Journal of Hospitality and Tourism Management xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Hospitality and Tourism Management

journal homepage: www.elsevier.com/locate/jhtm

Research impact benchmarks for tourism, hospitality and events scholars in Australia and New Zealand

Pierre Benckendorff*, Mengya (Lavender) Shu

UQ Business School, The University of Queensland, St Lucia, Qld., 4072, Australia

ARTICLE INFO

Keywords:
Research performance
Impact
Bibliometrics
Citation metrics
h-index
g-index

ABSTRACT

This paper provides academic impact benchmarks for tourism, hospitality and event scholars located in Australian and New Zealand to assist with the reporting requirements required for grant applications, performance appraisals, career advancement, job applications and research assessments. The paper also presents research metrics for several other countries that produce tourism, hospitality and events research. The analysis is based on Google Scholar citation metrics calculated from two different datasets. The findings confirm significant differences between scholars across different academic levels, research fields, institutional contexts and higher education systems. The key contribution of this paper lies in aggregating all of the available data to present objective and valid benchmarks for comparing the academic impact of individual scholars. The paper also considers some of the risks and limitations of relying solely on citation metrics to measure academic impact. The measurement of research performance at an individual, institutional or national level can lead to individual behaviours that may distort the production of knowledge and epistemological development of the field. We conclude by calling for a wider discussion about the assessment of research performance and impact in tourism, hospitality and events.

1. Introduction

Scholarly research is the *raison d'être* for many scholars and higher education (HE) institutions. However, the neoliberal ideology that has dominated most Western political systems has resulted in increasing scrutiny of the research performance and competitiveness of institutions. This has in turn imposed a performance driven culture on scholars, who are increasingly expected to meet performance targets related to research productivity, quality and impact (Feller, 2009). This is evidenced by an increasing focus on measuring and benchmarking the research performance of institutions and individuals in the United Kingdom (Research Excellence Framework), Australia (Excellence in Research for Australia) and New Zealand (Performance Based Research Fund). Recent attention has shifted away from measuring research productivity to assessing research impact. However, attempts to measure impact are not particularly useful without discipline-specific performance benchmarks.

The purpose of this paper is to benchmark the research performance of tourism, hospitality and event scholars located in Australia and New Zealand (NZ). The paper extends previous evaluations of research performance in tourism, hospitality and events by providing objective and valid benchmarks of well-established research impact metrics across different academic levels, institutional contexts and fields. The paper also builds on a similar study of Australian marketing academics conducted in 2013 and updated in 2015 (Soutar, 2013; Soutar, Wilkinson, & Young, 2015). Citation metrics and benchmarks vary enormously between different disciplines and fields because they are a function of the size of an academic community, the number of sub-disciplines or fields, the productivity and nationality of scholars, citation patterns and other disciplinary characteristics (Kelly & Jennions, 2006). This paper is the first attempt to establish benchmarks specifically for the tourism, hospitality and events fields.

The paper makes several contributions that have implications for knowledge production in the field. The analysis fills a pragmatic need for research impact benchmarks in the tourism, hospitality and events fields. These benchmarks assist tourism, hospitality and event scholars with the reporting requirements attached to grant applications, performance appraisals, career advancement or job applications. Benchmarks are particularly important when presenting individual research performance metrics to assessors and review panels from other fields because they provide data about how individuals compare with their peers. The intent is not to rank individuals or institutions, but rather to aggregate bibliometric data to offer contextual insights into research performance. The paper also makes a methodological

* Corresponding author.
E-mail addresses: p.benckendorff@uq.edu.au (P. Benckendorff), mengya.shu@uqconnect.edu.au (M.L. Shu).

https://doi.org/10.1016/j.jhtm.2018.04.005

contribution by using two bibliometric data sources to explore research impact benchmarks. While the results will be of particular interest to scholars based in Australian and NZ, the use of a second data source provides additional insights into the research metrics of tourism, hospitality and events scholars in other countries. We also seek to contribute to the wider debate about the measurement of research performance by considering how an obsession with citation metrics can distort the epistemological development of the field.

2. Literature review

The bibliometrics field has developed a substantial literature on the measurement of research performance. It is clear from this literature that there are three broad measures of individual research performance: productivity, impact and quality. Productivity is typically measured by the number of publications produced by a scholar, irrespective of whether these contributions are in high quality outlets. Some scholars are prolific writers and collaborators and produce a large number of papers but these papers may be poor quality and may have limited impact. Quality refers to the significance and originality of a publication, the sophistication and robustness of the methodology and the efficacy of the conclusions. Quality is typically measured in academic settings by counting the number of papers published in top tier journals (i.e. journals that have high impact factors). Impact is closely related to quality but quality on its own does not guarantee impact.

The Australian Research Council (ARC) (2015, p.1) defines impact as: “the demonstrable contribution that research makes to the economy, society, culture, national security, public policy or services, health, the environment, or quality of life, beyond contributions to academia.” Despite this wider definition of impact, the ARC acknowledges that impact is likely to be underpinned by excellent research. It is clear that expectations about research performance are shifting from an emphasis on productivity to evidencing impact (Becken, Miller, & Banhalmi-Zakar, 2016). The ARC (2015) predicts that impact will become increasingly important under constrained funding conditions and returns on research investment will need to be demonstrated in terms of environmental, economic and social impact.

Research impact is typically divided into two categories: academic impact and economic/societal impact. This study is concerned with academic impact, which refers to the extent to which other researchers value published research in advancing knowledge, methods, theory and application. This focus is not intended to limit the debate about impact – it is acknowledged that other indicators of impact are equally valid and appropriate when assessing research performance. Academic impact can be assessed holistically and qualitatively but is typically measured using citation metrics. The focus on citation metrics allows for a more objective reporting of the bibliometric indicators that are increasingly used by institutions to evaluate and benchmark research performance. In this context, the ability for scholars to benchmark their research with others in their field becomes a useful pragmatic exercise for career advancement.

Citation analysis involves counting the number of times a paper is cited by other authors to measure the impact of a scholar. The advantage of citation metrics is that they are based on objective data that are easy to measure, obtain and analyse (Becken et al., 2016). A number of citation metrics have been developed over the last decade for the purpose of measuring academic performance. The most ubiquitous of these is the Hirsch index, or h-index, first proposed by Jorge Hirsch (2005). According to Hirsch (2005) a scholar has index h if h of his/her N_p papers have at least h citations each, and the other (N_p-h) papers have no more than h citations each. In more pragmatic terms, a scholar with a h-index of 10 has published 10 papers that have at least 10 citations.

The h-index has a number of advantages: (1) it combines productivity and impact, (2) the data used to calculate an individual's h-index is readily available, (3) it is not sensitive to extreme values, and (4) it is difficult to artificially inflate or manipulate (Batista, Campiteli, Kinouchi, & Martinez, 2006). While the h-index was originally conceived as a metric for measuring individual research performance it has subsequently been applied to measure the performance of research groups, institutions, journals, disciplines and fields. It is therefore not surprising that within a year of being proposed, the h-index was adopted by leading journals such as *Science* and *Nature* as a measure of individual research performance (Bornmann & Daniel, 2007).

While the h-index is widely used, and has been described as a robust measure of research performance, it does have some shortcomings. Hirsch's own work demonstrated that the h-index varied enormously across different disciplines and fields, highlighting the need for discipline specific benchmarks. Egghe (2006) argued that the h-index can mask the impact of a small number of very highly cited contributions amongst an otherwise lacklustre collection of contributions. Some of these contributions may be the ‘one hit wonders’ of the academic world, having enormous impact that can move an entire field forward. To correct for this problem, Egghe (2006) proposed the g-index, which gives more weight to highly-cited articles. A scholar would have an index of g if their g most cited articles summed to g^2. For example, if a scholar's most cited articles had a cumulative total of 100 citations their g-index would be 10.

The h-index has also been criticised for not taking into account the age of an article, thereby failing to differentiate between active and inactive scholars. An active researcher's h-index tends to increase over time and therefore reflects the number of years they have been actively publishing. This puts new scholars at a disadvantage and favours established researchers (Glänzel, 2006). Some scholars may be ‘resting on their laurels’ by relying on citations generated by significant articles published many decades ago. However, in an academic context it is useful to identify scholars who have had a sustained impact over a long period of time, as well as brilliant emerging scholars publishing ground breaking papers. To address this issue Sidiropoulos, Katsaros, and Manolopoulos (2007) proposed the contemporary h-index (h_c-index), which gives less weight to older articles. A final criticism of the h-index is that it is not sensitive to co-authored publications. Batista et al. (2006) proposed the individual h-index (h_I-index) to address this issue. The h_I-index divides the standard h-index by the average number of authors in the articles that contribute to the h-index, in order to reduce the effects of co-authorship. While many other citation metrics are now available, these four approaches are used most commonly when benchmarking the research performance of scholars.

In the tourism literature both the h-index and g-index have been used by McKercher (2008) to identify the most cited tourism scholars. Chang and McAleer (2012) included the h-index in their comparison of several research assessment measures for ISI tourism and hospitality journals. The h-index was also cited briefly by Airey, Tribe, Benckendorff, and Xiao (2015) to comment on the quality of tourism journals relative to other fields. A recent study by Becken et al. (2016) reported that more than 80% of academics identified citations as a useful measure of academic impact. Over 60% of researchers also identified the h-index as a useful measure of impact, although 20% indicated that they did not know what this was or were unsure whether it was useful. The same study also reported that despite the widespread use of these metrics there was an acknowledgement from many researchers that they were imperfect measures. While these studies all make useful, and at times controversial, contributions they do not provide tourism, hospitality and events researchers with useful benchmarks for self-evaluation.

Beyond the tourism literature, Soutar (2013) provides research impact benchmarks for level B (Lecturer), level C (Senior Lecturer), level D (Associate Professor) and level E (Professor) marketing academics in Australia. His study of h-indices and g-indices revealed significant differences between marketing scholars at different academic levels, with more senior scholars having greater impact. Soutar et al. (2015) expanded and updated this analysis by reporting the citation

metrics for over 2000 marketing scholars in the top 500 research universities in the Academic Ranking of World Universities (ARWU). They note that “these kinds of data can and should inform government assessments of universities and can also aid universities and academics in judging their research performance” (p. 158). The purpose of this paper is to build on this work by benchmarking the research performance of Australian and NZ tourism, hospitality and event scholars.

3. Methods

The data required for citation analysis can be accessed from three sources. The Institute for Scientific Information (ISI) database (also known as the ‘Web of Science’) was traditionally the only tool available for conducting citation analysis. However, the introduction of Elsevier's Scopus and Google Scholar (GS) has resulted in more comprehensive and multidisciplinary coverage (Bar-Ilan, 2008; Meho & Yang, 2007). GS was selected to measure the performance metrics reported in this study because it is considerably more comprehensive in its coverage of publications in the broader business and management research field in which tourism is usually situated (Harzing & van der Wal, 2009). For this reason, an individual's GS h-index will usually be higher than h-indices based on Web of Science or Scopus datasets. The results presented in this paper draw on two GS datasets.

The first dataset consisted of citation metrics collected for tourism, hospitality and events scholars based in Australia and NZ. Scopus was initially used to identify Australian and NZ-based researchers publishing in A* and A ranked tourism, hospitality and event management journals (based on the 2016 ABDC list). This initial list was expanded by visiting the websites of affiliated institutions to identify additional scholars. This second step was necessary for identifying less established academics who may not have published in the top tourism journals. The preliminary list included a number of scholars who had retired as well as international scholars who held honorary positions with Australian and NZ universities. These names were removed from the list since the focus of this dataset was on active academics based primarily in Australia and NZ. The preliminary list also included scholars from other disciplines who occasionally publish in the tourism literature. Several of these scholars were excluded from the final list after examining their publications and determining that tourism, hospitality or events was not a primary research focus. The final list consisted of 197 scholars from 32 institutions, including 35 scholars from NZ and 162 scholars from Australia.

The GS metrics for tourism, hospitality and events scholars based in Australia and NZ were collected in September and October 2016 using Harzing's (2007) *Publish or Perish* software. The metrics obtained for this study were the h-index, the g-index, the h_c-index and the h_I-index. The average number of authors per paper was also collected for each scholar to accompany the h_I-index. The name of each scholar was entered into the software to obtain these metrics. A great deal of effort was expended to ensure that the results were based only on contributions published by the scholar that was the focus of the search, rather than by other scholars with the same name. This was particularly time consuming for scholars with common names, since each publication needed to be screened to verify its inclusion in the metrics. For less common names the search took as little as five minutes but this process of disambiguation took over an hour for some authors with common surnames. These data were cross-checked with the GS Profiles of individuals where available. The metrics were calculated based on books, book chapters, journal articles and conference papers. Given that most tourism, hospitality and events scholars were not active prior to 1960 the search was limited to contributions published between 1960 and 2016. This differs from the approach taken by Soutar (2013), which only examined papers published between 2001 and 2013 to reduce age effects. As noted earlier, the inclusion of the h_c-index in this study presents an alternative metric that controls for the age of contributions.

The academic level of each scholar was identified by visiting their institutional profile page. In cases where this information was not available other online sources, such as Google Scholar Profiles, LinkedIn and ResearchGate were consulted. Following the Australian academic classification system, post-doctoral researchers were grouped with associate lecturers (level A), research fellows were grouped with lecturers (level B), senior research fellows were grouped with senior lecturers (level C), and principal research fellows were grouped with associate professors (level D). Associate lecturers were underrepresented in the sample, partly because these academics are at the start of their career and were difficult to identify but in general there are also fewer appointments at this level.

The second dataset was based on data collected from the GS profiles of tourism, hospitality and events academics worldwide. A search was conducted for all GS profiles that included ‘tourism’, ‘hospitality’, ‘events’, ‘tourist’, ‘travel’, ‘visitor’, ‘destination’, ‘hotel’ and ‘hotels’ as areas of interest. An automated script was used to capture data from each GS profile, including the name, position, institutional affiliation and citation metrics for each scholar. Due to time constraints, the search was limited to GS profiles with more than 200 citations. The initial search produced a list of 500 GS profiles that met these criteria.

A GS profile must be manually created and curated by the profile owner. This impacts on the quality of the data in five ways. First, many tourism, hospitality and events scholars may not have a public GS profile and are therefore not included in the dataset. Second, GS Profiles are more likely to be created by established scholars who already have a number of publications to curate. Early career researchers may be reluctant to create a public profile if they only have a few publications or a small number of citations. As a result, it is likely that any bibliometrics harvested from GS Profiles will be skewed toward more senior scholars with more papers and citations. Third, tourism, hospitality and events academics who do have a profile will be excluded if the profile did not list one of the areas of interest used for the search. Fourth, some researchers do not indicate their position or institutional affiliation on their profile. Finally, publications are assigned to individual profiles automatically and require some curation by profile owners to merge duplicates or to remove incorrect entries. To overcome some of these limitations, the initial dataset was screened to eliminate duplicate entries and GS profiles that clearly did not belong to tourism, hospitality or events researchers. For example, a number of profiles listed ‘travel’ as an area of interest but in most cases this was not a research interest. The data were also scanned for inconsistencies, such as unusually high citation metrics and suspicious profiles were checked individually. Manual Internet searches were then conducted to complete remaining profiles with missing data. This was a laborious process, but in most cases the missing fields were able to be captured by scanning the institutional pages or ResearchGate profiles of individual researchers. The final dataset included 388 tourism, hospitality and events scholars from 47 countries, including 69 scholars from Australia and New Zealand who were also included in the first dataset. Almost 60% of these profiles came from the USA (n = 67), Australia (n = 54), Spain (n = 48), UK (n = 46) and New Zealand (n = 15).

It was necessary to align the academic levels in different countries with Australian and NZ positions so that metrics could be compared across different levels of seniority. Australia and New Zealand, like most Commonwealth countries, have traditionally adopted the UK system of academic ranks (see Table 1). Countries in North America and Asia typically follow the US system of academic ranks. Many European countries (notably France, Germany and Spain) have their own classification systems. While every effort was made to align positions as shown in Table 1, considerable differences remain. For example, in some countries, the title of Professor is more widely used and assigned to all scholars regardless of seniority or academic attainment. In Germany, Austria and some other Germanic countries the title is used to denote scholars who have completed a *Habilitation* by writing a second thesis. These international differences in academic nomenclature create a number of challenges and limitations when attempting to benchmark

Table 1
Alignment of academic levels.

Australia/NZ	United Kingdom	North America	Spain
Associate Lecturer	Assistant/Associate Lecturer	Teaching Assistant	Profesor Ayudante
Lecturer	Lecturer	Assistant Professor	Profesor Colaborador Profesor Ayudante Doctor
Senior Lecturer	Senior/Principal Lecturer	Associate Professor	–
Associate Professor	Reader/Associate Professor	Professor	Profesor Contratado Doctor Profesor de Universidad Privada
Professor	Professor	Full Professor (Chair)	Catedrático de Universidad Profesor Titular de Universidad

Table 2
Publish or Perish metrics for tourism, hospitality & events researchers.

	Count	h-index			g-index	h_c-index	h_I-index	Authors/Paper
		Median	Min	Max	Median	Median	Mean	Mean
Academic Level								
Associate Lecturer	7	1	0	14	2	2	1.34	1.53
Lecturer	51	5	0	13	9	5	2.00	2.62
Senior Lecturer	53	9	2	18	18	9	3.95	2.76
Associate Professor	48	15	3	34	27	12	6.48	2.84
Professor	38	27	4	97	58	23	13.31	2.77
All Levels	197	10	0	97	21	9	5.78	2.70
Country								
Australia	162	10	0	48	20	9	5.37	2.75
NZ	35	12	2	97	25	10	7.64	2.48
Region								
Queensland	73	10	0	48	20	9	6.06	2.79
New South Wales	37	9	0	34	19	8	4.56	2.72
Victoria	23	7	1	34	16	8	4.66	2.74
Western Australia	18	13	0	31	26	12	5.49	2.78
South Australia	10	7	1	35	13	9	4.58	2.32
South Island	19	12	2	97	21	12	7.82	2.63
North Island	16	12	3	59	25	9	7.43	2.31
Group								
Go8	28	16	0	43	30	13	6.39	2.49
IRU	50	10	0	48	18	9	6.44	2.84
ATN	27	10	1	35	25	10	5.34	2.65
RUN	22	9	0	33	17	8	4.74	2.79
Independent	28	7	2	24	16	7	3.56	2.82
TVET	7	6	1	31	10	8	2.99	3.06
NZ	35	12	2	97	25	10	7.64	2.48
ERA2015 Rank								
Not Assessed	32	7	0	33	15	7	3.77	2.82
2	19	10	3	34	20	8	4.98	2.84
3	54	11	0	48	24	10	5.89	2.61
4	36	8	0	46	16	9	5.34	2.90
5	21	15	1	43	30	14	6.91	2.63
Field								
Events	14	5	0	17	9	5	2.76	2.48
Hospitality	31	9	0	38	17	8	4.14	2.84
Tourism	152	11	0	97	23	10	6.39	2.69

University Groupings: Go8 = Group of Eight, IRU = Innovative Research Universities, ATN = Australian Technology Network, RUN = Regional Universities Network, Independent = Non-affiliated universities, TVET = Technical and Vocational Education and Training, NZ = NZ Universities. **ERA2015 Rankings:** 1 = Well below world standard, 2 = Below world standard, 3 = At world standard, 4 = Above world standard, 5 = Well above world standard.

performance across different countries.

4. Results and discussion

The first part of the results and discussion are focussed on the more comprehensive dataset consisting of 197 tourism, hospitality and events scholars based in Australian and NZ. Table 2 provides an overview of the various metrics based on academic level, country, region, university grouping and primary field of research. The table also includes a breakdown of the metrics based on the ERA2015 tourism field of research (1506) rankings for Australian institutions. Medians are presented for the h-index, g-index and h_c-index because the data were skewed and included a number of outliers. The calculation of the h_I-index and mean authors per paper is less likely to produce outliers and the means are presented as a more accurate measure. Metrics for individual scholars and institutions are not reported because this paper is not concerned with rankings.

The results indicate a clear pattern, with more senior scholars having higher indices across all of the metrics. The h_c-index and h_I-index indicates that even when controlling for the age of papers and co-authorship, more senior scholars still have greater impact. An Oneway ANOVA confirmed that these differences were significant ($p < 0.001$) for all metrics except papers/author. At the time of writing the highest h-index (97) was recorded for a Professor based in NZ. The low minimums for each level may indicate the 'long tail' of low performing scholars (Airey et al., 2015), but these metrics may also represent

Table 3
Publish or Perish h-index percentiles by academic level.

	5th Percentile	25th Percentile	75th Percentile	95th Percentile	99th Percentile
Associate Lecturer	0	0	4	–	–
Lecturer	0	3	6	10	13
Senior Lecturer	3	7	13	17	18
Associate Professor	4	11	18	23	34
Professor	10	21	35	59	97
All Levels	1	5	17	35	59

individuals at the start of their career (for levels A or B) or scholars who may hold teaching focussed or administrative appointments. Due to differences in data collection, these metrics are not directly comparable with Soutar (2013) work in marketing but the results do confirm the same patterns. It is clear that research performance does differ by academic level and that the benchmarks used by scholars to evaluate their own performance should be different for each academic level. To further explore these differences Table 3 presents the h-indices at various percentiles.

These results provide a number of benchmarks that can be used by tourism, hospitality and events scholars in Australia and NZ. For example, a lecturer (level B) with an h-index of 13 or higher would be in the top 1% of scholars in at this level but would also be in the top 25% of scholars who are at the senior lecturer level. Four level B academics had an h-index between 10 and 13. Similarly, a senior lecturer (level C) academic with an h-index of 17 or higher would be in the top 1% of scholars at that level but also close to the top 25% of scholars at the associate professor level. Again, there were four scholars from various universities with h-indices of 17 or 18. The pattern is repeated at level D, with two academics having h-indices of 34 but there was a much larger gap between the 95th and 99th percentiles. These benchmarks provide some indication of the level of research performance required for promotion from one level to the next.

While these metrics are useful for benchmarking, it is also likely that regional differences exist due to differences in funding opportunities and resources. The country and region/state for each scholar was recorded to explore this proposition. The results in Table 2 indicate that while NZ scholars perform slightly better than Australian scholars, a paired-samples *t*-test confirmed that the differences were not significant ($p < 0.05$). Furthermore, it appears that the extent of co-authorship is lower in NZ and this results in a higher h_I-index for NZ scholars. The regional comparison also shows the same pattern of results with no significant differences ($p < 0.001$), although it is interesting to note that two thirds of Australian tourism, hospitality and events scholars in the sample were located at Queensland institutions.

The next set of metrics in Table 2 explore whether differences exist between different types of institutions. Anecdotal observations would suggest that differences in teaching loads and funding for research might impact on research performance. A number of university groupings exist in Australia and these groupings are generally indicative of the history, style and focus of the institutions in each group. The NZ metrics are again included here for comparison. The metrics indicate that scholars located in Group of Eight institutions (primarily The University of Queensland, Monash, UNSW and UWA) perform better than colleagues in other institutions, particularly non-affiliated institutions and TVET institutions. However, it is worth mentioning that the top 10 scholars were Professors based at six different universities, with only two based at a Group of Eight institution. While the differences were not significant ($p < 0.05$) these results confirm what some academics have long suspected and highlight disparities in teaching loads and the resourcing of research. It is also useful for individual benchmarking purposes to explore how these institutional differences differ by academic level. Table 4 presents the median h-index by academic level and university grouping. Associate lecturers are excluded from this analysis due to the low sample size.

Table 4
Publish or Perish median h-index by academic level and university grouping.

	Go8	IRU	ATN	RUN	Independent	NZ
Lecturer	7	5	4	5	6	4
Senior Lecturer	13	8	8	9	5	10
Associate Professor	19	15	14	16	11	13
Professor	30	36	23	23	13	26

The results indicate a pattern that is consistent with what has already been observed, although it is interesting to note that regional universities (RUN) appear to harbour younger scholars who perform quite well at the lecturer, senior lecturer and associate professor levels. This advantage disappears at the professorial level, presumably because talented scholars move to institutions with more favourable research environments when they reach this level.

In Australia, research performance is measured at the institutional level through the Excellence in Research for Australia (ERA) assessment system. The last assessment was conducted in 2015. While many of the science disciplines used citation metrics to benchmark research performance the assessments for most of the social sciences (including tourism) were based on peer review. While ERA also considers other indicators, such as grant success and the esteem of scholars, it is useful to explore whether there is some alignment between ERA outcomes and the metrics reported in this paper. It might be expected that institutions with higher ERA rankings employ scholars with higher research performance metrics. The results show that while scholars at Australia's top ranked institution perform better across all metrics this pattern is not repeated for the institution that received a ranking of four. Careful interpretation is required in this case because the institution receiving a four is the largest employer of tourism, hospitality and events scholars in the sample. The staffing profile at this institution is skewed toward early career researchers (ECRs) and teaching focussed scholars who may not be included in an ERA assessment. On the other hand, the same institution also has a larger number of top professors and research only appointments than any other institution. These observations highlight the limitations of taking citation metrics at face value without considering institutional differences. Aside from this anomaly the metrics generally follow the expected pattern but statistically significant differences ($p < 0.05$) were only evident for the h_c-index.

The last set of metrics presented in Table 2 examined differences based on scholars' field of research. While some scholars publish across multiple fields, an attempt was made to categorise each scholar into tourism, hospitality or events based on the area reflected by a majority of their publications. An Oneway ANOVA indicated that significant differences existed between the three fields ($p < 0.05$) across all of the metrics except mean authors per paper. Generally, the metrics for tourism scholars were higher than for hospitality and events. This is likely to reflect the size and maturity of the research communities in each of these fields. *Ceteris paribus,* larger research communities usually generate more citations. These benchmarks highlight important differences between fields that need to be considered by academics as well as panels responsible for assessing research performance for new appointments and career advancement.

Although these results provide useful benchmarks of tourism, hospitality and events scholars based in Australia and New Zealand, benchmarking these results against a more global sample provides additional insights. Table 5 presents the median h-index of 388 scholars at various academic levels based on the second dataset collected from the Google Scholar profiles.

Although the second dataset included 69 scholars from Australia

Table 5
Google Scholar Profile h-indices by academic level.

	Count	Median	Min	Max
Lecturer	59	9	2	17
Senior Lecturer	65	10	6	29
Associate Professor	84	13	6	37
Professor	180	17	4	97

and New Zealand, the additional data from other countries has resulted in median scores that are higher for more junior academic levels (i.e. Lecturers and Senior Lecturers) but lower for Associate Professors and Professors. The higher medians for Lecturers and Senior Lecturers are most likely an artefact of our decision to collect data only for scholars who had received more than 200 citations in total. Associate Lecturers were not represented in this analysis because only two scholars at this level had more than 200 citations. When we exclude Australian and NZ scholars with less than 200 citations from our first dataset the median h-indices for lecturers (median = 8) and senior lecturers (median = 11) are similar to the international dataset presented in Table 5. On the other hand, the lower medians for more senior scholars may be due to a lack of alignment between Australian and NZ academic ranks and academic systems in other countries, particularly North American, Asian, Spanish and Germanic countries, where the position of 'professor' has a different meaning.

Notwithstanding these limitations, the results confirm that median h-indices increase with academic seniority and a Oneway ANOVA indicated significant differences in h-indices between these academic levels ($p < 0.05$). This on its own is hardly a startling finding, however the analysis also highlights the discrepancies that may exist between different national education systems and research environments. To explore these differences, the median h-indices were calculated for countries with more than ten scholars (see Table 6). These countries represented 68% of the sample.

Once again, when we exclude scholars with less than 200 citations from our first dataset in Table 2, the median h-indices for Australians (median = 13) and New Zealanders (median = 13) are very similar to the international dataset presented in Table 6. The results for other countries are insightful. An Oneway ANOVA indicated that there were significant differences between countries ($p < 0.05$). Countries with national research assessments (e.g. NZ, Australia, UK) or tenure requirements (e.g. USA, Canada) tend to exhibit higher median h-indices. It is also interesting to note that countries where English is not the dominant language had lower medians. Moving beyond the data in the table, Hong Kong, China, Turkey, Brazil, Germany and other European and Scandinavian countries are conspicuously absent despite the fact that many tourism, hospitality and events scholars are based in these countries.

5. Conclusions and implications

The results have several important implications for career planning and the measurement of individual research performance in the future.

Table 6
Google Scholar Profile h-indices for countries with more than 10 scholars.

	Count	Median	Min	Max
USA	67	15	6	72
Australia	55	14	6	48
Spain	48	12	7	24
UK	46	14	4	60
New Zealand	15	13	7	97
Canada	12	14	4	51
Portugal	12	9	6	14
Malaysia	10	10	7	11

The results provide objective and valid benchmarks for comparing the research performance of tourism, hospitality and event scholars across different academic levels, institutional contexts and fields. The metrics highlight that important differences exist and that a single average h-index or g-index for the entire academic community is unlikely to be a useful benchmark. Instead, the analysis presents benchmarks for a range of academic levels and contexts to assist scholars with the reporting requirements attached to grant applications, performance appraisals, career advancement or job applications. Using these benchmarks, an individual researcher is able to compare their own h-index with the appropriate level and context to make a case for promotion or grant funding. The pragmatic value of this paper lies in aggregating all of the available data – a time consuming task requiring in excess of 90 h of data collection and analysis. While these data are publically available the time and effort required to conduct a comprehensive analysis of all scholars in the field is impractical for individuals preparing applications for grants or career advancement.

The international comparison highlights some of the dangers in comparing research performance across different countries. Differences in national education systems, resources, workload and performance expectations make international benchmarking challenging. These differences are important to consider in the context of an increasingly mobile academic workforce. The results highlight a need for university recruiters and selection panels to exercise caution when interpreting and comparing the research performance of applicants from other education systems. The GS Profile data also raises questions about whether the monitoring and reporting of research metrics in countries that are well represented may motivate scholars to create GS Profiles to improve their online research visibility in an effort to attract further citations. These observations lead us to conclude that the measurement of research performance at an individual, institutional or national level can lead to individual behaviours that may distort the production of knowledge and epistemological development of the field.

Some caveats and limitations of the research also need to be acknowledged. Firstly, the data were collected in late 2016 and since the academic community and citations are constantly growing the benchmarks are time sensitive. In other words, if this analysis were repeated in 2020 it is likely that the benchmarks will be higher. Secondly, the analysis was largely focussed on scholars based in Australia and NZ. Some effort was made to compare these benchmarks with scholars working in other countries but the findings highlighted a number of challenges and a need for caution. Thirdly, the results highlight the risks of applying neoliberal approaches to manage research performance through national research assessments that focus exclusively on citation metrics without considering academic mobility or differences in national education systems, institutional contexts and resourcing.

As acknowledged earlier, there are many other quantitative and qualitative indicators of research performance. Indeed, the citation metrics used in this paper have been criticised because they are imperfect measures that do not account for the different audiences scholars may be targeting. A key argument advanced by some scholars is that research should facilitate change and innovation and that these outcomes can be difficult to measure using citation metrics (Smith, Crookes, & Crookes, 2013). As Harzing and van der Wal (2009) note, reducing an entire lifetime of work to a single metric is unlikely to provide a complete picture of a scholar's real impact. For this reason, we would not recommend the use of these metrics as the only measure of an individual scholar's research impact. Scholars seeking advancement would be well advised to develop a research portfolio that includes citation metrics and benchmarks alongside other measures of impact such as research-informed teaching, awards, grants, media commentary, membership of editorial boards and scientific academies, invitations for keynote presentations or panels, social media and Internet mentions, and other measures of esteem. Evidence of wider economic and societal impacts could be presented in the form of evidence-based impact narratives or statements that demonstrate the

significance of research outcomes beyond scholarly publications. These conclusions highlight that it may be timely for the academic community in tourism, hospitality and events to initiate a wider discussion and debate about the most appropriate methods for assessing research performance and impact in our fields.

References

Airey, D., Tribe, J., Benckendorff, P., & Xiao, H. (2015). The managerial gaze the long tail of tourism education and research. *Journal of Travel Research, 54*(2), 139–151.

Australian Research Council (2015). *Research impact principles and framework.* Retrieved October 30, 2016, from http://www.arc.gov.au/research-impact-principles-and-framework.

Bar-Ilan, J. (2008). Which h-index? A comparison of WoS, Scopus and Google scholar. *Scientometrics, 74*(2), 257–271.

Batista, P. D., Campiteli, M. G., Kinouchi, O., & Martinez, A. S. (2006). Is it possible to compare researchers with different scientific interests? *Scientometrics, 68*(1), 179–189.

Becken, S., Miller, G., & Banhalmi-Zakar, Z. (2016). *Making an impact – practices and views of tourism academics.* Griffith Institute for Tourism Research Report No 13Queensland, Australia: Griffith University.

Bornmann, L., & Daniel, H.-D. (2007). What do we know about the h index? *Journal of the Association for Information Science and Technology, 58*(9), 1381–1385.

Chang, C. L., & McAleer, M. (2012). Citations and impact of ISI tourism and hospitality journals. *Tourism Management Perspectives, 1*, 2–8.

Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics, 69*(1), 131–152.

Feller, I. (2009). Performance measurement and the governance of American academic science. *Minerva, 47*(3), 323–344.

Glänzel, W. (2006). On the opportunities and limitations of the h-index. *Science Focus, 1*(1), 10–11.

Harzing, A. W. (2007). *Publish or perish.* Retrieved October 30, 2016, from http://www.harzing.com/pop.htm.

Harzing, A. W., & van der Wal, R. (2009). A Google Scholar h-index for journals: An alternative metric to measure journal impact in economics and business. *Journal of the American Society for Information Science and Technology, 60*(1), 41–46.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the national academy of sciences of the United States of America* (pp. 16569–16572). arxiv:physics/0508025.

Kelly, C. D., & Jennions, M. D. (2006). The h index and career assessment by numbers. *Trends in Ecology & Evolution, 21*(4), 167–170.

McKercher, B. (2008). A citation analysis of tourism scholars. *Tourism Management, 29*(6), 1226–1232.

Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology, 58*(13), 2105–2125.

Sidiropoulos, A., Katsaros, D., & Manolopoulos, Y. (2007). Generalized Hirsch h-index for disclosing latent facts in citation networks. *Scientometrics, 72*(2), 253–280.

Smith, K. M., Crookes, E., & Crookes, P. A. (2013). Measuring research 'impact' for academic promotion: Issues from the literature. *Journal of Higher Education Policy and Management, 35*(4), 410–420.

Soutar, G. N. (2013). Australian marketing scholars' research impact: An update and extension. *ANZMAC annual conference, Auckland, NZ*.

Soutar, G. N., Wilkinson, I., & Young, L. (2015). Research performance of marketing academics and departments: An international comparison. *Australasian Marketing Journal, 23*(2), 155–161.