

2013

# Measuring research 'impact' for academic promotion: issues from the literature

Kylie M. Smith

*University of Wollongong, kyliesm@uow.edu.au*

Ellie Crookes

*University of Wollongong, ec945@uowmail.edu.au*

Patrick A. Crookes

*University of Wollongong, pcrookes@uow.edu.au*

---

## Publication Details

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.

---

# Measuring research 'impact' for academic promotion: issues from the literature

## **Abstract**

Around the world, governments and the higher education sector are being asked to become more accountable for the money they spend on research funding. Research quality measurement exercises, such as the Excellence in Research for Australia initiative, use a number of agreed indicators to measure, analyse and report on various institution and discipline-based research outputs. This emphasis on the outputs of research as opposed to its longer term outcomes is having an effect on internal university policy and processes which can often operate negatively on individual staff career development and promotion. This article reports on a project aimed at more clearly articulating and defining the idea of research impact for academics by preparing a promotion application. Phase one of the project was an extensive international literature review and this article sets out the findings from this review, considers the difficulties for articulating and evidencing impact at the individual level and makes some suggestions for how academic staff and units might begin to deal with the idea of research impact.

## **Disciplines**

Medicine and Health Sciences | Social and Behavioral Sciences

## **Publication Details**

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.

## Measuring Research 'Impact' for Academic Promotion: Issues from the Literature

### **Authors:**

Dr Kylie M Smith  
Faculty of Health & Behavioural Sciences  
University of Wollongong NSW 2522  
Tel: +61 2 4221 4278  
Email: kyliesm@uow.edu.au

Ms Ellie Crookes  
Faculty of Health & Behavioural Sciences  
University of Wollongong NSW 2522  
Tel: +61 2 4221 4278  
Email: ecrookes@uow.edu.au

Professor Patrick Crookes  
Faculty of Health & Behavioural Sciences  
University of Wollongong NSW 2522  
Visiting Professor  
Department of Health Studies, Faculty of Social Sciences, University of Stavanger  
and School of Human and Health Sciences, University of Huddersfield  
Tel: +61 2 4221 4208  
Email: prcookes@uow.edu.au

### **Affiliation:**

Faculty of Health & Behavioural Sciences, University of Wollongong; Department of Health Studies, Faculty of Social Sciences, University of Stavanger and School of Human and Health Sciences, University of Huddersfield

### **Abstract:**

Around the world, governments and the higher education sector are being asked to become more accountable for the money they spend on research funding. Research quality measurement exercises such as the Excellence in Research for Australia initiative use a number of agreed indicators to measure, analyse and report on various institution and discipline based research outputs. This emphasis on the outputs of research as opposed to its longer term outcomes is having an effect on internal university policy and processes which can often operate negatively on individual staff career development and promotion. This paper reports on a project aimed at more clearly articulating and defining the idea of research impact for academics preparing a promotion application. Phase one of the project was an extensive international literature review and this article sets out the findings from this review, considers the difficulties for articulating and evidencing impact at the individual level and makes some suggestions for how academic staff and units might begin to deal with the idea of research impact.

Key words: research quality, impact, bibliometrics, Excellence in Research for Australia, ERA

## **Introduction**

It has been argued that research quality measurement exercises such as Excellence in Research for Australia (ERA) are problematic and limited in their understanding of what constitutes quality research (Bennett et al., 2011; Butler, 2003; Cooper et al., 2011; Genoni et al., 2009; Haddow et al., 2010; Jarwal et al., 2009; Jasco, 2010). This is not least because ERA primarily uses quantitative bibliometrics that measure and define research outputs as numbers of academic publications and not as the longer term health, social or economic benefits that research might bring. This focus on particular kinds of academic outputs raises issues of fairness and transparency across disciplines, and doesn't account for the different audiences or targets at which much research is aimed. It also perpetuates biases against certain disciplines which have struggled to be taken seriously, causing problems for academics in those areas who produce quality work of high impact (Wilson, 2011).

Many academics are concerned with how their work changes lives, improves health, or brings increased stability or sustainability beyond the world of the academic journal. Some attempts are being made to address this concern, with the Australian Technology Network of universities looking to develop an alternative measuring system to ERA which would account for research resulting in products, services or innovations that are designed to address specific industry or social problems (Duryea et al., 2007). The rationale for this is that some kinds of research are designed to effect change, and sometimes that change should be measured in outcomes other than journal publications or grant income. However, a problem remains about how these kinds of outcomes would be measured. What constitutes valid evidence of impact beyond publication bibliometrics remains a contested issue.

This is a concern that has become more pressing due to the diversifying nature of the academic workforce (Coates et al., 2010). Universities are making policy and procedural changes in order to deal with the demands of a diverse workforce, many members of which feel disadvantaged by existing academic career development and promotion systems which appear to reward traditional research outputs over other kinds of outcomes or impact. If individuals are appointed to undertake specific kinds of work, or within specific disciplines that have different academic audiences, then these staff are still entitled to have excellence in their work recognised, and be offered a genuine career pathway on the basis of that work. For many of these staff, traditional promotion criteria do not overtly recognise the kinds of outcomes that their research produces, and while a promotions committee may be able to deal with these issues on a case by case basis, clearer articulation of outcomes related to impact are needed to facilitate career development, provide transparency across an institution and encourage staff to undertake this important work.

As part of a project looking at promotion criteria more broadly across all academic spheres and levels, researchers at the University of Wollongong identified this gap around identifying 'impact' in promotion documents and processes. In order to ensure best practice and learn from what others have done in this area, we undertook a major review of the existing literature. We found evidence indicating that bibliometric measures of impact are problematic and biased, as well as a substantial body of work lamenting the need for impact to be considered within research quality measures. We were also particularly interested in possible frameworks for the measurement and reporting of impact that could be adapted to our purposes. However, we discovered a significant gap in the literature, which was the inability of current frameworks to account for individual impact. This article sets out the findings of our literature review and considers the implications of the debates around

'impact' for the development of academic promotion documentation and processes within the prevailing higher education context.

### ***Search strategy***

The article search for the literature review was conducted using the databases of CINAHL, Scopus, ProQuest, Web of Science and Science Direct, as well as the search engine Google Scholar, using the keywords and phrases; 'beyond impact factors,' 'assessing impact,' 'academic reputations,' 'assessing academic esteem,' 'new ways of assessing impact,' 'beyond citation analysis,' 'beyond bibliometrics,' and 'beyond citation impact.' Further articles were then sourced from bibliographies of articles downloaded. From this search system, 128 articles were found. This number was reduced to 71 after careful reading and analysis of the content, culling all articles that did not focus specifically on the issue of impact in relation to current evaluative systems, or new ways of measuring scholarly impact.

The remaining seventy-one articles provide insight into the reasons behind, and the problems with the employment of current assessment systems, by suggesting new ways to measure impact, and highlighting ideas for the development of evaluative tools which use impact as an indicator of academic excellence. It is around these main themes that we have structured our literature review.

### **Current research assessment systems**

Bibliometric indicators, in particular journal impact factors, citation rates and H-Indexes, are currently used as the primary measures of academic excellence in Australia and across the international academic community. Bibliometrics do have the potential to reveal data that is reflective of some measure of scholarly impact in terms of citation rates; however, they are not indicators of excellence. Being focused solely on *amounts* of published outputs and citations, they do not, by themselves, provide an accurate measurement of a scholar's overall academic impact or quality. Yet, it appears as though they are increasingly being used in this way. Since its initial trial in 2009, the Australian Research Council (ARC) has been using the Excellence in Research for Australia initiative (ERA), which employs predominantly bibliometric measures to assess quality of the research outputs of Australian Universities. It is largely due to ERA, and its emphasis on impact factors and the citations rates of journals and articles, that bibliometric measures have become the principal way that academic accomplishment is measured in Australia (Bennett, et al., 2011). Yet the literature reveals some serious concerns with the development and use of bibliometric measures for research quality (Cooper, et al., 2011; Martin, 2011).

Of the seventy-one articles examined in the literature review, eleven of them provide background information about the instigation and utilisation of bibliometric exercises for the measurement and assessment of research and researchers (Al-Awqati, 2007; Crookes et al., 2010; Fava et al., 2000; Genoni, et al., 2009; Harzing et al., 2009; Jackson et al., 2009; King, 1987; Kostoff, 1997; Ortner, 2010; Smith, 2001). Al-Awqati and King both argue that it was due to the unreliability of the methods previous to bibliometrics, such as relying purely on peer review and the reputation of academics, that bibliometric indicators flourished (Al-Awqati, 2007; King, 1987). Prior to bibliometrics, academic impact was measured by a system that lent itself to subjectivity and bias, and thus 'one's reputation simply depended on the words of others' (Al-Awqati, 2007, p. 183). Al-Awqati also criticises the peer-review system, arguing that it is flawed due to the partiality of peers, 'the old boy network' and the 'halo effect' which hindered upward mobility for new researchers and certain research fields,

and thus impeded the equal attainment of academic acclaim and reward for individual academics, faculties and universities.

It was in reaction to these short-comings that academic institutions instigated bibliometric tools. These new bibliometric methods were designed to establish international benchmarks for 'impact' and 'quality,' and gauge the scientific value of the journal or researcher more easily, objectively and accurately (Fava, et al., 2000; Genoni, et al., 2009).

In the late twentieth century, there was a reduction in University research funding globally due to an impetus to only fund research that would have outcomes that affected the greater population, (Buxton et al., 2004; Moodie, 2006). This shift had an impact on the introduction of bibliometrics for assessment and measurement as, according to King 'reductions in research budgets... [have] led to the need for greater selectivity in research allocation' (1987, p. 261). Bibliometric methods were believed to be easily comparable and quantifiable and thus it was thought that they could quickly determine which individuals or groups 'deserved' to be funded (Smith, 2001). Use of the impact factor, citations, rankings and the h-index became prevalent, along with the belief that only 'quantitative indicators [are] applicable, meaningful and useful in the assessment process' (Kostoff, 1997, p. 110).

Bibliometric measures are now commonly used for staff and institution evaluation, reward and assessment and funding delegation (Bevan, 2004; Bordons et al., 2002; Favaloro, 2009; Genoni, et al., 2009; Hendrix, 2010; Kurmis, 2003). It has been argued that bibliometric measures are 'Increasingly, not only [used] amongst the bibliometric community, but amongst researchers and science policy makers' (Bordons, et al., 2002, p. 195), who employ bibliometrics to 'assess the research quality and productivity of their faculty and staff' (Hendrix, 2010, p. 183) and to identify the 'institutional strengths, collaboration among faculty, emerging areas of research, benchmarks and budget priorities' (Hendrix, 2010, p. 184) of an institution. Bibliometrics are currently being used to determine 'staff appointments, allocation of staff promotions and tenure' (Kurmis, 2003, p. 2449) and have also been used as the basis for continued funding of research institutes, groups or individuals (Bevan, 2004), which has resulted in 'research funding [being] concentrated on a small number of institutions and research units' (Genoni, et al., 2009, p. 7). The practice of restricting the assessment and evaluation of academic performance to bibliometric indicators alone, so as to produce a supposedly objective measurement of the quality of an institution, a scholar, a journal or an article, is flawed and potentially damaging for the equity of academic reward and evaluation. There are a number of well-known issues surrounding the use of such measures which make this reliance on them potentially problematic.

### ***The impact factor***

Internationally, the predominant bibliometric tool, (which also forms the basis of the assessment of publications in ERA), is the journal Impact Factor (IF). According to Garfield, the impact factor was created by Irving H. Sher and Eugene Garfield in the early 1960s and was developed as a system for the ISI (Institute for Scientific Information) to help select journals for the Science Citation Index (1999) and 'enable researchers and librarians to map the networks of journals' (Crookes, et al., 2010, p. 420). In no way was the impact factor created as an indicator of quality or academic impact, however, according to Adam 'the ISI journal citation report's impact factor has moved in recent years from an obscure bibliometric indicator to become the chief quantitative measure of a journal, its research papers, the researchers who wrote those papers, and even the institution they work in' (2002, p. 729).

Given this, much of the literature argues that bibliometric indicators such as the impact factor have been used to evaluate journals and articles, measure quality and impact, and assess staff and institutions, to a degree that is inappropriate. Garfield, one of the founders of the Impact Factor, states that although bibliometric tools can be used effectively for ranking articles and journals, they have also been ‘inappropriately used as surrogates in evaluation exercises’ (1996, p. 3). The academic community has seemingly forgotten the important fact that when it comes to research, ‘what counts should be its actual merits, not its impact factor’ (Zavos et al., 2006, p. 1034). Garfield states that ‘the relationship between quality and impact is not absolute’ (1996, p. 3) and so the popularity of using IFs as an ‘index of quality and academic success’ (Bevan, 2004, p. 65) is something that could be potentially damaging to academic assessment and evaluation. Journal impact factors are based more on technicalities than academic quality, and according to Kostoff they ‘indicate quantity of output, not quality’ (1997, p. 113). There have been no studies into the validity of the impact factor as a sound indicator of quality, and so the positioning of IFs as indicators of quality or impact by institutions around the globe, is based on little more than the convenience and the easy comparability of quantitative bibliometrics, rather than proof that the impact factor is the best and most appropriate measure of academic quality and impact.

Of the seventy-one articles that were examined in this literature review, seventeen of them contain long lists of the issues posed by the use of the impact factor as an evaluative and assessment tool. The impact factor, according to Adam is ‘so riddled with errors and biases, it can be worse than useless’ (2002, p. 729) and indeed, Bordons, Fernandez et al. (2002) state that the ‘abuse and incorrect use of IF measures’ (2002, p. 205) are the two underlying problems with the IF system. There are certain problems with the structure of the system itself, such as the fact that books often are not included in the database as a source of citation, even though they are centrally important to some fields (Crookes, et al., 2010; Seglen, 1997b); citation data is vulnerable to technical problems and so often contains errors (Adam, 2002; Frank, 2003; Kostoff, 1997; Kurmis, 2003; Opthof, 1997; Seglen, 1997a, 1997b; Smith, 1998); informal and negative influences are often not properly cited, or cited at all (MacRoberts et al., 1989); marginalia is often not included, such as ‘letters, news articles, book reviews, abstracts’ (Frank, 2003, p. 5); and referencing errors are rife (Bloch et al., 2001; Seglen, 1997a, 1997b).

Intentional abuse of the system is also a serious problem with the impact factor, such as the fact that review articles are heavily cited so as to inflate the impact factors of journals (Adam, 2002; Al-Awqati, 2007; Bevan, 2004; Crookes, et al., 2010; Seglen, 1997a, 1997b). Self-citation is rife (Bevan, 2004; Crookes, et al., 2010; Kostoff, 1997; MacRoberts, et al., 1989; Seglen, 1997a, 1997b; Zavos, et al., 2006); biased citing, in relation to in-house citations, is prevalent and is employed so as to boost the IF of a journal (MacRoberts, et al., 1989; Seglen, 1997b; Zavos, et al., 2006); and there is a language bias towards English. The ISI is also dominated by American publications which impacts unfairly upon citation counts (Adam, 2002; Bevan, 2004; Favaloro, 2009; Genoni, et al., 2009; Kostoff, 1997; Kurmis, 2003; MacRoberts, et al., 1989; Seglen, 1997a, 1997b; Smith, 1998), and journals with a high impact factor are more likely to cover broader areas of basic research rather than specialty disciplines (Bevan, 2004; Crookes, et al., 2010; Fava, et al., 2000; Smith, 1998). These are all issues that demonstrate the problem of relying on one single blanket measure which cannot account for discipline specific differences in research practice and outcomes. This can also be a problem when it comes to measures of individual impact, such as the more recent H-Index.

### ***The H-Index***

The h-index, developed by Professor Jorge Hirsch in 2005 is also a popular bibliometric tool, invented for the measurement of research practice, but unlike the impact factor, the h-index was invented specifically as a metric to measure and evaluate the output of individual researchers (Ortner, 2010). Since its invention in 2005, the h-index has become a ‘popular way for academics to rank themselves relative to their peers’ (Jackson, et al., 2009, p. 2537). Like many other bibliometric tools, the h-index was invented as a solution to the ‘growing demand to quantify research output’ and has since ‘tempted funding agencies, promotion committees and employers to treat numerical indices of research output’ (Kelly et al., 2006, p. 167) as innately revealing of a scholar’s impact and the quality of their work. The h-index is defined by Kelly and Jennions as ‘the maximum number of papers  $h$  by a scientist where each paper has received  $h$  or more citations’ (2006, p. 167), which means that it is based on the scholar’s most frequently cited papers and the amount of citations one has received in other publications.

A number of scholars have attempted to justify the use of the h-index, with some noting that it is an improvement on older indices (such as the impact factor) because it doesn’t assume that an individual researcher’s performance is equal to that of the journal in which they have been published (Ortner, 2010). Other justifications for its use are that it is easy to calculate; that it provides a fairer comparison across disciplines; that it is conceptually simple and more accurate; that it is more comprehensive than the impact factor; that it does not reward unfairly for highly cited papers and/or penalise for low or uncited papers; and that it provides a good estimate of the impact of one’s cumulative research (Harzing, et al., 2009; Jackson, et al., 2009; Ortner, 2010).

That said, employment of the h-index as the sole tool used to measure an academic’s work poses the problem of reducing the evaluation of a whole complex body of work (which may have taken different forms over time) to one single number, which is, according to Harzing & Wal (2009) ‘unlikely to provide a complete picture of a scholar’s real impact’ (2009, p. 45). A number of scholars have raised concerns about the use of the h-index, arguing that it houses the assumption that papers accumulate at a fixed rate of citation, when in reality most papers go through a limited period of active citation (Kelly, et al., 2006). It has also been argued that it only allows for scholars with similar publishing years to be fairly compared (Ortner, 2010) and that it can only rise or stay constant, it can never decrease and thus it cannot indicate periods of inactivity, retirement or death (Kelly, et al., 2006).

There are also issues associated with the h-index that are common to the problems posed by the impact factor and bibliometrics as a whole, such as the data problems encountered in relation to the misspellings of names, common names, changed names and those who publish in several areas (Jasco, 2010; Oppenheim, 2008; Watson et al., 2006), its inability to be compared fairly between disciplines (Kelly, et al., 2006); and the problem of inconsistent or manipulated data entry has also been discussed (Jasco, 2010). Bibliometric tools like the impact factor and the h-index are also, according to some, fundamentally flawed in relation to the fact that papers are sometimes cited ‘for reasons that are unrelated to the quality or utility of a study’ (Kelly, et al., 2006, p. 167). Thus, the verity of the information generated by the employment of these tools as indicators of academic excellence and impact is debatable. Bibliometrics alone are often flawed, and given they focus only on publications, do not present a whole picture of the impact of research beyond scholarly journals. With these problems in mind, we searched the literature for alternative ways of thinking about and recording and reporting impact.



## **Alternative approaches to measuring impact**

### ***Impact analysis frameworks***

In recent years, governments have begun investing in research ‘not for its intrinsic worth, nor to win esteem and still less to indulge researchers’ curiosity, but for its contribution to economic development’ (Moodie, 2006, p. 132). Internationally, there has been a growing trend whereby governments and universities are placing more emphasis on academic practices that have the potential for wider social impact. The literature demonstrates a gravitation towards policies and projects which focus on measures of impact, with Grant (2006), Moodie (2006), Buxton (2004), Hanney et al. (2004), Theus (1993) and Hanney (2004) all stating that a requirement for governments and organisations to be more accountable for their research expenditure, has led to pressure being placed upon academics to justify the social value of their academic practice.

An impetus has thus been placed upon academics to produce work that has the potential to be utilised by governments and organisations to benefit the greater community. Some scholars (Buxton, et al., 2004; Grant, 2006; Hanney et al., 2003; Moodie, 2006) have speculated about the development of a new mode of assessment, that focuses ‘not [on] how many reports have been done, but how many people’s lives have been bettered by what has been accomplished’ (Grant, 2006, p. 1). This betterment could be economic, behavioural, clinical, environmental or social; in a small research group or world-wide (Moodie, 2006). This would be, according to Moodie ‘a radically different orientation to cultivating research esteem’ (Moodie, 2006, p. 133) in that it would position those who use research and who are impacted by it, as the evaluators and judges of it.

Internationally, work is being done to try and develop systems which employ this kind of impact assessment. The research undertaken by RAND Europe for the Higher Education Funding Council of England (HEFCE) is presented in (Grant et al., 2010) and provides a concise and informed evaluation of international practice in relation to the development of measures of research impact. Grant et al. (2010) examine the policies discussed and implemented in the UK with the RAND/ARC impact scoring system (RAISS); in the US with the Program Assessment Rating Tool (PART); in the Netherlands with Evaluating Research in Context (ERiC) and in Australia in relation to the Research Quality Framework (RQF), now abandoned and replaced with ERA. They argue that all four policy frameworks involve the measurement and evaluation of impact through the use of non-bibliometric, quantitative methods, suggesting that globally, more value is being placed on the external impact (socially, economically, environmentally etc.) of research (Duryea, et al., 2007; Grant, et al., 2010).

As a result of the review undertaken by RAND for the HEFCE, the Research Excellence Framework (REF) was developed for the UK. It is based closely on Australia’s now defunct Research Quality Framework, and focuses on the measurement and evaluation of research impact across a number of different categories: social, cultural, economic, environmental, public policy, and quality of life (DEST, 2007; HEFCE, 2009). The framework provides suggestions for the implementation of methods to measure impact (highlighting in particular ‘impact statements’), but only at project, faculty or institutional levels, not as a tool for the measurement of the impact of individual scholars. This is the major gap in the literature which this review revealed.

### ***Measuring individual impact***

A number of scholars have attempted to articulate new ways of assessing the impact of academic practice that are not based purely on bibliometric indicators (DEST, 2007; Duryea, et al., 2007; Hanney, 2004; HEFCE, 2009; Hicks, 2005; Kalucy et al., 2007; Kellogg, 2004; Kuruvilla et al., 2007; Lavis et al., 2003; Becker Medical Library Library, 2011; Molas-Gallart et al., 2000; Sarli et al., 2010; Smith, 2001). For the most part, all of these publications only explore new methods in relation to measuring impact at a research project or faculty level. When combined, however, they provide a raft of possible assessment criteria and evidentiary sources for the measurement of impact which could be adapted at the individual level.

Lavis suggests surveys, structured interviews with decision makers, document reviews, observations of decision making process, analyses of data collected and the study of research organisations' files can all be used as evidence of an individual's impact (Lavis et al., 2002). Similarly, Beacham argues that the study of clinical guidelines, peer review, bibliometrics, patent analysis, study of administrative decisions, examination of official statements and examination of guidelines and evaluative criteria (Beacham et al., 2005). Hanney discusses semi-structured interviews with key informants, and the use of document and literature reviews (Hanney et al., 2004), while Kalucy argues for the inclusion of contact details of those who can confirm the use of research, the submission and citation of policy documents, organisational documentation, reports of minutes of meetings and the inclusion of statements made about the scholar's work by managers or decision makers (Kalucy, et al., 2007).

One particular method of reporting impact which is constant through all of these articles, and which draws together the individual measures of impact across the range of these articles, is the employment of 'impact narratives' or 'impact statements'. Both REF (HEFCE, 2009) and RQF (DEST, 2007) suggested the use of 'impact statements' or 'impact narratives' to demonstrate the significance of research outcomes beyond the scholarly journal, guided by these stipulated categories. Both systems provide an outline for developing impact statements, structuring them as a series of free-text case studies, each case study demonstrating a specific example of impact, guided by a template of suggested indicators (provided by the institution). Both REF and RQF contain draft templates for case studies that can be adapted to develop documentation and guidelines for the writing of impact statements for individual staff members. REF and RQF assert that within each case study, the range and significance of impact gained through the work that has been undertaken should be discussed, using examples which are easily identifiable, supportable and evidence-based, thus demonstrating that qualitative analysis can be as reliable a measure of a scholar's work as quantitative tools.

Most of the work which has been undertaken into developing new methods for measuring impact is concerned with the development of new tools at a research project or institution level, however, the same principles, methods of assessment and forms of evidence can be adapted to develop evaluative tools for the assessment of individual academics. Certainly, impact statements could be structured and supported in such a way as to demonstrate the specific role an individual may have had within each research project, the impact that that work has had on the intended audience, and the wider esteem with which the academic is held in their profession.

## **Conclusion**

It is significant that the literature surveyed in this review does not provide a ready-made approach to the measurement of individual impact and esteem. It is a complex issue, and the problems of subjectivity and bias around the traditional peer review system were what led to the development of bibliometrics initially. It is now well recognised, however, that an over reliance on bibliometrics has proven problematic because they may give the appearance of scientific objectivity but suffer from a number of inherent flaws. More than this, they do not tell the whole story as they are only able to report on the impact a person is having within scholarly publications. Many disciplines are concerned with the impact a research project or individual researchers are having in the wider community, especially the actual intended audience of the research. This is not an easy thing to measure; there is no one single neat way of doing so.

However, existing and newly developed impact analysis frameworks provide a possible key to the solution. By using the idea of ‘impact statements’ which draw on a number of sources (including bibliometrics) for evidence of impact, across a range of categories, individuals could be provided with a tool by which to demonstrate their impact and the esteem with which they are held in their discipline (both academic and professional). By facilitating the use of such statements, tailored to reflect an institution’s performance expectation criteria, university promotion systems can become more able to account for the diversity of academic work, and can do so in a rigorous and evidence based way. More broadly, this literature review has demonstrated the increasing global concern with the importance of ‘impact’ for research evaluation and has signalled the urgency of developing ways in which it can be measured and reported beyond bibliometrics, for the benefit of both research projects and individual researchers.

## References

- Adam, D. (2002). The Counting House *Macmillan Magazine*
- Al-Awqati, Q. (2007). Impact Factors and Prestige *Kidney International* 71, 183-185
- Beacham, B., Kalucy, L., & McIntyre, E. (2005). Understanding and measuring research impact. *Primary Health care research and information service*(2), 1-12.
- Bennett, D., Genoni, P., & Haddow, G. (2011). FoR codes pendulum: Publishing choices within Australian research assessment. *Australian Universities Review*, 53(2), 88-98.
- Bevan, D. (2004). Impact Factor 2002 - too much "impact"? *Clinical Medical Association* 27(2), 65-66.
- Bloch, S., & Walter, G. (2001). The Impact Factor: time for change *Australian and Newzeland Journal of Psychiatry* 563-568.
- Bordons, M., Fernandez, M. T., & Gomez, I. (2002). Advantages and limitations in the use of impact factor measures for the assessment of research performance in a peripheral country *Scientometrics* 53(2), 195-206.
- Butler, L. (2003). Modifying publication practices in response to funding formulas. *Research Evaluation*, 12(1), 39-46.
- Buxton, M., Hanney, S., & Jones, T. (2004). Estimating the economic value to societies of the impact of health research: a critical review *Bulletin of the World Health organization* 82(10), 733-739.
- Coates, H., & Goedegebuure, L. (2010). *The Real Academic Revolution: Why we need to reconceptualise Australia's future academic workforce, and eight possible strategies for how to go about this*. Carlton VIC: L H Martin Institute.
- Cooper, S., & Poletti, A. (2011). The new ERA of journal ranking: The consequences of Australia's fraught encounter with 'quality'. *Australian Universities Review*, 53(1), 57-65.
- Crookes, P. A., Reis, S., & Jones, S. C. (2010). The Development of a ranking tool for refereed journals in which nursing and midwifery researchers publish their work *Nurse Education Today*, 30, 420-427.
- DEST. (2007). *Research Quality Framework: Assessing the quality and impact of research in Australia - RQF Submission Specifications*. Canberra: Australian Government: Dept of Education, Science and Training.
- Duryea, M., Hochman, M., & Parfitt, A. (2007). Measuring the impact of research. *Research Global*, 8-10.
- Fava, G. A., & Ottolini, F. (2000). Impact Factors versus Actual Citations *psychotherapy and Psychosomatics* 69(6), 287-286.
- Favaloro, E. J. (2009). The Journal Impact Factor: don't expect its demise any time soon *Clinical Chemical Laboratory Medicine* 47(11), 1319-1324.
- Frank, M. (2003). Impact Factors: arbiter of excellence? . *Journal of Medical library Association* 91(1), 4-6.
- Garfield, E. (1996). Fortnightly Review: How can impact factors be improved? *BMJ*, 313(411).
- Garfield, E. (1999). Journal Impact Factor: a brief review. *Canadian Medical Association Journal*, 161(8), 979-980.
- Genoni, P., & Haddow, G. (2009). ERA and the Ranking of Australian Humanities Journals *Australian Humanities Review* 1(46), 7-26.
- Grant, J. (2006). *Measuring the Benefits from research* Cambridge RAND.
- Grant, J., Brutscher, P.-B., Kirk, S. E., Butler, L., & Wooding, S. (2010). *Capturing research impacts: A review of international practice*. Cambridge UK: Higher Education Funding Council for England.

- Haddow, G., & Genoni, P. (2010). Citation analysis and peer ranking of Australian social science journals *Scientometrics*, 85, 471-478.
- Hanney, S. (2004). Personal Interaction with Researchers or Detached Synthesis of the Evidence: Modelling the Health Policy Paradox. *Evaluation and Research in Education*, 18(1), 72-82.
- Hanney, S., Gonzalez, M., Buxton, M., & Kogan, M. (2003). The utilisation of health research in policy-making: concepts, examples and methods of assessment *Health Research Policy and Systems*, 1(2), 1-28.
- Hanney, S., Grant, J., Wooding, S., & Buxton, M. J. (2004). Proposed methods for reviewing the outcomes of health research: the impact of funding by the UK's 'Arthritis Research Campaign' *Health Research Policy and Systems* 2(4), 1-11.
- Harzing, A.-W., & Wal, R. v. d. (2009). A Google Scholar h-index for journals: An Alternate Metric to Measure Journal Impact in Economics and Business *Journal of the American Society for Information Science and Technology*, 60(1), 41-46.
- HEFCE. (2009). *Research Excellence Framework: Second consultation on the assessment and funding of research*. London.
- Hendrix, D. (2010). Tenure Metrics: Bibliometric Education and Services for Academic Faculty *Medical Reference Services Quarterly* 29, 183-189.
- Hicks, V. (2005). *Developing of CIHR Framework to Measure The Impact of Health Research*. Ontario CIHR
- Jackson, D., Haigh, C., & Watson, R. (2009). Nurses and Publications - the impact of the impact factor *Journal of Clinical Nursing* 18(18), 2537-2538.
- Jarwal, S. D., Brion, A. M., & King, M. L. (2009). Measuring research quality using the journal impact factor, citations and 'Ranked Journals': blunt instruments or inspired metrics? . *Journal of Higher Education Policy and Management*, 31(4), 289-300.
- Jasco, P. (2010). Pragmatic issues in calculating and comparing the quantity and quality of research through rating and ranking of researchers based on peer reviews and bibliometric indicators from Web of Science, Scopus and Google Scholar *Online Information Review* 34(6), 972-982.
- Kalucy, L., Bowers, E. J., & McIntyre, E. (2007). Primary Health Care Research Impact Project.
- Kellogg, W. K. (2004). *Logic Model Development Guide* Michigan: W.K Kellogg Foundation
- Kelly, C. D., & Jennions, M. D. (2006). The h index and career assessment by numbers. *TRENDS in ecology and evolution*, 21(4), 167-170.
- King, J. (1987). A review of bibliometric and other science indicators and their role in research. *Journal of Information Science*, 13(1), 261-275.
- Kostoff, R. N. (1997). Use and Misuse of Metrics in Research Evaluation *Science and Engineering Ethics*, 3(2), 109-120.
- Kurmis, A. P. (2003). Understanding the Limitations of the Journal Impact Factor *The Journal of Bone and Joint Surgery* 85(12), 2449-2454.
- Kuruvilla, S., Mays, N., & Walt, G. (2007). Describing the impact of health services and policy research. *Jornal of Health Services Research Policy* 12, 23-31.
- Lavis, J., Ross, S., McLeod, C., & Gildiner, A. (2003). Measuring the impact of health research. *Journal of Health Services Research and Policy*, 8(3), 165-170.
- Lavis, J., Ross, S. E., Hurley, J. E., Hohenadel, J. M., Stoddart, G. L., Woodward, C. A., & Abelson, J. (2002). Examining the Role of Health Services Research in Public Policy Making *The Milbank Quarterly* 80(1), 125-154.
- Library, B. B. M. (2011). Bernard Becker Medical Library: Model for Assessing the Impact of Research. from <https://becker.wustl.edu/impact/assessment/index.html>

- MacRoberts, M. H., & MacRoberts, B. R. (1989). Problems of Citation Analysis: A Critical Review. *Journal of the American Society for Information Science*, 40(5), 342-349.
- Martin, B. (2011). ERA: adverse consequences. *Australian Universities Review*, 53(2), 99-102.
- Molas-Gallart, J., Tang, P., & Morrow, S. (2000). Assessing the non-academic impact of grant-funded socio-economic research: results from a pilot study. *Research Evaluation*, 9(3), 171-182.
- Moodie, G. (2006). Vocational education institutions' role in national innovation *Research in Post-Compulsory Education*, 11(2), 131-140.
- Oppenheim, C. (2008). Out With the Old and in With the New: the RAE, Bibliometrics and the New REF *Journal of Librarianship and Information Science* 40(3), 147-149.
- Opthof, T. (1997). Sense and nonsense about the impact factor *Cardiovascular Research* 33, 1-7.
- Ortner, H. M. (2010). The impact factor and other performance measures - much used with little knowledge about *International Journal of Refractory Metals and Hard Materials* 28, 559.
- Sarli, C. C., Dubinsky, E. K., & Holmes, K. L. (2010). Beyond citation analysis: a model for assessment of research impact. *Journal of the Medical Library Association*, 98(1), 17-23.
- Seglen, P. O. (1997a). Why the impact factor of journals should be used for evaluating research. *BMJ*, 314(497).
- Seglen, P. O. (1997b). Citations and Journal impact factors: questionable indicators of research quality *Allergy* 52, 1050-1056.
- Smith, R. (1998). Unscientific practice flourishes in science: impact factors of journals should not be used in research assessment *BMJ*
- Smith, R. (2001). Measuring the social impact of research *BMJ*, 323, 528.
- Theus, K. T. (1993). Academic Reputations: The Process of Formation and Decay *Public Relations Review* 19(3), 277-291.
- Watson, R., & Thompson, D. R. (2006). Use of factor analysis in *Journal of Advanced Nursing: literature review Journal of Advanced Nursing* 55(3), 330-341.
- Wilson, J. (2011). Creative arts research: A long path to acceptance. *Australian Universities Review*, 53(2), 68-76.
- Zavos, C., Kountouras, J., & Katsinelos, P. (2006). Impact factors: looking beyond the absolute figures and journal rankings *Gastrointestinal Endoscopy* 64(6), 1034.