

Getting cited: A reconsideration of purpose

Martin Davies & Angelito Calma

University of Melbourne

Michael Calver's recent exhortation 'Please don't aim for a highly cited paper' (AUR, 57(1): pp. 45-49) is welcome and a timely reminder of the problems associated with seeking citations at any cost. While not disagreeing with the concerns he raises we offer another way of looking at citation-seeking; thereby outlining a reconsideration of its purpose. We suggest that citations indirectly help to shape the terrain of a discipline. By providing an analysis of citation data from two key higher education journals, we show how citations are a measure of the 'geography' of a discipline, i.e., the networks of influence of key thinkers and the keywords that reveal scholarly interests and practices. This, in turn, provides us with information that is revealing about the nature of disciplines themselves. This paper provides a summary of data from an ongoing research program we are conducting that analyses the citation metrics of key journals in the field.

Keywords: Citation networks, Gephi™, citation analytics

Introduction

Michael Calver (AUR, 57, No. 1) has recently reminded us of the problems associated with aiming for highly cited papers. We summarise these as follows:

1. It narrows the scope of research undertaken. Calver reminds us that some areas of worthy research (e.g. taxonomy) attract few citations. Aiming purposively for high citations is tantamount to choosing 'hot topics' over worthy ones.
2. It reduces intrinsic rewards and true collaboration and fosters misconduct. He reminds us that intrinsic motivation declines when rewards are offered. Put simply: researchers are more likely to pursue 'normal science' than 'revolutionary science' simply because the rewards for the former are greater, in the form of high citations, and where there is little evident benefit of the latter. Moreover, the 'contrived congeniality' resulting from reduction in intrinsic rewards leads to back-scratching in the pursuit of citations, marginalising junior colleagues in writing papers (so they may cite them later on) and misconduct – e.g., manipulation of research data and violation of ethical standards.
3. Citations are controversial. Non-existent papers or errors in citations can result in high citation rates; deliberately nonsensical papers can too. We are reminded of the Sokal Hoax: a paper constructed using computer-generated semi-literate nonsense that was published in a refereed journal before the hoax was finally revealed.
4. Successful research careers are based less on highly cited papers and more on persistence and influence. While highly cited papers can be influential, it does not follow that low-cited papers are not influential. Indeed, as Calver points out, 'significant influence on a discipline is more likely to follow from consistent performance' (p. 47). He uses the example of more

influential novelists (Hemingway, Faulkner) who sold fewer copies than their less influential peers.

5. Citation counting impedes innovation. Obsession with citations leads to obsession about one's output and narcissism both about oneself and one's career. This mitigates against good research as it focuses on quantity rather than quality, and loses sight of the immeasurable benefits of working on important topics, irrespective of their output in terms of citations.

Calver's conclusion from all this is the exhortation: 'please don't aim for a highly cited paper'. His argument is: High citations lead to problems so don't aim for a highly cited paper.

We agree with all of the above as far as it goes. However, we'd like to offer a corrective to the argument while not implying that we disagree with Calver's concerns. Aiming for high citations *ne plus ultra* is not an end in itself. However, it does not follow that data on excessive citation rates are useless. Indeed, we think that citation data can help reveal the nature of the scholarly enterprise. Let us explain.

'Geographies of Influence'

For decades now, the aim of collecting citation data has been mainly for administrative purposes, and a way of 'bean counting' the output of staff. It has been used secondarily for hiring and firing decisions. This is widespread, and academic staff are required to obtain points for their publications in an effort to meet requirements for decisions about tenure and promotion. Citation metrics have also been used to provide 'league tables' of scholarly journals in an attempt to steer research output into areas of specialisation. The now abandoned ERA ranking tables was an example of this (Moosa, 2011).

However, recently there has been a move in some disciplines to use citation data to tell us something about the character and history of scholarly disciplines themselves (in much the same way, perhaps, as animal droppings tell us something about the diet and social life of animals). This work broadly falls under the emerging field of citation network analysis, itself a sub-field of research analytics.

We submit that networks of citation data, if analysed through the entire life cycle of a journal's history, can tell us a great deal about the 'geography' of a discipline. This gives another reason for seeking high citations for papers besides personal aggrandisement. High citations, and citation networks are more than a (flawed) measure of influence; they also tell us about the terrain of a discipline,

and this has a number of benefits. We provide examples of this from two key higher education journals: *Studies in Higher Education* (based in the UK) and *Higher Education* (a European journal based in the Netherlands). We suggest that these data are revealing about the nature of scholarly practices. We use the discipline of higher education in what follows but the same analysis can be applied to any discipline, from Astrophysics to Zoology.

Surprisingly, this has not been attempted before to our knowledge – at least, not for the discipline of higher education. The only exceptions to this have been studies that are fragmentary. These studies are noted below.

- Westbury analysed the citation data over a ten-year period for the *Journal of Curriculum Studies* using only six years of available data (1972-78) (Westbury, 1980). He found, amongst other things, that the journal did not show expected widespread evidence of influence across national boundaries.
- Budd (1990) investigated the published material commonly cited in the higher education literature and the most frequently cited. Basing his analysis exclusively on US journals, he determined that US-based academics Astin, Pascarella, Feldman, Centra and Cameron were the most commonly cited authors in the discipline of higher education.
- Budd and Magnuson (2010) followed-up this study two decades later and found some expected variation in the most cited authors, but still US-based academics (Astin, Pascarella, Tinto, Kuh, Cabrara).
- Kandlbinder (2012) assessed the citations from 15 consecutive issues over a five-year period of the Australian journal *Higher Education Research and Development* (1982-1986). He compared this to 16 issues published in a second tranche (2008-2010). Comparing three prominent journals (*Higher Education* (HE), *Studies in Higher Education* (SHE), and *Teaching in Higher Education*) with the data from HERD, he found that one author (Biggs) appeared in all four journals. He also noted that Marton, Entwistle and Ramsden appeared in three of the journals. Other highly-cited authors – namely, Clark, Becher, Barnett, Boud, Slaughter, and Rowland – featured in one or other journal, but not all four. Unlike Budd's studies, Kandlbinder's data suggested no nationality bias.
- Tight (2008) attempted a synoptic, longitudinal study over one year of publication looking at several higher education journals. Investigating 17 English language higher education journals published outside North America in the year 2000, he found that Clarke, Ramsden, Becher, Biggs and Boud occupied the top-five

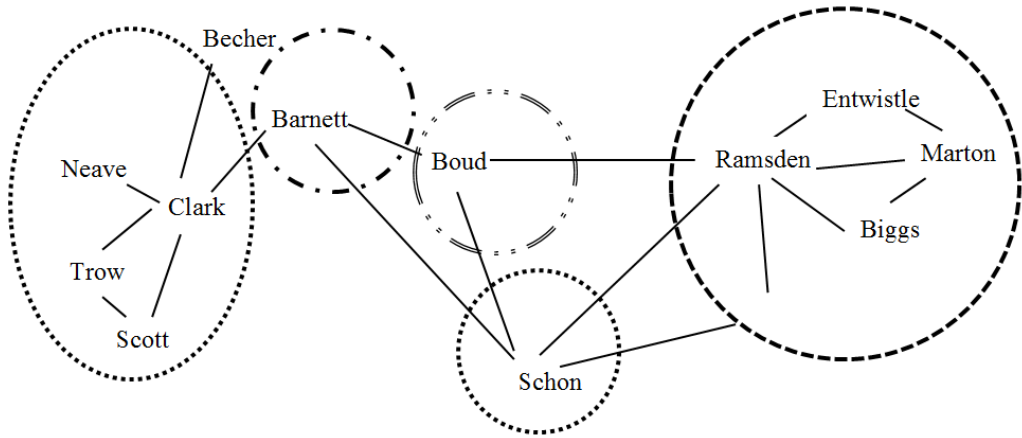


Figure 1: Clusters of relationships among selected authors (Tight, 2008)

positions of 'most cited' authors. A disproportionately large Australian-based contingent of scholars featured in the top 20 'most cited' (Ramsden, Biggs, Boud, Prosser, Trigwell). Tight proposed that scholarly influence can be seen in terms of 'clusters of relationships', with some researchers lying outside the major spheres of influence. See Figure 1 (circles have been added to show the clusters).

- Finally, MacFarlane (2012) devised a tongue-in-cheek 'map' of the field of higher education from his recollections from working in the

discipline (Fig. 2). Drawn from experience, rather than hard data, he notes a division in 'policy' research and 'teaching and learning' research - each forming 'islands' of scholarship that have little to do with each other. Between them a 'sea of disjuncture' has formed, and various 'reefs' associated with specialist areas such as 'identity', 'philosophy' and 'professional development'. MacFarlane's map is presented in Figure 2.

We decided to bring empirical data to bear on MacFarlane's 'archipelago' and Tight's 'clusters of relationships'. We also decided to go further than the

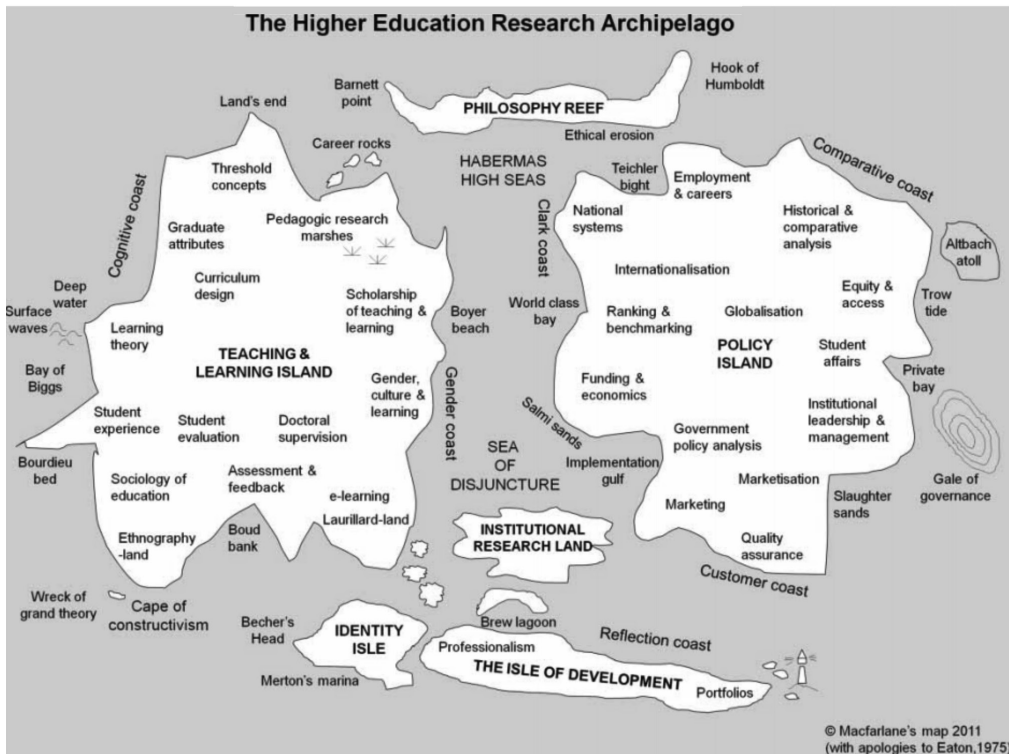


Figure 2: Macfarlane's higher education research archipelago (2012)

fragmentary analyses of Westbury, Budd, Budd and Magnusen, Kandlbinder and Tight. We wanted to analyse all citations from all extant volumes and issues of some key journals in the field, both in terms of what was published, and who cited whom (in the case of SHE we also looked at which keywords were used). Naturally, we had to narrow our focus so we chose two of the key journals in the field, SHE and HE. What would we find, and what might it tell us about citation practices? What would it tell us about key influences? Key scholars? What would it tell us about nationality biases, if any? What would the terrain of the discipline look like?

Studies in Higher Education (1976-2013)

SHE is one of the leading journals in higher education. To analyse it, 38 years of data was available, from the journal's inception in 1976 until our arbitrary cut-off point of 2013 (see Calma & Davies, 2014).

We analysed 32,738 references from 1,056 articles using Web of ScienceSM and Excel, and visualised them as network diagrams using the software GephiTM. The search was performed in July 2013 to include all SHE publications from the default year of 1900 to July 2013. In order to create a citations network for analysis, 'nodes' and 'edges' files were prepared. The nodes contained all citing authors and cited authors while the edges file contained information about the relationship between the two. Nodes and edges files were prepared in Excel and imported into the GephiTM software for each of the 'authors' and 'keywords' GephiTM files.

Unlike the authors' GephiTM files, the keywords files were incomplete. This is because only articles from 2010 onwards appeared to have associated keywords, leaving us only 218 articles with keywords. However, this still represents the entire list of SHE articles with keywords. The keywords GephiTM file resulted in 1,248 nodes and 991 edges.

The result was a complex web of authors and cited references too big to fit on paper. (For those interested in the gritty detail we provide a dynamic web-based version of the author data at <http://tinyurl.com/pnby8xq>).

The diagram shows the entire history of the journal in terms of its citation networks (i.e. who cited whom). It is, as it were, the citation 'geography' of *Studies in Higher Education*. To see how these connections clustered, we focus on just a few authors with family names starting with A. The resulting image is also available online for detailed viewing: <http://tinyurl.com/p30lkr6>.

Drilling down further in this online diagram, we can see that authors such as Adelman, Anderson, Andrews, Archer

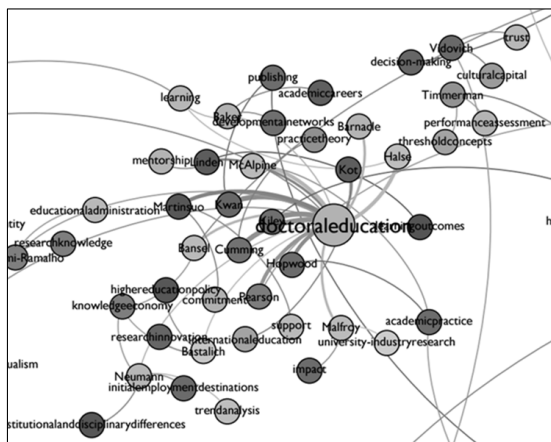


Figure 3: A detailed section of the doctoral education sub-network in SHE.

and Argyris, are connected to, or were cited by, a number of other authors. It is fair to infer from these clusters that research topics are shared; this assumes authors cite only articles relevant to their research.

We drew the following conclusions from the data:

- Most published. We found out that Richardson, Kember, Becher, Boud and Elton have been the most published in the journal's history.
- Most cited. The most cited, on the other hand, were Entwistle, Martin, Ramsden, Biggs and Becher.
- Single or multiple authorship. There is a trend away from solo towards multiple authorship. Single author works accounted for a greater proportion of articles in the 80s/90s than in recent decades. The demands from academics to publish under a 'publish or perish' regime may have changed how researchers reconceptualised the value of working with peers.

We recently ran another analysis using Web of Science on 9th June 2015, using the years 1900-July 2013. We found 1,132 articles and the following presents some additional information about the data:

- Top countries. The top publishing countries - in decreasing order - are England, Australia, Scotland, USA and South Africa.
- Top years. The top publications years are 2013, 2012, 2009, followed by 2011 and 2010.

We did another analysis of keywords. The online diagram (see <http://tinyurl.com/p39cs3e>) shows all the keywords and their connections to authors. (Discrete, marginally or unconnected keywords - i.e. those used only once by a single author - appear at the periphery of the circle, with the most strongly connected, most used keywords appearing in the centre.)

We found the following:

Table 1: Top 10 publishing countries

No.	Country/Territory	Records	% of 2,167
1	USA	431	19.889
2	Australia	358	16.521
3	England	344	15.874
4	Canada	126	5.814
5	Netherlands	114	5.261
6	South Africa	80	3.692
7	Israel	70	3.230
8	Scotland	56	2.584
9	Sweden	52	2.400
10	Spain	50	2.307

- Most discussed. The most frequently listed keywords (since their inception) throughout the journal's history (excluding the keyword 'higher education') were doctoral education, assessment, phenomenography, student learning and identity. From this we infer that these describe the most frequently discussed topics.
- Top topic. Doctoral supervision is the single most discussed topic in SHE. A dynamic, interactive version of our keywords data-map is available online (<http://tinyurl.com/oqt59w4>) (a fragment is provided above). By zooming in, a detail of the doctoral education topic network shows that it branches out to sub-networks of authors who coined the topic and the specific topics of interest of those authors. It reveals a cluster of seven major sub-topics/issues: university management, assessment, transition and internationalisation, student

Table 2: Top 10 publishing institutions

No.	Institution	Records	% of 2,2167
1	Open University	45	2.077
2	University of Lancaster	44	2.030
3	Monash University	33	1.523
4	University of Sydney	32	1.477
5	University of Twente	30	1.384
6	University Melbourne	30	1.384
7	University of New England	27	1.246
8	University of Queensland	26	1.200
9	University of Cape Town	26	1.200
10	University of Hong Kong	25	1.154

identity, university leadership and student experience (See Figure 3).

The subtle patterns of influence deserve more analysis than we can attempt here. Indeed, mining these data to show the precise relationships has the potential to be a cottage industry.

Higher Education (1972-2014)

We also analysed the entire history of publication in the journal HE. This amounted to 2,176 articles and 68,009 references. Like SHE, we were interested in the most published authors and the most cited authors. We have not yet done a keyword analysis. Instead, we were particularly interested in the most cited articles, the top publishing countries, the top publishing institutions and the top publishing years.

Table 3: Top 10 most cited articles

No.	Publication year	No. of citations	% of 68,009
1	Clark, B. R. (1983). Higher education system: academic organisation in cross-national perspective.	109	.160
2	Marton, F. & Saljo, R. (1976). On qualitative differences in learning (I, II): outcome and process. <i>British Journal of Educational Psychology</i> , v46.	103	.151
3	Slaughter S. & Leslie, L. (1997). Academic capitalism: politics, policies and the entrepreneurial university.	83	.122
4	Entwistle N. J. & Ramsden, P. (1983). Understanding student learning.	82	.121
5	Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P. and Trow, M. (1994). The new production of knowledge: the dynamics of science and research in contemporary societies	81	.119
6	Clark, B. R. (1998). Creating entrepreneurial universities: organisational pathways of transformation	58	.085
7	Becher, T. (1989). Academic tribes and territories: intellectual enquiry and the cultures of disciplines.	57	.083
8	Becher, T. and Trowler, P. (2001). Academic tribes and territories: intellectual enquiry and the cultures of disciplines.	52	.076
9	Prosser M. & Trigwell, K. (1999). Understanding learning and teaching: the experience in higher education.	46	.068
10	Biggs J. B. (1987). Student approaches to learning and studying	45	.066

Table 4: Top 10 most cited authors (as single or first author)

No.	Author	No. of times cited	% of 68,009
1	Entwistle, N. J.	472	.694
2	Clark B. R.	396	.582
3	Marton, F.	391	.575
4	Biggs, J.	369	.543
5	Ramsden, P.	258	.379
6	Becher, T.	222	.326
7	Slaughter, S.	189	.278
8	Prosser, M.	117	.172
9	Gibbons, M.	107	.157
10	DiMaggio, P.J.	60	.088

We found that USA, Australia and England make up 52 per cent of 2,167 articles while Open University, University of Lancaster and Monash University provided the greatest number of published articles (122 of 2,167). Further, the top 20 universities included 10 from Australia – indicating, again, a disproportionate influence. *Higher Education* also achieved its greatest annual output of articles in 2012 followed by 2009 and 2011. The top eight years of annual article output were all from the 2000s while 1973 ranked ninth and 1994 tenth.

B.R. Clark's 1983 article 'Higher education system: academic organisation in cross-national perspective' was the most cited article in the entire history of HE (109 citations). See Table 3. Another article by Clark in 1988, 'Creating entrepreneurial universities: organisational pathways of transformation' can be seen in Table 3, also belonging to the top 10, and ranks 6th.

From Table 4, Entwistle is the single most cited author (472 times) followed by Clark, Marton and Biggs.

Empirical evidence is brought to bear on Macfarlane's notion of an 'archipelago'. It shows the 'islands' of influence and the geography of the discipline – or at least

Table 5: Most published authors

No.	Author	As first or single author	As co-author	% of 2,167
1	Meyer JHF	11	6	.784
2	Kember D	9	5	.646
3	Richardson JTE	6	6	.554
4	Enders J	7	4	.508
	Prosser M	2	9	.508
5	Kyvik S	7	3	.461
6	Trigwell K	7	2	.415
	Marginson S	8	1	.415
	Watkins D	6	3	.415
7	Teichler U	7	1	.369
	Altbach PG	7	1	.369
	Harman G	7	1	.369
	Over R	6	2	.369

the terrain captured from the citation data from one key journal in the field. We found Meyer, Kember, Richardson as the most published authors, while Enders and Prosser share the fourth spot.

Further findings

Besides showing the geography of influence in two of the best higher education journals, SHE and HE, we also discovered something else. Our investigation showed that the lists of 'most cited' authors in Australian and UK/European journals were almost identical (see Table 6). In Table 6 also illustrate the considerable overlap between the findings in our two studies and those conducted by Kandlbinder and Tight.

Surprisingly, when comparing these studies with that of US journals we found that there was an international 'split' among 'most cited' authors. It appears there are regional differences when comparing US journals to

Table 6: Most cited authors comparison between US, UK and Australian journals

	US		UK	Australian		
	Budd (1990)	Budd and Magnuson (2010)	Tight (2008)	Kandlbinder (2012)	Calma and Davies (2015)	Calma and Davies (2017)
1	Astin	Astin	Clarke	Marton	Entwistle	Entwistle
2	Pascarella	Pascarella	Ramsden	Biggs	Marton	Clark,
3	Centra	Tinto	Becher	Ramsden	Ramsden	Marton
4	Baldrige	Kuh	Biggs	Entwistle	Biggs	Biggs
5	Feldman	Cabrara	Boud	Prosser	Becher	Ramsden

UK/European and Australian journals. US journals tend to exclusively cite US-based authors. They are parochial in their citation patterns.

Implications of this finding

Calver jokes that 'a narrowing of research diversity is as valuable to scholarship as atherosclerosis is to the cardiac patient, yet one follows citation hunger as surely as the other follows a fatty diet'.

We think this is a gloomy assessment of the situation. True, aiming for high citations as an end in itself is not necessarily a good thing for the reasons he mentions. However, there is another way to look at the enterprise of an academic seeking citations. The narrowing of research diversity also means the refinement and shaping of the entire enterprise of scholarship. To use a different metaphor, like the forces of nature of erosion by wind and water, citations effectively provide a useful map of the 'geography' of a discipline. The democracy of citation making ensures good ideas become of central importance, and less worthy ideas become marginalised or ignored altogether. This provides a measure of the history of a discipline – where it has been, and potentially where it is headed. They arm us with useful information about what academics find important. Citation maps also provide a concrete representation of key influences. If an idea – no matter how old – becomes influential, it has the potential to change the landscape of the discipline. To use Calver's analogy, narrowing of research interests does not 'clot' a discipline as much as shape its geography, free it from unwanted rubble, allow it to grow in productive directions.

Of course, it is true that important work can be infrequently cited, and contrariwise, frequent citations need not equate to influence or importance, as Calver notes (Calver, 2015, p. 47). Non-existent papers can also be cited through referencing errors, misinterpreted information can be promulgated, and citations can be 'inessential and perfunctory' (Calver, 2015). This is no doubt the case, especially in the short-term. In the main, however, these infelicities are usually addressed in the normal process of slow, methodical, self-correcting scholarship and attrition. Over longer time spans 'the truth will out' (we don't see citations anymore to phlogiston research, phrenology, luminiferous aether, or a host of other previously viable, and quite legitimate, research domains.) An adequate understanding of the worth of cited information, can only be conducted longitudinally surveying the entire landscape of a discipline – as we have tried to do here.

If an academic's work does not feature as part of the landscape of a discipline, it is unclear to what extent he or she is making a difference – or even whether they are working in areas that count. The key question to ask about citations, is not 'how many do I have?' but 'are my citations part of an important, and growing, terrain in my discipline?' It is only by analysing citation data that we can make this determination. Aiming for high citations is important.*

*NB: No unnecessary citations have been used in the production of this paper.

Martin Davies is Principal Fellow in Higher Education at the Melbourne Graduate School of Education, University of Melbourne, Australia and Senior Learning Advisor at Federation University, Australia

Contact: wmdavies@unimelb.edu.au

Angelito Calma is a Senior Lecturer at the Williams Centre for Learning Advancement, Faculty of Business and Economics, The University of Melbourne.

References

- Budd, J. M. (1990). Higher Education Literature: Characteristics of Citation Patterns. *Journal of Higher Education*, 61(1), 84-97.
- Budd, J. M., & Magnuson, L. (2010). Higher Education Literature Revisited: Citation Patterns Examined. *Research in Higher Education*, 51(3), 294-304.
- Calma, A., & Davies, M. (2017). Geographies of Influence: A Citation Network Analysis of Higher Education 1972-2014. *Scientometrics*, doi:10.1007/s11192-016-2228-3.
- Calma, A., & Davies, M. (2015). Studies in Higher Education 1976–2013: a retrospective using citation network analysis. *Studies in Higher Education*, 40(1), 4-21.
- Calver, M. (2015). Please don't aim for a highly cited paper. *Australian Universities' Review*, 57(1), 45-49.
- Kandlbinder, P. (2012). Recognition and Influence: The Evolution of Higher Education Research and Development. *Higher Education Research and Development*, 31(1), 5-13.
- Macfarlane, B. (2012). The Higher Education Research Archipelago. *Higher Education Research and Development*, 31(1), 129-131.
- Moosa, I. (2011). The demise of the ARC journal ranking scheme: An ex post analysis of the accounting and finance journals. *Accounting and Finance* 51(3), 809-836.
- Tight, M. (2008). Higher education research as tribe, territory and/or community: a co-citation analysis. *Higher Education*, 55(5), 593-605.
- Westbury, I., (1980). The Impact of the Journal of Curriculum Studies: A Citation Analysis. *Curriculum Studies*, 12(2) 149-156.