publishing, 2018) 177

Coase, Ronald H, 'The Problem of Social Cost: The Citations' (1996) 71 *Chicago-Kent Law Review* 809

Council of Australian Law Deans, *Statement on the Nature of Legal Research* (Position Statement, October 2005) 1 <https://cald.asn.au/wp-content/uploads/2017/11/cald-statement-on-the-nature-of-legal-research-20051.pdf>

Dibadj, Reza, 'Translantic Publication Fashions In Search of Quality and Methodology in Law Journal Articles' in Rob van Gestel, Edward Rubin and Hans-W Micklitz (eds), *Rethinking Legal Scholarship* (Cambridge University Press, 2017)

Directorate-General for Research and Innovation, *Towards a Reform of the Research Assessment System* (Scoping Report, European Commission, 2021) <https://data.europa.eu/doi/10.2777/707440>

Dumay, Benedict Sheehy and John, 'Examining Legal Scholarship in Australia: A Case Study' (2021) 49(1) *International Journal of Legal Information* 32

Duncan, Terry Hutchinson and Nigel, 'Defining and Describing What We Do: Doctrinal Legal Research' (2012) 17(1) *Deakin Law Review* 83

Duncan, Terry Hutchinson and Nigel, 'Defining and Describing What We Do: Doctrinal Legal Research' (2012) 17(1) *Deakin Law Review* 83

ERA EI Review Advisory Committee, *ERA EI Review 2020–2021* (Final Report, Australian Research Council, 2021)

Ferber, Marianne A, 'Citations: Are They an Objective Measure of Scholarly Merit' (1986) 11 *The University of Chicago Press* 381

Gestel, Rob van, 'Quality, Methodology, and Politics in Doctrinal Legal Scholarship' [2022] *Law and Method*

Gestel, Rob Van, 'Ranking, Peer Review, Bibliometrics and Alternative Ways to Improve the Quality of Doctrinal Legal Scholarship' in Hans-W Micklitz and Edward L Rubin Rob van Gestel (ed), *Rethinking Legal Scholarship* (Cambridge University PRess, 2017) <https://drive.google.com/file/d/1vKuMGdPOqorMi54IozGGVsI52AXyem4u/view?usp=share_link>

Gestel, Rob Van, 'Ranking, Peer Review, Bibliometrics and Alternative Ways to Improve the Quality of Doctrinal Legal Scholarship' in Rob van Gestel, Edward Rubin and Hans-W Micklitz (eds), *Rethinking Legal Scholarship* (Cambridge University Press, 2017)

Gestel, Rob van, 'Sense and Non-Sense of a European Ranking of Law Schools and Law Journals' (2015) 35(1) *Legal Studies* 165

Gestel, Rob van, 'Sense and Non-Sense of a European Ranking of Law Schools and Law Journals' (2015) 35(1) *Legal Studies* 165

Gestel, Rob van and Jan Vranken, 'Assessing Legal Research: Sense and Nonsense of Peer Review versus Bibliometrics and the Need for a European Approach' (2011) 12(3) *German Law Journal* 901

Giblin, Kimberlee Weatherall and Rebecca, *Inoculating Law Schools against Bad Metrics* (Working Paper, 25 January 2021)

Gingras, Yves, *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact*, ed Blaise Cronin and Cassidy Sugimoto (The MIT Press, 2014)

Group, Research Quality Framework Development Advisory, *Research Quality Framework: Assessing the Quality and Impact of Research in Australia* (2006)

Harrison, Jeffrey L and Amy R Masburn, 'Citations, Justifications, and the Troubled State of Legal Scholarship: An Empirical Study' (2015) 3(1) *Texas A&M Law Review* 45

Hicks, Diana et al, 'The Leiden Manifesto for Research Metrics' (2015) 520 *Nature* 429

Horne, Mehmet Pinar and Timothy J, 'Assessing Research Excellence: Evaluating the Research Excellence Framework' (2021) 31(2) *Research Evaluation* 173

Horne, Mehmet Pinar and Timothy J, 'Assessing Research Excellence: Evaluating the Research Excellence Framework' (2021) 31(2) *Research Evaluation* <https://drive.google.com/file/d/1jxBQcgIsFV6i0Z2e1r9EqMeTJeVmF540/view?usp=share_link>

Kissam, Philip C, 'The Evaluation of Legal Scholarship' (1988) 63(2) *Washington Law Review* 221

Leiter, Brian, 'Measuring the Academic Distinction of Law Faculties' (2000) 29(1) *Journal of Legal Studies* 451

Lord Nicholas Stern, *Building on Success and Learning from Experience: An Independent Review of the Research Excellence Framework* (Report, Government of the United Kingdom, July 2016)

Lucas Jr, Gary M, 'Measuring Scholarly Impact: A Guide for Law School Administrators and Legal Scholars' (2017) 165 *University of Pennsylvania Law Review Online* 165

McNamara, Lawrence, 'Understanding Research Impact in Law: The Research Excellence Framework and Engagement with UK Governments' (2019) 29(3) *King's Law Journal* 437

Mehmet Pinar and E Unlu, 'Determinants of Quality of Research Environment: An Assessment of the Environment Submissions in the UK's Research Excellence Framework in 2014, Research Evaluation' (2020) 29 *Research Evaluation* 231

Merton, Robert K, 'The Matthew Effect' (1968) 159(3810) *Science* 56

Molly, Morgan Jones et al, *Assessing Research Impact: An International Review of the Excellence in Innovation for Australia Trial* (RAND Europe, 2013)

Murray, Ian and Natalie Skead, 'Who Publishes Where?: Who Publishes in Australia's Top Law Journals and Which Australians Publish in Top Global Journals' (2020) 47(2) *University of Western Australia* 220

Nous Group, *Australian Universities Accord Discussion Paper Submission Analysis* (Report, Nous Group, 12 May 2023) <https://www.education.gov.au/australian-universities-accord/resources/australian-universities-accord-discussion-paper-submission-analysis>

Osterloh, Margit and Bruno S Frey, 'Ranking Games' (2015) 39(1) *Evaluation Review* 102

Perez, Oren et al, 'The Network of Law Reviews: Citation Cartels, Scientific Communities, and Journal Rankings' (2019) 82(2) *Modern Law Review* 240

Pinar, Mehmet and E Unlu, 'Evaluating the Potential Effect of the Increased Importance of the Impact Component in the Research Excellence Framework of the UK' (2020) 46 140

Quality Framework Development Advisory Group, *Research Quality Framework: Assessing the Quality and Impact of Research in Australia* (Final Report, 2006) 1

Rachel Klesch, Guzyal Hill, and David Price, 'The Academy and the Courts: Citation Practices' (2023) 42(1) *University of Queensland Law Journal* 103

Ramsay, Ian and Geofrey P Stapledon, 'A Citation Analysis of Australian Law Journals' (1997) 21(2) *Melbourne University Law Review* 676

Rehaag, Sean, 'Are Publication and Citation Counts Reliable Indicators of Research Productivity or Impact?' (2018) 43(2) *Canadian Law Library Review* 26

Research Excellence Framework, *Summary Report Across the Four Main Panels* (Report, 2021)

Research Excellence Framework 2021: REF Director's Report (Report, REF, October 2023) <https://archive.ref.ac.uk/media/1918/ref-directors-report.pdf>

Roux, Theunis, 'Judging the Quality of Legal Research: A Qualified Response to the Demand for Greater Methodological Rigour' (2014) 24 *Legal Education Review* 173

Rubin, Edward L, 'On Beyond Truth: A Theory for Evaluating Legal Scholarship' (1992) 80(4) *California Law Review* 889

Rubin, Edward L, 'The Practice and Discourse of Legal Scholarship' (1998) 86(8) *Michigan Law Review* 1835

Russell Smyth, 'Academic Writing and the Courts: A Quantitative Study of the Influence of Legal and Non-Legal Periodicals in the High Court' (1998) 17(2) *University of Tasmania Law Review* 164

Russell Smyth, 'Citing Outside the Law Reports: Citations of Secondary Authorities on the Australian State Supreme Courts over the Twentieth Century' (2009) 18(3) *Griffith Law Review* 692

Russell Smyth, 'Other than "Accepted Sources of Law"?: A Quantitative Study of Secondary Source Citations in the High Court' (1999) 22(1) *UNSW Law Journal* 19

Russell Smyth, 'What Do Intermediate Appellate Courts Cite? A Quantitative Study of Citation Practice of Australian State Supreme Courts' (1999) 21(1) *Adelaide Law Review* 51

Russell Smyth, 'What Do Trial Judges Cite? Evidence from the New South Wales District Court' (2018) 41(1) *UNSW Law Journal*

Siems, Mathias M, 'Legal Originality' (2008) 28(1) *Oxford Journal of Legal Studies* 147

S.K Coombs and I Peters, 'The Leiden Manifesto under Review: What Libraries Can Learn from It' (2017) 33(4) *Digital Library Perspectives* 324

Smyth, Russell, 'Who Publishes in Australia's Top Law Journals?' (2012) 35(1) *University of New South Wales Law Journal* 201

Steinbuch, Robert E, 'On the Leiter Side: Developing A Universal Assessment Tool for Measuring Scholarly Output by Law Professors and Ranking Law Schools' (2011) 45 *Loyola of Los Angeles Law Review* 87

Stojanovski, Jadranka et al, 'Peer Review in Law Journals' (2021) 6 *Frontiers in Research Metrics and Analytics*

Susan Kiefel AC, 'The Academy and the Courts: What Do They Mean to Each Other Today?' (2020) 44(1) *Melbourne University Law Review* 447

Svantesson, Dan Jerker B and Paul White, 'Entering an Era of Research Ranking - Will Innovation and Diversity Survive?' (2009) 21(3) *Bond Law Review* 173

Terry Hutchinson, 'Legal Research in the Fourth Industrial Revolution' (2017) 43(2) *Monash University Law Review* 567

Terry Hutchinson, 'Setting the Scene: Research and Writing in Context' in *Researching and Writing in Law* (Thomson Reuters (Professional) Australia Pty Limited, 3rd ed, 2010)

Thomson, Stephen, 'Letterhead Bias and the Demographics of Elite Journal Publications' (2019) 33(1) *Harvard Journal of Law & Technology* 203

UNSW Law Faculty, *What Makes You Tick?: Report on a Survey of the Factors That Condition High Quality Research* (UNSW Law Faculty, 2012)

Warren, Dennis, 'Australian Law Journals: An Analysis of Citation Patterns' (1996) 27(4) *Australian Academic and Research Libraries* 261

Williams, Kate et al, *An Evaluation of the Impact Component of the Australian Research Council's 2018 Engagement and Impact Assessment* (Report, 2020) <https://www.arc.gov.au/sites/default/files/final_arc_report.pdf>

Osterloh, Margit and Bruno S Frey, *Research Governance in Academia: Are There Alternatives to Academic Rankings?*, CESIFO Working Paper 2797, September 2009

Australian Government Department of Education, 'Australian Universities Accord', *Australian Universities Accord* (Web Page, 19 July 2023) <https://www.education.gov.au/australian-universities-accord>

Australian Government Department of Education, 'Australian Universities Accord Panel' (Web Page, 2 January 2024) <https://www.education.gov.au/australian-universities-accord/panel>

Australian Research Council, 'Engagement and Impact Assessment' <https://www.arc.gov.au/evaluating-research/ei-assessment>

Australian Research Council, 'ERA 2023' <https://www.arc.gov.au/evaluating-research/excellence-research-australia/era-2023>

Australian Research Council, 'Introduction' <https://dataportal.arc.gov.au/EI/NationalReport/2018/pages/introduction/index.html?id=background>

Bennett Moses, Lyria, 'Legal Scholarship and Q1 Journals'

Birmingham City University, 'How REF Is Measured' (2021) <https://www.bcu.ac.uk/research/ref-2021/introduction-to-ref/how-ref-is-measured>

Catherine Renshaw and Lyria Bennett Moses, 'Council of Australian Law Dean's Survey'

Catherine Renshaw and Lyria Bennett-Moses, 'Research Leader Questionnaire'

Coalition for Advancing Research Assessment, 'Agreement on Reforming Research Assessment' <https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf>

Coalition for Advancing Research Assessment, 'CoARA' <https://coara.eu/>

Council, Australian Research, 'ERA EI Review' (2022)

Council of Australian Law Deans, 'CALD Standing Committee on Research and Scholarship: Recommendation 1.6' <https://cald.asn.au/wp-content/uploads/2017/11/CALD-Statement-giving-effect-to-Recommendation-1.pdf>

DORA, 'Home', *DORA* (Web Page, 1 August 2023) <https://sfdora.org/>

Elsevier, 'Our Commitments to Responsible Research Evaluation', *www.elsevier.com* <https://www.elsevier.com/en-au/insights/icsr/responsible-research-evaluation>

'Goodhart's Law', *Oxford Reference* <https://www.oxfordreference.com/display/10.1093/oi/authority.20110803095859655>

Group of Eight Australia, 'About the Go8' <https://go8.edu.au/about/the-go8#:~:text=The%20Group%20of%20Eight%20(Go8,Monash%20University%20and%20UNSW%20Sydney.>

Innovative Research Universities, 'Our Universities' (Web Page) <https://iru.edu.au/our-universities/>

'Leiden Manifesto for Research Metrics' <http://www.leidenmanifesto.org/>

Research Excellence Framework, 'REF 2029' (Web Page) <https://www.ref.ac.uk/>

Research Excellence Framework, 'Your Simple Guide to REF' <https://www.exeter.ac.uk/media/universityofexeter/ref2021/Simple_Guide_to_REF2021.pdf>

Ruhl, JB, 'The Hierarchy of Legal Scholarship', *Jurisdynamics* (2006) <http://jurisdynamics.blogspot.com/2006/09/hierarchy-of-legal-scholarship.html>

Sawczak, Ksenia, 'Assessing Impact Assessment – What Can Be Learnt from Australia's Engagement and Impact Assessment', *London School of Economics and Political Science* (2019) <https://blogs.lse.ac.uk/impactofsocialsciences/2019/05/16/assessing-impact-assessment-what-can-be-learnt-from-australias-engagement-and-impact-assessment/>

Tertiary Education Commission, ‘Performance Based Research Fund: Guidelines for the Quality Evaluation 2026 Assessment Process’ <https://www.tec.govt.nz/assets/Publications-and-others/Guidelines-for-the-Quality-Evaluation-2026-assessment-process.pdf>

The World Conferences on Research Integrity Foundation, ‘Hong Kong Principles’ <https://www.wcrif.org/guidance/hong-kong-principles>