

Australian University Heads of English (AUHE) submission to Australian University Accord Panel on measuring impact of research and Engagement in Literary Studies.

The following summary and article relate directly to the Australian Government Response to point 10 ‘Evaluation of Excellence and Impact’, in ‘Trusting Australia’s Ability: Review of the Australian Research Council Act 2001’. This response indicates the Minister has directed to Australian University Accord Panel to recommend new models of evaluation.

The AUHE (the peak body for the discipline of English literary studies at Australian Universities) submission here is focused entirely on this point, and we would like the panel to consider our arguments with relation to it. The submission includes a summary of some of the main points that concern us, followed by a full discussion paper authored by a representative of AUHE offering more detail on these points. This paper has been peer reviewed and will be published in the scholarly journal *Australian Literary Studies* later in 2023.

Summary

The ARC must recognise scholarship that confirms that major providers of research citations, such as Scopus/Scimago and Clarivate/WoS, simply do not count the majority of citations from books, book chapters and journals in many humanities disciplines.

Research indicates only between 20-27% of humanities journals are covered by these commercial databases so that according to the 2017 KPMG report to FASS UNSW, p 22, 'Any research evaluation using citation data supplied by Scopus or WoS would be based on very small samples of an academic's work, and it is unlikely that this would provide for robust and verifiable assessment of their portfolio' (see also Aksnes and Sivertsen; Mongeon et. al; Martín-Martín et.al). Google Scholar potentially provides a more robust measure of research citations in the humanities and can be applied to all disciplines, and its application should be examined (Martín-Martín et.al). However, we question any data driven approach that insists on being equally applicable to 'all disciplines'. This simply does not recognise or value publishing and research traditions that are specific to different fields. 'The Norwegian Model' which runs successfully in many European countries is designed to account for such field specific differences and we encourage the ARC to examine this model (Aksnes and Sivertsen). For example, humanities book publishing traditions involves research undertaken over many years that is then relevant within its field for many years, even decades for strong work.

So too, commercial databases that attempt to rank journals in a field (such as Scimago's ranking of 'Literature and Literary Theory') are highly problematic. They include many journals that do not sit predominantly within 'Literature and Literary Theory' for example but rather come from high citations disciplines (such as the top ranked journal 'Criminology and Public Policy'). Including a high number of such journals distorts the rankings and citation expectations of those journals that primarily or exclusively relate to the discipline of Literary Studies and thereby radically pollute the data related to 'journal quality'. Rankings lists of this kind are problematic in a discipline such as 'Literary Studies' (FoR 4705) which is divided into numerous sub-fields (something 6 digit FoR codes attempt to partially capture) since these sub-fields will each have their own journals of record, so that rather than a simple hierarchy of acknowledged quality, quality is dispersed into branches in the field. Again we recommend the ARC/ERA look to the methods adopted by the Norwegian Method, which factors in 'field specific publishing traditions' into its ranking methodology (Aksnes and Sivertsen, 3-5).

We strongly contend that in the absence of genuine peer review the ARC/ERA must develop field-specific modes of analysis and evaluation. We strongly contend that disciplinary peak bodies be allowed some time (at least 3 but ideally 6 months) to consult with their fields to respond to problems posed by the ARC/ERA and attempt to provide advice specific to their own fields.

We also recommend that some form of peer review be maintained, even if this a reduced version of peer review.

Evaluating Literary Studies [FOR 4705] in Australia: Bad Data, Bad Peer Review

Professor Anthony Uhlmann (FAHA), Western Sydney University

[On behalf of The Australian University Heads of English]

Abstract

In 2022 the new Labor government cancelled the ERA (Excellence in Research, Australia) round and instructed the ARC to find new ways of accounting for research excellence that are more cost effective than the peer review model that had served in previous iterations of the ERA. To this end the ARC conducted a consultation process that recommended the ERA be discontinued. What is clear in this emerging new order is the imminence of a move from peer review processes (which have been abandoned because of the cost burden placed on universities) to data driven approaches ‘relevant to all disciplines.’ In this light of this it is essential that potential pitfalls with current data driven systems currently operating in relation to and directly affecting research in the Humanities be clearly brought into the light, so that errors that might negatively impact disciplines in the Humanities might be avoided. This paper (part of which was sent to the ARC review as part of the AUHE submission), sets out to examine some of the flaws with current systems of accounting for the value of literary studies. The failings occur in two main ways. One of these is to do with the collection and propagation of what might objectively be qualified as ‘bad data’; as a way of representing these disciplines. This is seen in the major journal rankings and citation ranking systems that currently operate in the university sector. The other is to do with the collection and propagation of what might objectively be called ‘bad peer review’ as a way of representing humanities disciplines in international university ranking systems. Some suggestions are made concerning potentially less flawed systems of accounting.

In the middle of 2022 the new Labor government cancelled the current Excellence in Research Australia round.¹ The ‘ERA’ as it has been known since it began (copying a similar model developed in the UK) was tasked with ranking disciplines in Australian universities through a process that involved an evaluation of publication and other metrics of research achievement, alongside (in disciplines that have historically depended on peer review to underline quality), detailed peer-review of materials submitted to the ERA process by universities and their qualifying disciplines. In calling for a suspension of ERA 2023 the minister for Education, Jason Clare wrote a letter to the Australian Research Council that comprised a ‘Statement of Expectations.’ Specifically with regard to the ERA process this statements reads as follows:

In light of the sector’s concerns about workload, I ask that you discontinue preparations for the 2023 ERA round and commence work to develop a transition plan, in consultation with the sector and my Department, to establish a modern data driven approach informed by expert review. In addition, I ask you to continue your

¹ <https://www.arc.gov.au/evaluating-research/excellence-research-australia>

work with my Department on developing research engagement and impact indicators to inform the Engagement and Impact assessments.

I ask you provide me with a transition plan by the end of 2022, which in addition to any recommendations from the ARC review, can be considered for implementation in 2024-2025.²

This statement was welcomed by the ARC who posted a public response citing the CEO of the ARC Ms Judy Zielke. This reads in part:

In response to the Statement, the ARC will be prioritising the development of a modern data driven approach to Excellence in Research for Australia (ERA) informed by expert review, for implementation in 2024-25. This means an ERA evaluation round will not be undertaken in 2023.³

Both the Minister's statement and the CEO of the ARC's response underlined the importance of data 'informed by expert review'. The ARC then established 'a working group drawing on experts from across the university sector, peak bodies and the Department of Education to provide advice on this transition.'

The review took place, receiving 223 submissions, and its recommendations, authored by Margaret Shiel, Susan Dodds, and Mark Hutchinson, entitled, *Trusting Australia's Ability: Review of the Australian Research Council Act 2001* was sent to the minister and released publicly in April 2023.

A version of the current paper was developed for the Australian University Heads of English's 'The Value of Literature' conference in Melbourne in December 2022, and it was edited down to key points as part of the AUHE submission to this ARC review (as the peak body for the study of literature at Australian universities) in response to those questions specifically related to the ERA. Because the ERA assessment process was felt to place too heavy a burden on universities, the review, as expected, recommended that the ERA process be scrapped. The AUHE working party were pleased that arguments we, among other

² <https://www.arc.gov.au/about-arc/our-organisation/statement-expectations-2022>

³ <https://www.arc.gov.au/news-publications/media/media-releases/arc-welcomes-new-direction-government>

submissions, made about the flaws with metrics based systems used in isolation were clearly accepted by the review panel as outlined below. I quote recommendation 10 from this report in full here:

Recommendation 10: Evaluation of Excellence and Impact

We recommend that:

- i. the role of the ARC in relation to evaluation of excellence, impact and research capability within Australian universities be re-affirmed by inclusion in the ARC Act
- ii. the Excellence in Research for Australia (ERA) and Engagement and Impact (EI) exercises be discontinued
- iii. the resourcing for evaluations be maintained so the ARC retains and internationally recognised expert evaluation capability.
- iv. the ARC collaborates with TEQSA to develop assessment processes that enable TEQSA to draw on the expertise or the ARC to make decisions on the extent to which current and future higher education provider meet research provider standards;
- v. the ARC develops a framework for regular evaluation and reporting on the outcomes of the NCGP program over a timeframe that allows the full impact of research funding to be assessed and the public benefit explained; and
- vi. the ARC develops a program to evaluate current and future research capabilities within Australian universities, giving priority in the first instance to the capability of Aboriginal and Torres Strait Islander researchers and research that impacts on Indigenous Australians.

We do not recommend that ERA and EI be replaced by a metrics-based exercise because of the evidence that such metrics can be biased or inherently flawed in the absence of expert review and interpretation. (Sheil, Dodds, Hutchinson, *Trusting*, 57)

In August 2023 the Government issued a response to this report which agreed to these recommendations ‘in principle’, stating that ‘Research remains a key element of all Australian universities, and reform of the performance measurement and management of university research is critical to ensure that the future contribution of universities is driven effectively’. The Government further indicated that it ‘requested the Australian University Accord Panel’ to ‘consider the recommendations on measuring impact and engagement in university research’ tasking them with providing further recommendations ‘on a new model’

to the Minister of Education ‘alongside the Universities Accord Panel’s Final Report, due in December 2023’.⁴

The recommendations and the response are interesting not only for the clear statement concerning the limitations of applying existing data driven systems to research in Literary Studies, and the Humanities more generally. They also underline that the process of review and building systems of analysis in our fields is ongoing, though will now be passed on to or developed via recommendations from the Australian Universities Accord Panel in collaboration with TEQSA (Tertiary Education Quality and Standards Agency).⁵ The debate remains very much ‘live’, then, and arguments developed in this paper remain relevant to this process. Further, the arguments remain very much relevant to ongoing processes of assessment developed by individual universities throughout Australia (and elsewhere), where, in many if not most cases, data driven metrics are predominantly used to assess the value and impact of research in literary studies.

In light of this it is essential that potential pitfalls with current data driven systems currently operating in relation to and directly affecting research in the Humanities be clearly brought into the light. TEQSA and the ARC, and the research offices of Australian universities, must take into account the failings of certain prominent ‘data based’ systems of ranking, and to clearly recognise the drawbacks of these systems when they are applied to certain disciplines within the humanities. Here I will be focussing on English, but the problem is equally significant for History, and Philosophy and the Creative Arts.

It must in turn be clearly recognised that these failings are often specific to these humanities and creative arts disciplines which work through qualitative methodologies and which value research outputs that are published not just in refereed journals but in refereed books and book chapters and for creative arts in NTROs (Non-Traditional Research Outputs, such as novels, collections of poetry, plays, musical compositions, works of visual art and so on). Here I will leave aside the problem of NTROs which is a separate issue and concentrate on refereed academic publications in Literary Studies [ARC Field of Research Category 4705]. In this paper I will argue that TEQSA and the ARC must consult with peak bodies that

⁴ <https://www.education.gov.au/higher-education-reviews-and-consultations/resources/australian-government-response-trusting-australias-ability-review-australian-research-council-act>

⁵ <https://www.teqsa.gov.au>

represent disciplines in the humanities to work towards developing approaches that better represent those disciplines.

The failings of data driven metrics occur in two main ways. One of these is to do with the collection and propagation of what might objectively be qualified as ‘bad data’ as a way of representing these disciplines. This is seen in the major journal rankings and citation ranking systems that currently operate in the university sector and in particular Scopus/Scimago,⁶ and Clarivate Analytics/Web of Science.⁷ While these systems are well accepted and accurate in regard to many citation-heavy disciplines that publish in journals captured by these providers, they are, as is well recognised, poor in relation to the humanities (see Bornmann, Thor, Marx, & Schier, 2016; Hammarfelt, 2016; Hammarfelt & Haddow, 2018a; Linmans, 2010; Raughvargers, 2011; Gómez-Sancho & Pérez-Esparrells, 2012; Hazelkorn, 2011, 2013).

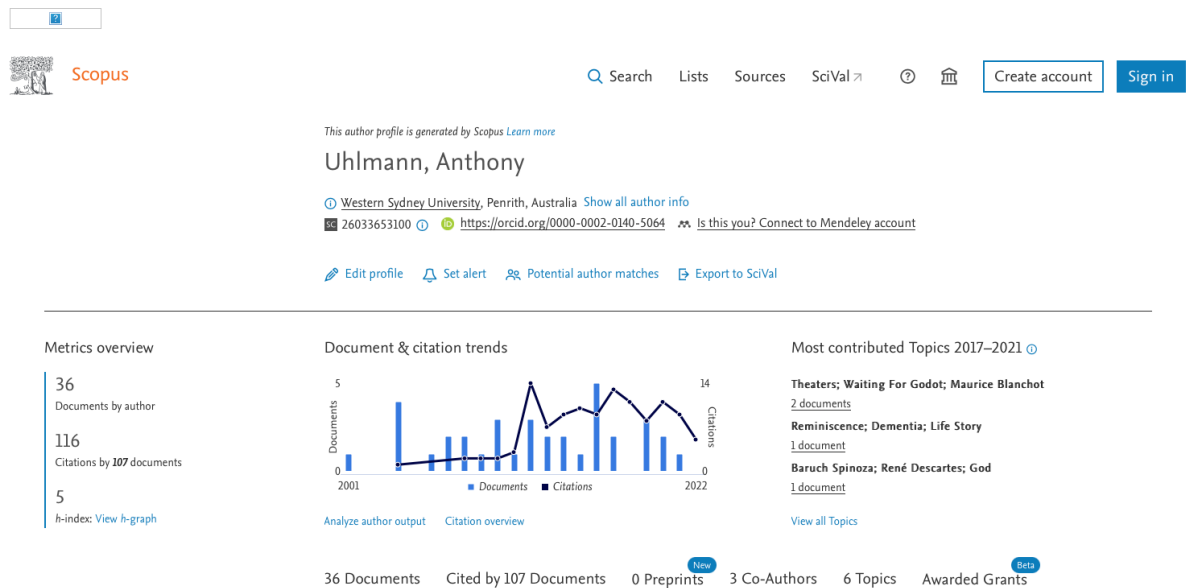
The other is to do with the collection and propagation of what might objectively be called ‘bad peer review’ as a way of representing humanities disciplines in international university ranking systems. This is seen in major university ranking systems that currently operate in the university sector, most prominently in Australia the QS Survey, and the Times Higher Education Survey. In both cases the major flaws are simple to see and are openly admitted to both in the University Rankings and the citation tracking, and have been pointed out by researchers examining these systems of ranking (see Reale, et. al., 2018; Scott, 2013; Sowter, 2013).

The problems, indeed, are somewhat connected. With regard to the tracking of citations of publications in Literary Studies there are many problems but perhaps the most telling is that neither Scopus nor Clarivate (Web of Science) comprehensively count books (academic monographs) or book chapters (peer reviewed chapters in thematically organised academic collections of essays). In addition, while Scopus is estimated to index twice as many journals as Clarivate/Web of Science, according to a 2017 KPMG report, Scopus only counts ‘around 22% of the published work of Australian humanities researchers in a calendar year’ (KPMG UNSW, 22).

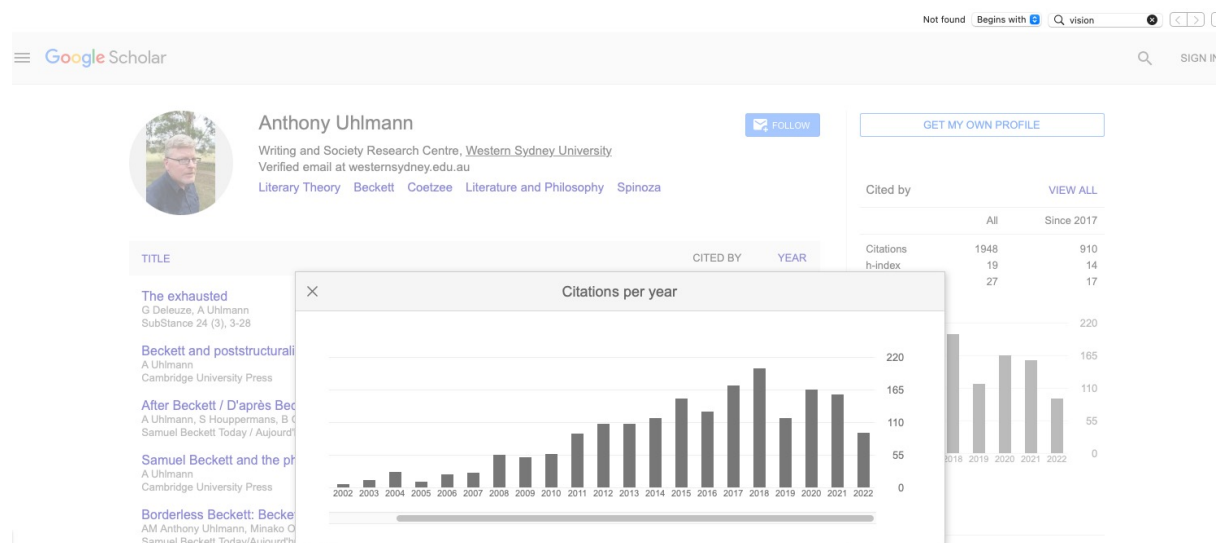
⁶ <https://www.scopus.com/home.uri>; <https://www.scimagojr.com>

⁷ <https://clarivate.com>; <https://webofscience.com>

This means that the kind of citation report Scopus, say, might offer about an individual academic in Literary Studies will not include comprehensive reference to the citations accrued through books they have written or been published in, and will further exclude any citations to journals not included in Scopus. This of course, leads to compounding errors and a major under counting of true citations. I offer an example of the disparity here, using my own individual citations (to avoid ethics concerns), as an example.



Scopus/Web of Science Report, Accessed December 2022



Google Scholar Report, Accessed December 2022

Compare the citations of Scopus/Web of Science and Google Scholar (for a detailed comparative analysis of these databases in relation to Humanities research see Prins, et al., 2016). Note how much lower the stated citations are in Scopus (116 total citations listed) which excludes most of my books and edited collections as opposed to Google Scholar (1948 total citations) which does collect reasonably comprehensive references to books and book chapters. Further detail can be entered into, but this, I think, is sufficient to demonstrate the merit of the conclusion made by KPMG in their 2017 report for UNSW which, citing numerous studies, that states:

Any research evaluation using citation data supplied by Scopus or WoS would be based on very small samples of an academic's work, and it is unlikely that this would provide for robust and verifiable assessment of their portfolio. The limited coverage means that commercial bibliometric databases are not suitable for research evaluation in the humanities (KPMG for UNSW, 2017, 22)

Similar issues then occur for the journals that are included by Scimago and ranked as being the best in particular disciplines as these relate to the disciplines of the humanities. Again, here, I will focus on our discipline which is called 'Literature and Literary Theory' in Scimago.

As we have seen, KPMG cites studies that conclude only about 22% of journals in humanities disciplines are captured by Scopus, which is relied upon by Scimago. It is clear, through a comparison of Journals in 'Literature and Literary Theory' and Project Muse (which aggregates many of the most important journals in the discipline that are largely based in the United States), that many Project Muse journals are simply not captured in Scimago. And yet these same Project Muse journals are often widely considered to be leaders in their fields. This is also true of important journals in Australian literature, which are not captured by Scimago. It is also true of JSTOR which houses many of the most prestigious journals in the humanities in an archive that stretches back over a century. That is, many of the most prominent journals in the field of Literary Studies **do not appear at all** in Scimago. To illustrate this point I began to compare Scimago with Project Muse. Stopping at the letter 'C' in an alphabetic process undertaken in December 2022 the following journals held by Project Muse were found to not be listed in Scimago:

- American Literature
- Antipodes
- ariel: A Review of International English Literature
- boundary 2
- The Byron Journal
- The Cambridge Quarterly
- Canadian Review of Comparative Literature / Revue Canadienne de Littérature Comparée
- Cervantes: Bulletin of the Cervantes Society of America
- Children's Literature
- Christianity & Literature
- The Comparatist
- The Cormac McCarthy Journal

Other methodological decisions further undermine the reliability of the Scimago list as an accurate source to rank the quality of journals in the humanities. Firstly, Scimago looks to short time frames for the citation of sources and ranks its Q1, Q2, etc, journals on an annual basis based on captured citations. However, humanities research, unlike some research in high citations disciplines, takes place over much longer timeframes, with researchers in our fields often engaging with works that are decades and in the case of primary source materials centuries old, and researchers often cite scholarly publications that are many years old given the nature of research in our disciplines (this would be the focus of a separate discussion, but it is crucial to what is going on). This again can be simply illustrated by an example drawn from my own research. My monograph *Beckett and Poststructuralism*, published in late 1999 has continued to draw citations each year since publication, and this is typical of similar monographs in the field.



Anthony Uhlmann

[Edit](#) [Delete](#)

Beckett and poststructuralism

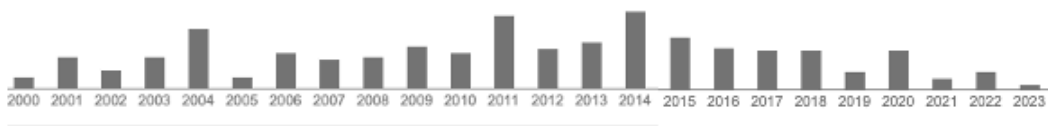
Authors [Anthony Uhlmann](#)

Publication date 1999/9/2

Publisher Cambridge University Press

Description In *Beckett and Poststructuralism*, Anthony Uhlmann offers a reading of Beckett in relation to French philosophy, particularly the work of Foucault, Deleuze and Guattari, Levinas, and Derrida. Uhlmann offers a work of literary criticism that is also a piece of intellectual history, emphasizing how Beckett develops a kind of critical thinking which differs from yet is just as powerful as that of philosophers who, along with Beckett, found themselves faced with sets of ethical problems which were thrown into sharp relief in post-war France. Uhlmann explores the links between ethics and physical existence in Beckett, Foucault and Deleuze and Guattari, and between ethics and language in Beckett, Derrida and Levinas, showing how post-war French philosophy was powerfully affected by Beckett's work. Literature is not reduced to philosophy or vice versa; rather Uhlmann considers how they interrelate and overlap, informing and deforming one another, and how both encounter history.

Total citations [Cited by 240](#)



Scholar articles [Beckett and poststructuralism](#)
A Uhlmann - 1999
[Cited by 237](#) [Related articles](#) [All 4 versions](#)

[Beckett and Poststructuralism \(Cambridge: Cambridge UP, 1999\) *](#)
A Uhlmann - Notes on Contributors
[Cited by 3](#) [Related articles](#)

[Beckett and Postmodernism *](#)
A Uhlmann - 1999
[Cited by 2](#) [Related articles](#)

(Source: Google Scholar, Accessed May 2023)

Secondly, Scimago allows journals to nominate multiple disciplines and on occasion this can radically distort the spread of results. See here the list of what are supposed to be the top 34 journals in Literature and Literary Theory (as accessed in December 2022).

Title	Type	↓ SJR	H index	Total Docs. (2021)	Total Docs. (3years)	Total Refs. (2021)	Total Cites (3years)	Citable Docs. (3years)	Cites / Doc. (2years)	Ref. / Doc. (2021)	
1 Criminology and Public Policy	journal	1.618 Q1	37	29	171	2177	619	141	4.32	75.07	
2 Translation Spaces(Netherland)	journal	1.313 Q1	8	16	44	753	67	42	1.64	47.06	
3 Journal of Writing Research	journal	1.205 Q1	23	14	53	982	149	53	3.38	70.14	
4 Plant Phenomics	journal	1.154 Q1	12	29	53	1771	365	52	6.89	61.07	
5 Poetics	journal	1.102 Q1	67	79	143	5347	334	140	1.92	67.68	
6 Perspectives: Studies in Translation Theory and Practice	journal	0.959 Q1	26	93	189	3789	274	180	1.38	40.74	
7 Written Communication	journal	0.948 Q1	54	17	51	1128	147	50	2.66	66.35	
8 Men and Masculinities	journal	0.848 Q1	63	70	135	3438	425	127	3.05	49.11	
9 Translation and Interpreting Studies	journal	0.697 Q1	14	24	75	1276	72	72	0.86	53.17	
10 Latin American Research Review	journal	0.684 Q1	50	56	143	2756	180	136	1.19	49.21	

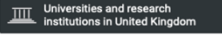
11 South Atlantic Quarterly	journal	0.638 Q1	42	69	175	1801	269	170	1.56	26.10	
12 Empirical Studies of the Arts	journal	0.525 Q1	17	21	44	1049	75	42	1.82	49.95	
13 Cambridge Classical Journal	journal	0.474 Q1	9	9	24	338	11	24	0.33	37.56	
14 Early American Literature	journal	0.454 Q1	14	39	92	1315	27	85	0.29	33.72	
15 Jahrbuch der Osterreichischen Byzantinistik	book series	0.451 Q1	6	22	27	1438	11	26	0.21	65.36	
16 College Composition and Communication	journal	0.436 Q1	37	13	71	422	46	54	0.76	32.46	
17 Translation and Interpreting	journal	0.436 Q1	14	16	54	754	67	54	1.03	47.13	
18 Language and Dialogue	journal	0.421 Q1	11	21	67	933	36	60	0.41	44.43	
19 Narrative Inquiry	journal	0.418 Q1	46	20	56	1146	61	56	1.24	57.30	
20 English in Education	journal	0.401 Q1	21	34	84	1106	65	63	0.79	32.53	
21 Journal of Cultural Analytics	journal	0.398 Q1	2	26	7	1126	4	7	0.57	43.31	
22 L1 Educational Studies in Language and Literature	journal	0.397 Q1	15	32	94	1615	120	93	1.25	50.47	

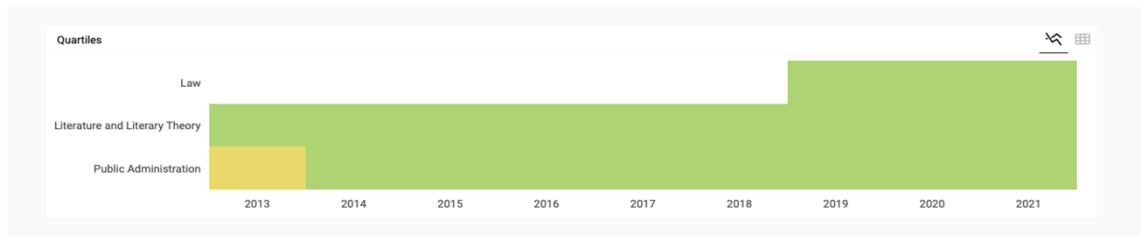
23	Journal of Biblical Literature	journal	0.392 Q1	25	45	186	2899	92	178	0.68	64.42	
24	Rhetoric Review	journal	0.365 Q1	17	24	81	1052	46	81	0.58	43.83	
25	Journal of African Cultural Studies	journal	0.343 Q1	16	57	112	1665	101	78	0.84	29.21	
26	Acta Linguistica Academica	journal	0.341 Q1	11	10	67	529	45	63	0.64	52.90	
27	Performance Measurement and Metrics	journal	0.341 Q1	22	14	49	457	65	48	1.47	32.64	
28	Literature Compass	journal	0.335 Q1	6	33	116	1581	52	112	0.56	47.91	
29	3L: Language, Linguistics, Literature 	journal	0.326 Q1	14	68	149	3089	181	149	1.33	45.43	
30	Studies in American Humor	journal	0.325 Q1	3	19	47	646	61	39	1.81	34.00	
31	International Journal of Education and the Arts 	journal	0.324 Q1	19	10	71	470	45	71	0.51	47.00	
32	Classical Quarterly	journal	0.323 Q1	33	77	253	2145	50	203	0.20	27.86	
33	Vetus Testamentum	journal	0.323 Q1	20	57	129	2590	58	128	0.40	45.44	
34	OCNOS 	journal	0.319 Q1	12	28	66	1106	52	66	0.84	39.50	

(Source, Scimago, Accessed December 2022)

I do not have space here to go through this list exhaustively but will note that many of these journals are from high citations disciplines using quantitative methodologies which for unstated reasons claim ‘Literature and Literary Theory’ as one of their disciplines. However, a quick analysis of a number of these journals indicates that articles related to literary studies are rare within their pages. See for example the number one ranked journals, *Criminology and Public Policy*. The journal’s title clearly situates it within disciplines outside Literary Studies and two other disciplines ‘Law’ and ‘Public Administration’ are also cited in relation to it.

Criminology and Public Policy

COUNTRY United Kingdom 	SUBJECT AREA AND CATEGORY Arts and Humanities └ Literature and Literary Theory Social Sciences └ Law └ Public Administration	PUBLISHER Wiley-Blackwell Publishing Ltd	H-INDEX 37
PUBLICATION TYPE Journals	ISSN 15386473, 17459133	COVERAGE 2012-2021	INFORMATION Homepage How to publish in this journal cpp@gmu.edu



(Source, Scimago, Accessed December 2022)

Further, while it is no doubt the case that some articles are published in the journal that draw upon literature or use works of literature as examples, or perhaps make use of methodologies from ‘literary theory’, clearly many issues include **no references at all** to works of literature. An example is offered below from publications listed on the journal website in 2022. Rather than being the most prominent journal in the field of literary studies, then, this journal seems to only be related to the field in a marginal way.

Articles

Most Recent

Most Cited

ORIGINAL ARTICLE | [Open Access](#)

Are progressive chief prosecutors effective in reducing prison use and cumulative racial/ethnic disadvantage? Evidence from Florida

Ojmarrh Mitchell, Daniela Oramas Mora, Tracey L. Sticco, Lyndsay N. Boggess

Pages: 535-565 | First Published: 4 August 2022

[Abstract](#) | [Full text](#) | [PDF](#) | [References](#)

ORIGINAL ARTICLE

De-prosecution and death: A synthetic control analysis of the impact of de-prosecution on homicides

Thomas P. Hogan

Pages: 489-534 | First Published: 7 July 2022

[Abstract](#) | [Full text](#) | [PDF](#) | [References](#)

ORIGINAL ARTICLE

Clearing crimes in the aftermath of police lethal violence

Aki Roberts, John M. Roberts

Pages: 619-648 | First Published: 7 June 2022

[Abstract](#) | [Full text](#) | [PDF](#) | [References](#)

ORIGINAL ARTICLE | [Open Access](#)

“Culture eats strategy for breakfast”: An in-depth examination of police officer perceptions of body-worn camera implementation and their relationship to policy, supervision, and training

James J. Willis

Pages: 713-737 | First Published: 25 May 2022

[Abstract](#) | [Full text](#) | [PDF](#) | [References](#)

Source: *Criminology and Public Policy* (Journal Website accessed December 2022)

That this problem seems common within the top 100 journals listed in ‘Literature and Literature Studies’ in Scimago attests to there being distorting factors in the application of ranks of Q1, Q2, etc in Scimago, as the high citations journals that publish predominantly in other disciplines that are included will distort the average citations across Literature and Literature Studies generally and push those journals that publish exclusively or predominantly in Literature and Literature Studies to lower rankings. I would argue, then, on

this basis alone, that just as Scopus and Web of Science are not reliable sources for judging citations, Scimago is not a reliable source for judging journal quality.

This is stating the problem, and these problems are already known in our discipline. They need to be better known, however, and better explained to the research offices of universities in Australia and elsewhere. Firstly, then, I merely wish to restate as forcefully as possible that TEQSA, the ARC and the research offices of universities cannot rely on this data to judge citations in the humanities, or to rank journals within the humanities. I will offer suggestions as to what might be done in their place at the end of this paper.

Bad Peer Review

The correspondence related to the ERA cited at the beginning of this paper also mentions data being ‘informed by peer review’ but it is clear that the amount of peer review to be undertaken jointly by TEQSA and the ARC will be greatly diminished from what was in place in previous iterations of the ERA. In the next part of this paper I want to further underline that the ARC/TEQSA and research offices cannot rely on particular publicly available ‘peer review’ of university discipline rankings in the humanities. Again, I underline that their systems are not universally bad. Rather, they work much better for those disciplines that have access to good or excellent citation data. Yet again, working well in one area does not justify using them in relation to other areas where they do not work well.

Note the ranking measures the QS Survey regarding universities in general or the Arts and Humanities in general: there are several categories and these seem comprehensive and involve sophisticated engagement with the data available. However, this approach shifts radically between those disciplines that have clean and consistent data re citations and humanities disciplines. When I wrote the first iteration of this paper in 2022 these methodologies were clearly outlined on the QS Survey webpage under the sub-heading ‘Methodology’. However, the information on methodology has been altered since then and less information is now easily accessible. However, the statement of methodology from December 2022 (accessed now via ‘The Way Back Machine’ which archives web content) stated:

Weightings

As research cultures and publication rates vary significantly across academic disciplines, the QS World University Rankings by Subject applies a different weighting of the above indicators in each subject. For example, in medicine, where publication rates are very high, research citations and the h-index account for 25% of each university's total score. On the other hand, in areas with much lower publication rates such as history, these research-related indicators only account for 15% of the total ranking score. Meanwhile in subjects such as art and design, where there are too few papers published to be statistically significant, the ranking is based solely on the employer and academic surveys.

Please see our [Support Page](#) for more details.⁸

This shows that while 50% of the ranking for a discipline like medicine is based on objective data related to research outputs, only 15% is based on data of this kind for some humanities disciplines. However, other disciplines, including English Language and Literature, are 'based solely on the employer and academic surveys'. These surveys, were, in 2022, described under 'Academic Reputation' and 'Employer Reputation'. These are both 'reputation' surveys.

So then, what, in fact, counts as 'peer review' in ranking disciplines, such as Literary Studies, within universities when no citations are factored in? The answer is two reputation surveys. For academic reputation, for example, email surveys are sent to 'experts in the discipline'. These surveys ask these experts to rank 'up to 10 domestic and 30 international institutions'. They are of course not allowed to include their own institution.

While these 'experts' no doubt *are* experts in their fields, no evidence is provided to suggest they might therefore have detailed knowledge of the work being done at other institutions in their country or region or internationally at any given time, or that they are consulting data of any kind in making these assessments. Of course, they will know something, but it is unlikely

⁸ (accessed via <https://web.archive.org/web/20230129015901/https://www.topuniversities.com/subject-rankings/methodology>) Please note: the link to the Support page no longer functions. The original link <https://www.topuniversities.com/subject-rankings/methodology> now takes you to a different article.

to be knowledge based on a recent analysis of evidence presented as has traditionally been the case, with peer review in our discipline in the ERA.

Rather, these surveys rank universities on historical reputation, whether or not these institutions are currently supporting the discipline in question and helping it to thrive, or, for example, giving little by way of resources to that discipline. The logic of reputation surveys of this kind, which I call 'bad peer review' is one that make it difficult if not impossible for less established institutions to be ranked more highly no matter how much these institutions and the staff working in the discipline invest into the discipline. This in turn will tend to discourage investment in these areas, since no progress is recognised through the surveys.

So too, and I base these observations on work I was asked to undertake to review English departments at a number of institutions, those who work in the disciplines in the higher reputation institutions can and have found that their institution might not invest in their discipline despite impressive results in these ranking systems. This is because, based on the logic of a reputation survey, these results will remain consistently high because of the historical strengths of the discipline rather than its current health. Given the rankings remain high even in the face of extensive cuts to the fields, what is the incentive of universities of this kind to invest in these disciplines?

Again, I will quickly illustrate the kinds of discrepancies that occur with a series of slides, here looking at the QS Survey. I choose this over the Times Higher Education Rankings since QS is the most pertinent to the individual disciplines of the Humanities as it offers disciplinary distinctions that resemble the recognised distinctions that pertain historically in humanities. The THE rankings, do not individually recognise the discipline of English but bundles it together under the heading of 'Languages, Literature and Linguistics' and does the same for most other disciplines making it very difficult to glean meaningful data regarding FoR codes in Australia. Their categories are listed here:

- Art, performing arts and design
- Languages, literature and linguistics
- History, philosophy and theology
- Architecture
- Archaeology

THE also rely on reputation surveys for at least 50% of rankings in the key areas of teaching and research. [<https://www.timeshighereducation.com/world-university-rankings-2023-subject-arts-and-humanities-methodology>]. In order to demonstrate some of the distortions these systems generate I will steer away from Australian universities and use examples from universities from the United States.

It quickly becomes clear that certain institutions are given surprising rankings because of the QS methodology built upon reputation surveys. In the discipline field of ‘English Language and Literature’ some no doubt excellent universities based in non-English-speaking countries are given high rankings, no doubt based on their being ranked highly in the regions they come from. This is not in any way surprising in itself, yet what is surprising is how they rank *above* institutions that have outstanding reputations specifically in English Language and Literature.

Examples of QS Rankings for ‘English Literature and Language’ (December 2022)

21 National University of Singapore	= 79 Georgetown University
=23 The University of Hong Kong	= 83 Rutgers
31 The University of Tokyo	= 84 University of Sussex
36 The University of Amsterdam	= 87 University of North Carolina at Chapel Hill
= 39 Freie Universitaet Berlin	= 94 University of California, Santa Barbara
= 42 Peking University	97 University of Washington
44 Humboldt Universitat zu Berlin	100 University of California Davis
= 48 Sorbonne University	
= 50 Ludwig-Maximilians Munchen	101-150 Birkbeck, University College London
= 55 The Chinese University of Hong Kong	101-150 City University of New York
= 58 Nanyang Technological University, Singapore	101-150 Dartmouth College
= 60 Universiti Malaya	101-150 Purdue University
= 63 Seoul National University, South Korea	101-150 Queen’s University Belfast
= 66 Yonsei University, South Korea	101-150 Royal Holloway University of London
= 68 Albert-Ludwigs-Universitaet Freiburg	101-150 University of Cape Town
= 70 Leiden University	101-150 University of East Anglia
= 71 National Taiwan University	101-150 University of Massachusetts Amherst
= 75 Kyoto University	101-150 University of Reading
= 77 University of Copenhagen	
	151-200 Boston College
= 79 Ruprecht-Karls-Universitat Heidelberg	151-200 Emory University
= 84 University of Helsinki	151-200 Queen’s Canada
	151-200 Rice University
101-150 Beijing Normal University	151-200 University of Buffalo SUNY
101-150 Fudan University	151-200 University of California, Santa Cruz
101-150 Ghent University	151-200 University of Florida
101-150 Lomonosov Moscow State University	151-200 University of Notre Dame, United States
101-150 Tsinghua University Beijing	151-200 University of Ottawa, Canada
101-150 Universidad Nacional Autónoma de México (UNAM)	
101-150 Universidad de Buenos Aires (UBA)	201-250 University of Aberdeen
101-150 Universitat de Barcelona	201-250 Fordham University
101-150 Universiti Kebangsaan Malaysia (UKM)	201-250 Northeastern University, Boston
101-150 Waseda University	201-250 Tufts University
101-150 Zhejiang University Hangzhou	301-320 Stony Brook University SUNY
	301-320 Syracuse University
151-200 Al-Farabi Kazakh National University, Kazakhstan	301-320 University of Manitoba
151-200 Ateneo de Manila University	301-320 University of Oklahoma
151-200 Beijing Foreign Studies University, Beijing	301-320 University of Waterloo, Canada
151-200 Chulalongkorn University, Bangkok Thailand	
151-200 Concordia University, Montreal Canada	
151-200 De La Salle University Manila Phillipines	

Here we can see how some universities with outstanding track records and reputations in English Literature and Language teaching do not rank well in the QS system. Just to take a few examples, John Hopkins is ranked 73 in QS when it is ranked 13 in the US News Best

English Programs. Rice University is ranked 151-200 in QS and 41 in US News Best English Programs. SUNY Buffalo is ranked 151-200 in QS and 46 in US News. Fordham University is ranked 251-300 in QS and 53 in US News.

These excellent schools, all of which would claim high status and standing in the field of English Language and Literature, are ranked well below what one might expect in QS. While Hong Kong and Singapore are no doubt excellent universities, their English Departments do not match or surpass those of John Hopkins or Rice University. This in part is due not only to the nature of the reputation survey, but the particular methodology adopted by QS which asks those who undertake its survey to rank ‘10 domestic [...] institutions in their field’

Drawing on responses from over 130,000 academics, respondents are asked to list up to 10 domestic and 30 international institutions which they consider to be excellent for research in the given area. The results of the survey are then filtered according to the narrow area of expertise identified by respondents. (<https://www.topuniversities.com/subject-rankings/methodology>, accessed December 2022)

This method, statistically, will have the effect of causing those institutions that fall just outside, say, a reputationally recognised top ten, to gather very few points notwithstanding their actual excellence. They will then in turn be disadvantaged against other institutions in different countries that gather votes related to their own domestic order of rankings.

While the US News rankings are also 100% based on reputational surveys and so equally subject to error and reputational bias, their methodology is at least more nuanced. Here the surveys are sent to departmental heads (rather than individual experts who may choose not to respond as in QS) and rather than being asked to list a top 10 they are asked to give grades to each institution.

The questionnaires asked respondents to rate the academic quality of the programs at other institutions on a 5-point scale: outstanding (5), strong (4), good (3), adequate (2) or marginal (1). Individuals who were unfamiliar with a particular school's programs were asked to select ‘don't know.’

(<https://www.usnews.com/education/best-graduate-schools/articles/social-sciences-and-humanities-schools-methodology>, accessed December 2022)

While not solving the inherent problems of the reputation survey it is arguable that the method of *requiring* heads of departments, who are more likely to take an interest in the structure of their own schools and those of rivals and feel obligated to respond in a timely and conscientious way, is superior to that of choosing generally defined experts in the fields who might not represent all institutions and are not obliged to respond. Secondly, the scale used by US News, while reasonably crude, at least avoids the problem of giving very few points to *what might be excellent schools* (because they fall outside of the generally recognised top ten). Indeed, it will further require the respondent to reflect somewhat on the quality of the programs offered, rather than to call to mind institutions with historically strong reputations in the field.

What Might Be Done?

So then, these are the problems but what might be done about them? Firstly, I would argue that the ARC, TEQSA, and Federal Government in recognising the limitations of data sets, and of limited peer review, allow peak bodies in the humanities that represent certain disciplines to consult within their disciplines for a reasonable period about rankings in those disciplines. This consultation process will then allow disciplines to develop strategies and methodologies that might viably replace any reliance on Scopus, Web of Science or University rankings systems. This at least would allow for an informed consideration of the pros and cons of proposed methodologies.

My own suggestions are as follows. While it has certain drawbacks, Google Scholar allows for a much more accurate collection of citations than Scopus or Web of Science. While the KPMG report argues that using Google Scholar is not possible for two reasons, (one that manual processing is required, and two because it uses an ‘open’ rather than ‘closed’ system of collection, KPMG, 23; Linmans, 2010; Hammarfelt, 2016) I believe neither of these arguments are compelling. Academics are already required to provide individual data to their institutions and maintaining their own Google Scholar pages and providing access to them via links on university webpages is already done in some institutions. Secondly, the ‘closed system’ of Scopus and WoS which only collect from journals within their systems is, as I have stated above, a *problem* for humanities academics rather than a benefit. Google’s ‘open’

system in Literary Studies, at least, is readily justifiable, and, at least with regard to refereed publications (journal articles, book chapters and books), which are the standard in Literary Studies, the citations that accrue to them on Google scholar are almost exclusively drawn from other refereed publications.

Other issues might arise related to harvesting this data, yet web-scraping and data trawling is extremely common in business and government and many systems exist to allow for coding that facilitates this process. If universities provide links to pages (since Google's own links do not always work for sites that wish to announce themselves as public) then it should be possible or be made possible for universities and the ARC to harvest this data.

What else can be done? One must question any data driven approach that insists on being equally applicable to 'all disciplines'. This simply does not recognise or value publishing and research traditions that are specific to different fields. 'The Norwegian Model' which runs successfully in many European countries is designed to account for such field specific differences and we encourage the ARC and TEQSA to examine this model (see detailed outlines of this model, Siversten, in Ochsner, Hug and Daniel; Aksnes and Sivertsen; Sivertsen 2018; Hammarfelt 2018b; Engels and Guns). Rather than inventing an entirely new system, it makes much more sense to look to a model such as this, which both seeks to directly address the kinds of distorting factors outlined in this paper, and which has been running with apparent success for some time in a number of countries.

However, I will venture some further suggestions that might be potentially useful at the discipline level within universities as a means of evaluating the performance of our disciplines. Here again the methodologies will need to be clear.

Through the numbers of publications, grants received, consideration of citations from Google Scholar of a number of staff in the discipline, Postgraduate completions, objective measures could be accumulated. The publications might achieve points, for example, through markers of quality, added to Google citations. Such markers could be suggested and refined by disciplinary working groups within peak bodies taking into account what is and is not practical in relation to gleaning of data. These recommendations could then be used to develop something approaching the 'modern data driven approach informed by peer review' the Minister requested in 2022.

Works Cited

- Aksnes, Dag W., and Gunnar Sivertsen. 'A Criteria-based Assessment of the Coverage of Scopus and Web of Science.' *Journal of Data and Information Science*, vol. 4, no. 1, 2019, pp. 1-21.
- Bornmann, L., Thor, A., Marx, W., and Schier, H. 'The application of bibliometrics to research evaluation in the humanities and social sciences: An exploratory study in normalized Google Scholar Data for the publication of a research institute.' *Journal of the Association for Information Science and Technology*, vol. 67, no. 11, 2016, pp. 2778-2789. doi:10.1002/asi.23627
- Engels, Tim C. E., and Raf Guns. 'The Flemish Performance- based Research Funding System: A Unique Variant of the Norwegian Model.' *Journal of Data and Information Science*, vol. 3 No. 4, 2018, pp 45-60. doi: 10.2478/jdis-2018-0020.
- Hammarfelt, B. 'Beyond coverage: Toward a bibliometrics for the humanities.' *Research Assessment in the Humanities: Towards Criteria and Procedures*, edited by M. Ochsner, S. Hug and H. Daniel, pp. 115-132. Springer Open, 2016.
- Hammarfelt, B., and Haddow, G. Conflicting measures and values: How humanities scholars in Australia and Sweden use and react to bibliometric indicators. *Journal of the Association for Information Science and Technology*, vol. 69, no. 7, 2018b, pp. 924-35. doi:10.1002/asi.24043.
- Hammarfelt, Björn. 'Taking Comfort in Points: The Appeal of the Norwegian Model in Sweden.' *Journal of Data and Information Science*, vol. 3, no. 4, 2018b, pp 85–95. doi: 10.2478/jdis-2018-0023
- KPMG Australia. 'Identifying Research Performance Measures and Performance Management Mechanisms for the Faculty of Arts and Social Sciences.' *Final Report for UNSW*, 2017.
- Linmans, A. Why with bibliometrics the humanities does not need to be the weakest link: Indicators for research evaluation based on citations, library holdings, and productivity measures. *Scientometrics*, vol. 83, 2010, pp. 337-354. doi:10.1007/s11192-009-0088-9
- Rauhvargers, A. 'Global university rankings and their impact.' *Brussels: European University Association*, 2011. Retrieved from <https://eua.eu/resources/publications/384: global-university-rankings-and-their-impact.html>

- Martín-Martín, Alberto, and others. 'Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories.' *Journal of Informetrics*, vol. 12, 2018, pp. 1160-1177.
- Mongeon, Philippe and Adele Paul-Hus. 'The journal coverage of Web of Science and Scopus: a comparative analysis,' *Scientometrics*, vol. 106, 2016, pp. 213-28.
- Pölönen, Janne. 'Applications of, and Experiences with, the Norwegian Model in Finland.' *Journal of Data and Information Science*, vol. 3 no. 4, 2018, pp. 31-44. doi: 10.2478/jdis-2018-0019
- Prins, A.M., and Rodrigo Costas, Thed N. van Leeuwen, Paul F. Wouters. 'Using Google Scholar in research evaluation of humanities and social science programs: A comparison with Web of Science data.' *Research Evaluation*, vol. 25, no. 3, July 2016, pp. 264-70, <https://doi.org/10.1093/reseval/rvv049>
- Reale, E., and others. 'A review of literature on evaluating the scientific, social and political impact of social sciences and humanities research.' *Research Evaluation*, vol. 27, no. 4, 2018, pp. 298-308. doi:10.1093/reseval/rvx025
- Scott, P. 'Ranking higher education institutions: A critical perspective.' *Rankings and accountability in higher education: uses and misuses*, edited by P. Marope, P. Wells and E. Hazelkorn, pp. 113-28. UNESCO, 2013.
- Shiel, Margaret, Susan Dodds, and Mark Hutchinson. *Trusting Australia's Ability: Review of the Australian Research Council Act 2001*, 20 April, 2023.
- Sivertsen, Gunnar. 'The Norwegian Model in Norway'. *Journal of Data and Information Science*, vol. 3, no. 4, 2018, pp 3-19. doi: 10.2478/jdis-2018-0017
- Sowter, B. 'Issues of transparency and applicability in global university rankings.' *Rankings and accountability in higher education: Uses and misuses*, pp. 55-69.