Reflections: On Publishing

Can research quality be measured quantitatively? On quality of scholarship, numerical research indicators and academic publishing – experiences from Norway

MICHAEL JONES

Jones, Michael (2017). Can research quality be measured quantitatively? On quality of scholarship, numerical research indicators and academic publishing – experiences from Norway. *Fennia* 195: 2, pp. 164–174. ISSN 1798-5617.

In this article I reflect on ways in which the neoliberal university and its administrative counterpart, new public management (NPM), affect academic publishing activity. One characteristic feature of NPM is the urge to use simple numerical indicators of research output as a tool for allocate funding and, in practice if not in theory, as a means of assessing research quality. This ranges from the use of journal impact factors (IF) and ranking of journals to publication points to determine what types of work in publishing are counted as meritorious for funding allocation. I argue that it is a fallacy to attempt to assess the quality of scholarship through quantitative measures of publication output. I base my arguments on my experiences of editing a Norwegian geographical journal over a period of 16 years, along with my experiences as a scholar working for many years within the Norwegian university system.

Keywords: bibliometrics, editing, impact factor (IF), new public management (NPM), Norway, peer review, publication points, research quality

Michael Jones, Department of Geography, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway. E-mail: michael.jones@ntnu.no

Introduction

A debate over publication points arose in Norwegian media after the death of the distinguished Norwegian political scientist, professor Frank Aarebrot, in September 2017. Publication points are allocated for published scholarly works with the aim of making research measurable as a means of allocating funding to higher-education and other research institutions on the basis of scholarly production rather than, for example, on the basis of the number of academic positions in an institution (Hansen 2015). Professor emeritus in education Svein Sjøberg (2017) pointed out that although Aarebrot was widely acclaimed as one of the most visible, productive and significant researchers in public debate and considered as an inspirational lecturer by students, only two articles that gave publication points were reported during his last ten years – yet he had 841 registered contributions in the form of newspaper articles, contributions in other media and public lectures. Sjøberg argued that Norway needs such academics, who 'live up to the best ideals of the academic's role as an informed participant in a living democracy'. However, the publication points system favours articles written in English in peer-

URN:NBN:fi:tsv-oa66602
DOI: 10.11143/fennia.66602

reviewed journals. Sjøberg claimed that this was used as an indicator of research quality when academics apply for jobs, promotion and pay rises.

Sjøberg's article provoked a reply from undersecretary at the Ministry of Education and Research (Kunnskapsdepartmentet), Bjørn Haugstad (2017). He argued that lack of publication points had not hindered the recognition of Aarbrot's work in teaching and public dissemination, although he stated that in this respect Aarebrot's contribution was untypical. However, attempts to introduce an indicator for public dissemination had been abandoned because it was found to involve too much reporting and administration without a clear benefit. Moreover, Haugstad claimed that publication points were not intended to serve as a means of assessing an individual researcher's productivity and quality. Their use was simply to measure the volume of peer-reviewed publications as part of the funding system of universities, and moreover accounted for only 1.6% of the universities' total funding. One of the instigators of the Norwegian publication indicator, professor Gunnar Sivertsen (2009, 25, 29), has similarly argued that it was not intended as a measure of quality or meant to replace the qualitative evaluation of scholarly work.

Nonetheless, if publication points are not intended as a measure of quality but only of quantity, one is left wondering how publication volume can logically be justifiable as an instrument for allocating part of university funding (beyond the fact that the indicator is simple and easy to use for bureaucrats). However much it is denied, the underlying assumption in practice seems to be that a high rate of publication in selected journals is the most important determinant of research quality, without taking into consideration the complexity of different types of academic endeavour by a multiplicity of scholars that is necessary to secure quality.

The neoliberal university and new public management

Despite strong protests from many academic staff, the university sector throughout Europe as well as elsewhere in the world has in recent years been undergoing a radical process of transformation from public service institutions to what is variously termed as the neoliberal university, academic capitalism, or the commodification of knowledge (Kjeldstadli 2010; Paasi 2013; Halffman & Radder 2015, 2017; Lund 2015; van Reekum 2015; Myklebust 2017a, 2017b). According to neoliberal ideology, universities are expected to adapt to market thinking and serve economic growth. Universities should function organizationally like businesses with strong top-down managerial steering. The aim is to promote efficiency by adopting principles of management from the private sector. The emphasis is on performance indicators, competition and cost-effectiveness. The field of education is still to a large extent financed through taxation and public funding rather than through the free market, but pseudo-market indicators have been developed to measure performance. Ranking and scores are examples of pseudo-market indicators, designed to stimulate competition between actors as if they were acting in a market. New public management (NPM) is the bureaucratic instrument whereby universities are to be pushed in a neoliberal direction.

Among features of NPM that directly impinge on publication activity is the increasing use of performance indicators in the allocation of funding along with increased reporting and auditing of results. Performance indicators are measurable values that are intended to show how effectively a company or organization is in achieving its key objectives. One of the key objectives of universities and other research organizations is to achieve high quality research.

In this article, I question whether the objective of research quality can be validly measured through quantitative performance indicators. I discuss publication points and journal impact factors (IF) as two commonly used and mutually related performance indicators. I contrast these with types of scholarly work that are generally not 'counted' by performance indicators but which are nevertheless important, often vital, for ensuring the quality of scholarship. I discuss intended and unintended consequences for academic publication activities of this differential use of performance indicators. I use examples from my experiences of editing *Norsk Geografisk Tidsskrift–Norwegian Journal of Geography (NGT–NJG)* from 1999 to 2014 and my wider experiences as a scholar working in the Norwegian university system from 1973 to the present.

Publication points – the Norwegian case

Bibliometrics – the measuring and weighting of publications (also termed in Norway the 'publication indicator'[1]) – was introduced into the Norwegian university sector in 2006 as an indicator of research productivity. Bibliometrics originated in the mid-20th century within the history of science and sociology of knowledge as a tool for studying the development of modern science and social networks within it, and only later became a tool of research evaluation, to begin with in natural sciences (Roll-Hansen 2009, 72–73). The background to the introduction of the Norwegian publication indicator has been given by Gunnar Sivertsen (2009). Models based on it have subsequently been adopted in other Nordic countries, including Finland (Paasi 2013, 6; Sivertsen 2016, 79).

The allocation of publication points is intended to stimulate increased scholarly publication. To be allocated publication points, research must be published in approved scientific or scholarly channels, meaning that articles are to be published in authorized journals, and books and book chapters are to be published by authorized publishers. Authorized publication channels are approved in two categories, level one and level two. Those considered as the 20% 'most prominent and influential' publication channels are credited as level two (Hansen 2015).

Proposals for publication channels can be made by researchers to the Norwegian Centre for Research Data (NSD, Norsk senter for forskningsdata). To be approved, publication channels must be identifiable by a valid ISSN or ISBN. Journals must have an editorship comprising researchers as well as routines for external peer-reviewing, while book publishers must have routines for evaluation by external consultants. The range of authors must be national or international (NSD 2017). To be eligible for level two, proposals sent by the national committees for different disciplines are assessed by the National Publications Committee of the Norwegian Association of Higher Education (UHR, Universitets- og høgskolerådet). The National Publications Committee updates the list of publication channels on level two annually (UHR n.d.).

Researchers register their publications in the database of the Current Research Information System in Norway (CRIStin, Det nasjonale forskningsinformasjonssystemet). Predecessors of the CRIStin database originated as a general register for all types of output by academics, a function it still has, and it was later developed into a bibliometric instrument for allocating funding. Hence a distinction is practised between general registration of output and reporting for bibliometric purposes. Bibliometric reporting applies to a strictly defined category of approved scientific or scholarly publications, referring to those that are admissible to level one or level two publication channels. There are routines intended to prevent double registration, for example of digital and printed versions of a work, original and revised versions, and original works and translations (CRIStin 2017). The publication points allocated for different types of approved publication are shown in Table 1.

Types of output that are not accredited with publication points include textbooks, books aimed at the general public, popular science works, debate books, working reports, memorandums, reference works, translations, factual prose and technical literature not based on original research, and fiction. Not all types of article in approved publications are eligible for publication points. Editorials, leaders, commentaries, debate articles, obituaries, interviews, bibliographies, and book reviews are examples of academic output that are not accredited in this way. However, peer-reviewed overviews and review articles in journals receive the same publication points as ordinary research articles (CRIStin 2017; Fossum-Raunehaug 2017).

Table 1. Publication points given to different types of approved scientific or scholarly publication in the Norwegian model (CRIStin 2017).

Publication type	Level one	Level two
Article in periodical or series (journal article) (ISSN)	1	3
Monograph (book) (ISSN/ISBN)	5	8
Chapter in anthology (ISBN)	0.7	1

In 2013 the Danish Centre for Research Analysis (Dansk center for forskningsanalyse) at Aarhus University was commissioned by the Norwegian Association of Higher Research Institutions to undertake an evaluation of the Norwegian publication points system. It concluded that the Norwegian system was simple, academically well thought-out, and relatively cheap, and had resulted in a considerable increase in output in relation to the available resources. However, adjustments were proposed to stimulate greater cooperation between researchers nationally and in particular internationally, and the weighting system was changed in 2017 to encourage more co-authorship between universities and countries. A new formula for weighting co-authorship between institutions was devised. This was intended to reduce discrepancies that had arisen in the allocation of points between different types of research institution, to counteract the tendency to list more authors than strictly necessary, and to make it more difficult to use the publication points as an indicator at the level of the individual researcher. Other problems identified in the evaluation were lacking transparency and legitimacy of the annual journal nomination process, and use of the publication indicator for allocating funding and incentives internally within institutions, faculties, departments and even at individual level (Aagaard *et al.* 2014, 2015; Kunnskapsdepartementet 2016, 69–70; Sivertsen 2016, 85–87).

Critiques of bibliometric methods have raised a number of issues, summarized in a report to the Norwegian Academy of Science and Letters (DNVA – Det Norske Videnskaps-Akademi) (Østerud 2009). It was argued that publication points weight research quantity rather than quality. They do not sufficiently consider differing publication traditions in different disciplines. They may result in strategic adaptation towards certain types of publishing that does not necessarily lead to better research. They do not reward research communication in non-scientific channels. There is a danger that they may in practice be used as a measure of quality for assessing the work of individual researchers. The quality of the measuring methods is uncertain, and high journal rankings may reflect popularity rather than quality. In the same report, Nils Roll-Hansen (2009, 76), professor in the history and philosophy of science, argued that the introduction of publication points reflects increasingly bureaucratic management of research, which is based not so much on an understanding of the research being undertaken as on applying formal measures of productivity.

In Europe generally, a move to classify journals into levels by the European Science Foundation through the European Reference Index for the Humanities (ERIH) led to a joint response from editors of 56 journals in the history of science, technology and medicine. They stated that classifying journals in this way was doubtful, arbitrary and formalistic, and undermined the broad debate that academic renewal and critical evaluation of research quality is dependent on. On the subject of creativity, it was pointed out:

> Truly ground-breaking work may be more likely to appear from marginal, dissident or unexpected sources, rather than from a well-established and entrenched mainstream (Andersen *et al.* 2009, 4, cited in Roll-Hansen 2009, 79).

The use of quantitative measures for research assessment in the humanities as well as in social sciences has raised widespread concern among scholars in a range of European countries. A recent report coordinated by the University of Zürich revealed a diversity of views regarding how best to assess research in the humanities but found there was no easy way. Regarding bibliometrics, it was concluded that for the humanities 'bibliometric analysis of publications cannot be used as a sole assessment tool' as it is 'an instrument that is too simplistic and one-dimensional to take into account the diversity of impacts, uses and goals of humanities research' (Ochsner *et al.* 2016, 8).

I have observed a number of issues of contention regarding publication points in discussions with colleagues at my university. A general issue is that since the number of publication points has increased without a corresponding increase in resources, the funding allocation per publication point has decreased over time. Another issue is the expectation that research articles should aim at publication in high-status international English-language journals, referring to journals published principally in the USA and the UK. Overlooked is the fact that English-language journals published in small countries outside this core may be more international in submissions and article provenance than many Anglo-American journals. Yet the latter tend to have higher IFs and constitute the majority

of level two journals. A further matter is that discussions and decisions regarding which 20% of journals should be considered as level two journals is subject to negotiation, compromise and political positioning between disciplines and sub-disciplines more than on the unbiased quality assessment of journals, although IF clearly plays a part in these decisions. Which journals are to be on level two is reviewed annually, and this can result in journals changing status from one level to another from one year to the next. Due to the time lapse between acceptance and publication, an article accepted for a level two journal may end up being published at level one if the journal is demoted in the meantime, or vice versa.

A further issue concerns the classification of publications. Books that have the word 'encyclopedia' in their title are deemed to be reference works and hence such contributions are ineligible for publication points. Yet writing them often involves a disproportionate amount of work in relation to their length, and is not infrequently based on considerable original research. Entries in *The Dictionary of Human Geography*, for example, are in many cases similar to overview or review articles published in journals, but unlike the latter do not receive points.

As journal editor, I experienced that it could be difficult to find book reviewers, as book reviews do not earn points and therefore tend to be given reduced priority in a busy work schedule. Yet book reviews are an important source of information for researchers and their inclusion in academic journals can make a significant contribution to the critical discussion of research quality. Again, obituaries are not allocated points, yet require not inconsiderable research and are a potentially useful source for the history of knowledge. The exclusion of textbooks and books aimed at the general public from the publication points system, despite the amount of work and not infrequently research that they involve, renders writing for this type of output less attractive for many university scholars. There is no doubt that the system of publication points influences what is given priority to and what is not.

The current system of publication points results in several paradoxes. The number of points given to journal articles is disproportionate to those given to monographs in terms of size and work involved. At level one, five short journal articles are given the same weight as a book and, at level two, three journal articles weigh more than a book, even if the latter may have required several years of painstaking and detailed research (Sandnes 2016).

Norwegian medical professor Per O. Seglen (2009, 40) states that there is no documentation to support the notion that articles in journals on level two are of higher quality than level one. He argues that the difference in weightings between levels one and two is arbitrary and has no justification.

A recent study found that, although as whole articles in level two journals are cited more often than those at level one, the frequency of citation varies considerably and not all articles at level two are cited. The same study found that articles in open access journals are cited more frequently than those in other types of journal – as one would expect because of their greater accessibility. However, few of the open access journals were on level two (Aksnes 2017).

The question arises as to how far do bibliometrics for research funding lead to high-quality research and how far merely to mass production of low-quality research (Gjengedal 2017). In Sweden, bibliometric methods of research evaluation were critically discussed at a conference arranged by the Academy of Letters, Antiquities and History in Stockholm (KVHAA – Kungl. Vitterhets Historie och Antikvitets Akademien). It was pointed out that, especially in natural sciences, publication points may encourage 'salami publication', that is 'slicing research results into the "smallest possible publishable units"' to get as many publications as possible from one study (Waarenperä 2011, 29). Researchers are rewarded for frequent publication but not for devoting time to peer-reviewing the research of others. The conference report further emphasized that research creativity and productivity were two different things, but only the latter is measured (Waarenperä 2011, 36, 39).

Impact factors

Academics are expected to aim to publish their research in the internationally most prominent and influential journals in their field. It is widely considered that an academic journal's prominence and influence are indicated by its impact factor (IF). Impact factors were introduced in 1975 by the Institute for Scientific Information (ISI), now part of the commercial organization Web of Science, which in 2014

indexed more than 12,000 journal titles. IF is used as a proxy to measure the importance or rank of a journal by calculating how often its articles are cited in indexed journals. It measures the frequency with which articles in the journal are cited in a given period. A journal's IF is normally calculated by first counting the number of times articles published in the journal during two consecutive years are cited in the following year in all journals in an indexed database (A); the number of cited articles is then divided by the total number of citable articles published in the journal in the same two-year period (B), that is IF = A/B. 'Citable articles' are those that present original research and have undergone peer review before publication.

Critical voices warned twenty years ago of the dangers of using the impact factor of journals to evaluate research. Seglen (1997) pointed out that high impact factor is implicitly considered an indicator of journal prestige, which is widely used as an evaluation criterion. Nonetheless, he stated, a journal cannot be regarded as representative of an individual article published in the journal. Apart from being non-representative, he argued, journal IF has technical shortcomings, such as bias caused by the manner of calculation. He referred to studies showing that while the database of citable articles was limited to standard research articles, research notes and review articles, the number of cited articles included in addition those cited in, for example, editorials, letters and conference reports. This favoured journals that included 'meeting reports, interesting editorials and lively correspondence' (Seglen 1997, 500). The fact that the citation count refers only to articles published in the previous two years also leads to temporal fluctuations in a journal's IF. Furthermore, research fields where a significant part of scientific output consists of books, which are not included in the journal article database, may be discriminated against in a research evaluation based on IF. Such technicalities are unrelated to the scientific quality of research output. Another aspect is that the dominance of English-language journals in the citation index database contributes to low IF for the comparatively few non-English language journals that are included 'since most citations to papers in languages other than English are given by other papers in the same language' (ibid., 500). Large English-language research communities such as North America tend to cite mainly other papers in English, thus increasing the citation rate and journal impact of their own community (ibid., 500–501). Seglen (1997, 502) concluded:

> '...citation impact is primarily a measure of scientific utility rather than of scientific quality, and authors' selection of references is subject to strong biases unrelated to quality. For evaluation of scientific quality, there seems to be no alternative to qualified experts reading the publications.'

It is paradoxical that a citation count may be boosted by poor quality articles that are cited when they are criticized for poor research. In Sweden, it was argued at the conference on bibliometric methods that especially in the humanities citation ranking can lead to scholarly overstatement to gain attention and that publication of controversial work may result in increased citations. A fixation on citations and ranking has led discussions away from what good research entails. There is uncertainty regarding whether research quality actually benefits from a citation index (Waarenperä 2011, 29, 34, 36).

Geography professor Anssi Paasi (2013) has summarized in *Fennia* the debate concerning how the Web of Science and IFs have come to determine what are considered as the most significant international journals. Inclusion in the Web of Science database has become a '"synonym" for quality' (Paasi 2013, 2) among researchers, university managers, science policy-makers, and commercial academic publishers. Publishing in journals with a high IF is used as an indicator in the international rankings of universities (ibid., 4–5). Since 'most of the ostensibly "internationally significant" journals' (i.e. those in the Web of Science) are published by major publishing houses in the UK and USA, this maintains the global hegemony of English as 'a global synonym for "international"' (ibid., 3). Although the Web of Science has widened its range to include more journals outside the Anglophone world, journals published in non-English language countries tend to come at the bottom of the IF hierarchy. These journals thus underscore the (assumed) 'excellence of the very established, high impact factor journals coming from the solid core of Anglophone publishing businesses' (ibid., 7). Paasi argued that journals are in effect 'classified according to their impact factors and in practice "quality" is still related to the journal's position in this hierarchy' (Paasi 2013, 4), resulting in demands from research administrators specifying which journals researchers should publish in, that is those at the top of the Web of Science hierarchy.

The journal that I was formerly editor-in-chief for, *NGT–NJG*, was admitted to the Web of Science index in 2008. In the nature of things, the IF was low in the years immediately after *NGT–NJG* became an indexed journal. Hence, raising the journal's IF became a regular topic of discussion at the editors' meetings. It was suggested that review articles should be encouraged as they often had relatively high citation rates, indicating the usefulness of this type of article, although they give an overview of a research field rather than presenting the results of a particular, specialized research project. It was further suggested that the inclusion of the journal in the citation database was contributory to increased manuscript submissions. On the other hand, it was found that an increase in the size of the journal from four to five issues a year to allow publication of more articles appeared to be counter-productive as it led, given the same number of citations (A), to a lowering of the IF due to the increased number of citable articles (B). Another consequence of the increasing focus on IF was that a generalist geography journal such as *NGT–NJG*, which initially aimed to include a relatively even balance between the number of articles in physical geography and those in human geography, no longer received the same amount of submissions in physical geography as earlier. Physical geographers preferred to submit to specialist journals in geosciences with higher IFs. This phenomenon has also been noticed in other, more prominent generalist geographical journals, such as *Annals of the Association of American Geographers* and *Transactions of the Institute of British Geographers* (despite their higher IF compared with *NGT–NJG*). This appears to have occurred less in specialist fields within human geography (although Paasi (2013, 7) has provided evidence to suggest that it has occurred to some extent in the case of economic geography).

The underlying assumption is that journals with high IF will attract the best scholars and hence the highest quality research will be published in them. This assumption is based on the reputation of such journals. Yet this assumption is not accompanied by a genuine discussion of what constitutes research quality nor of any real assessment of the quality of articles published in journals with high IF compared with those published in journals with lower IF.

Academic 'volunteer work' – peer-reviewing and editing

Kirsi Pauliina Kallio, in a recent editorial in *Fennia*, pointed out 'that the process of publishing a referee journal article contains a significant amount of academic "volunteer work" by authors, editors and reviewers' (Kallio 2017, 2). To illustrate this voluntary work, she outlined 20 steps in the interaction between author, editor and reviewer in the process from manuscript submission to final publication. In the following I focus on the role of reviewers and editors.

A rigorous peer-review process by independent and unbiased fellow researchers is designed to ensure that the research is of sufficient quality to be worthy of publication. Peer reviews are qualitative and discursive rather than quantitative. They receive their legitimacy through intersubjective understandings among scholars of what makes for research quality, and involve often informal rules to ensure fairness within disciplines (Lamont & Guetzkow 2016, 31–32). Without reliable and fair peer reviews, the system of quality control of scholarly output would collapse. Peer-reviewing is a taken-for-granted part of academic work, and often forgotten in the allocation of funding and budgeting of time for teaching and research. Peer-reviewing is undertaken voluntarily and gains no merits in the form of publication points, yet the academic publishing system is entirely dependent on it.

Answers to a recent brief e-mail questionnaire survey that I conducted among colleagues at the Department of Geography in Trondheim indicate how peer-reviewing is perceived as a work task. The survey was sent to those in tenured academic positions. With one exception, the respondents replied that they had undertaken at least one and up to seven peer reviews during the first nine months of 2017. In deciding whether to accept an invitation to undertake a peer review, most replied that they primarily considered the title and abstract of the manuscript. Just under a third of the respondents said they generally accepted review invitations, although it depended for many on time available and work pressure otherwise. Several emphasized that their competence in the research field of the manuscript was decisive. The majority paid some or significant consideration to which journal the invitation came from, but in all cases the journal's IF was considered of little or no importance. One respondent replied that the focus on publication points resulted in more manuscripts circulating that

need peer reviews or re-reviews, adding to work pressure or resulting in some academics choosing not to review manuscripts because this work took time away from writing articles. Most respondents had clear criteria for what they emphasized as important in undertaking a peer review. Frequently mentioned criteria included originality, topicality, correspondence between research questions and research findings, methodological soundness, conceptual or theoretical foundation, soundness of arguments, good structure, clarity of language, reader friendliness, and referencing of relevant and up-to-date literature. Several referred in addition to the individual journals' guidelines. The responses indicate that the task of peer-reviewing is considered as a necessary part of academic work, it is taken seriously, and conducted systematically on the basis of qualitative criteria.[2]

Editing journals and books is another task that is voluntary and for which publication points are not awarded. Through my personal experience of having edited or co-edited eleven books on geographical and social-science topics, and five special issues of *NGT–NJG*, in addition to serving as the journal's editor-in-chief, I can confirm that this is time-demanding work. Increasing submissions meant that the work load of journal editing increased over time, and in my last year as editor-in-chief I spent more than 560 hours on the journal. In addition came the work put in by the co-editors. Neither the editor-in-chief nor the co-editors received direct remuneration. The editor's royalty went towards paying a part-time editorial assistant, whose work was invaluable. However, my work as editor-in-chief received recognition from the department in the form of reduced teaching obligations. Considerable free time was also used. The work of editor-in-chief involved contact with the publisher, reading and making decisions on manuscripts, allocating as appropriate submitted manuscripts to co-editors, overseeing special issues, following up deadlines, and corresponding with authors, reviewers, co-editors, and the journal's owner (the Norwegian Geographical Society). The journal is international – 55% of the submissions in 2014 came from outside Norway, the great majority from non-English-speaking countries. Even though responsibility lies with authors for ensuring that manuscripts have been language-checked and are in accordance with the journal's guidelines, considerable editorial work went into correcting language, improving style, reducing wordiness, fact-checking, and checking references. The less costly and less satisfactory alternative would be to accept a lower quality of writing. The work involved in editing and its contribution to ensuring research quality tends to be little understood by those who have not done it and is frequently little acknowledged. The Norwegian publication register does not even offer a category for journal editing. It is a paradox that the rhetoric of improving quality and publishing more gives so little thought to the role of editors, without whose work there would be no journals or anthologies to publish.

Difficult though it might be to explain to colleagues, there is nonetheless a certain satisfaction to be gained from editing, derived from facilitating imperfect but promising research manuscripts to reach publication and hence realize their qualitative potential. Australian geography professor and former journal editor Iain Hay (2015, 159) has summarized the challenges and paradoxes of journal editing as well as its professional and personal rewards as follows:

> Although journal editing is central to scholarly enterprise, helping to maintain academic standards and shape disciplines, it is frequently discouraged within the academic assemblages that depend on it. … Despite strong disincentives, journal editing offers valuable opportunities for self-development and deepening professional networks, as well as for refining the discipline.

Despite having a 'problematic place in "academic capitalism"' (Hay 2015, 159), peer-reviewing together with journal and book editing belong alongside writing book reviews, obituaries, commentaries and debate articles to the sphere of academic services that contribute to the dynamics of a scholarly research community. These activities can be regarded as part of a care ethic, in which members of the research community strive collectively to promote maximum quality, as an alternative to the individualistically oriented, competitive academic ethos of the neoliberal university.

Concluding remarks

Impact factors and the weighting of publications favour production of articles in international, peer-reviewed journals at the expense of time-consuming books, which receive relatively few points in

relation to their size and the work involved. Similarly, popular science and presentation of research to the general public through communication and interpretation receive only limited recognition. Further, writing book reviews is no longer given priority because book reviews do not give publication points; they are not valued – except by the readers. Moreover, publication points are not given to the work of peer-reviewing and editing, which provides an important guarantee of quality. Without peer reviewers and editors, articles cannot be published in trustworthy academic journals.

The number of publication points is determined by whether an article is published in a first-level or second-level journal. Yet the process of deciding what is a top-level journal is determined by impact factors and negotiation rather than any real quality assessment. A high IF is claimed to be a mark of a journal's quality, yet it is calculated in such a way that, if a new or expanding journal increases the number of articles published in a year, the IF goes down. The reliance on IF as a measure of quality favours the long-established and most well-known English-language journals, which are mostly published by large international commercial publishers.

Pressure to publish, stimulated by the publication points system, does not provide a guarantee that the published research is of high quality, and might even be detrimental to achieving the best quality. Market indicators can tell something about the quantity and popularity of the products being sold, but not their quality as such. The quasi-market indicator of IF appears to rest on the false assumption that popularity attracts and reinforces the best research. IF is an indicator that was invented by and continues to serve commercial interests in the dominant English-speaking world. Despite attempts to widen its use to non-English language journals, its practice remains discriminatory to publication in other languages than English. In many cases, this also applies to publication of specialized research in regions far from the major world centres of research activity. Such research is not infrequently considered to be only of 'local' interest regardless of whether it is published in English or not and regardless of the quality of the research. That an article has a potentially high number of readers by being published in a journal with a high IF is no guarantee of quality in itself. The guarantee is only provided by a properly functioning peer-review system.

Measuring quality by quantitative measures ends up as measurement of quantity rather than quality. Quality is poorly amenable to quantification but requires critical reflection. Critical scholarship should be the hallmark of the university, but it can be questioned whether critical reflection and scholarship are best served by the system of control and auditing by numerical performance indicators that is favoured in new public management. Control mechanisms and auditing require increasing managerial resources. Critical scholarship requires on the other hand autonomous scholars working in an atmosphere of trust and democratic debate.

Quality and quantity are two different things and one cannot logically be substituted for the other. Quality is expressed through verbal discourse, while quantity is expressed by numbers and statistics. Furthermore, numbers require interpretation though qualitative discussion. I argue that it is a contradiction in terms, indeed a fallacy, to act as if research quality can validly be expressed by numerical measures.

NOTES

[1] The publication indicator is more commonly known in Norwegian as the *tellekant* system. The term *tellekant* (literally 'counting edge') is derived from the clothing business, where piled items of folded clothing can easily be counted when the folds or edges are neatly placed on top of one another.

[2] The e-mail survey was undertaken between 25 September and 5 October 2017. Questions were circulated to 16 colleagues and responses received from all 16.

References

Aagaard, K, Bloch, C. & Schneider, J. W. (2015) Impacts of performance-based research funding systems: the case of the Norwegian publication indicator. *Research Evaluation* 24, 106–117. https://doi.org/10.1093/reseval/rvv003

Aagaard, K., Bloch, C., Schneider, J. W., Henriksen, D., Ryan, T. K. & Lauridsen, P. S. (2014) *Evaluering af den norske publiceringsindikator*. Dansk Center for Forskningsanalyse, Aarhus Universitet, Aarhus. <https://npi.nsd.no/dok/eval2014/Evaluering_af_den_norske_publiceringsindikator_2014.pdf> 2.11.2017.

Aksnes, D.W. (2017) Artikler I nivå 2-tidsskrifter blir mest sitert. Forskerforum. 5.10.2017. <http://www.forskerforum.no/artikler-i-niva-2-tidsskrifter-blir-mest-sitert/> 2.11.2017.

Andersen, H., Ariew, R., Feingold, M., Bag, A. K., Barrow-Green, J., Dalen, B., Benson, K., Beretta, M. Blay, M., Bleker, J., Borck, C., Bowker, G., Leigh Star, S., Buccianti, M., Camerota, M., Buchwald, J, Gray, J., Cappelletti, V., Cimino, G., Carson, C., Clark, M., Keller, A., Cline, R., Clucas, S, Gaukroger, S., Cook, H., Hardy, A., Corry, L., Metraux, A., Renn, J., Dolan, B., Luckin, B., Duerbeck, H., Orchiston, W., Epple, M., Hård, M., Rheinberger H-J., Roelcke, V., Farber, P., Fissell, M., Packard, R., Fox, R., Frasca Spada, M., French, S., Good, J., Hackmann, W., Hllieux, R., Holmqvist, B., Home, R., Hoskin, M., Inkster, I., Jardine, N., Levere, T., Lightman, B., Lüthy, C., Lynch, M., McCluskey, S., Ruggles, C., Morris, P., Rhys Morus, I., Nelson, E. C., Perez, L., Rigden, J., Stuewer, R. H., Samsó, J., Schaffer, S., Schappacher, N., Staudenmaier SJ, J., Strom, C., Unschuld, P., Weingart, P., Zamecki, S. & Zuidervaart, H. (2009) Journals under threat: a joint response from history of science, technology and medicine editors. *Centaurus* 51 1–4. https://doi.org/10.1111/j.1600-0498.2008.00140.x

CRIStin [Det nasjonale forskningsinformasjons-systemet] (2017) *Reporting of academic publications in the health, institute and HE sectors*. 28.3.2017. CRIStin – Current Research Information System in Norway, Oslo. <http://www.cristin.no/english/resources/reporting-instructions/> 1.11.2017.

Fossum-Raunehaug, S. (2017) *Publication points and reward of publications at level 1 and 2*. 3.7.2017. NMBU – Norwegian University of Life Sciences, Ås. <https://www.nmbu.no/en/research/for_researchers/publishing-abc/node/25300> 2.11.2017.

Gjengedal, K. (2017) Kvalitet er meir enn siteringar. *Forskerforum* 49(8) 6–7.

Halffman, W. & Radder, H. (2015) The academic manifesto: from an occupied to a public university. *Minerva* 53(2) 165–187. http://dx.doi.org/10.1007/s11024-015-9270-9

Halffman, W. & Radder, H. (eds.) (2017) *International responses to the Academic Manifesto: reports from 14 countries*. Social Epistemology Review and Reply Collective, Special Report 2017. <http://wp.me/p1Bfg0-3FV> 2.11.2017.

Hansen, T. I. (2015) Tellekantsystemet. In Store norske leksikon. 20.2.2015. <https://snl.no/tellekantsystemet> 2.11.2017.

Haugstad, B. (2017) Om Aarebrot og tellekanter. Khrono. 27.9.2017. <https://khrono.no/debatt/snodig-om-aarebrot-og-tellekanter> 2.11.2017.

Hay, I. (2015) 'Why edit a scholarly journal? Academic irony and paradox. *The Professional Geographer* 68(1) 159–165. https://doi.org/10.1080/00330124.2015.1062704

Kallio, K. P. (2017) Subtle radical moves in scientific publishing. *Fennia* 195(1) 1–4. https://doi.org/10.11143/fennia.63678

Kjelstadli, K. (2010) *Akademisk kapitalismen*. Forlaget Res Publica, Oslo.

Kunnskapsdepartementet (2016) *Orientering om statsbudsjettet 2017 for universitet og høgskolar: etter vedtak i Stortinget 17. desember 2016: Mål for universitet og høgskolar, budsjett og endringar i løyving og finansieringssystemet*. Kunnskapsdepartementet, Oslo. <https://www.regjeringen.no/contentassets/31af8e2c3a224ac2829e48cc91d89083/orientering-om-statsbudsjettet-2017-for-universiteter-og-hoegskolar_ny-versjon160217.pdf> 2.11.2017.

Lamont, M. & Guetzkow, J. (2016) How quality is recognized by peer review panels: the case of the humanities. In Ochsner, M., Hug, S. E. & Daniel, H-D. (eds.) *Research assessment in the humanities: towards criteria and procedures*, 31–41. Springer Open. https://doi.org/10.1007/978-3-319-29016-4_4

Lund, R.W.B. (2015) Doing the ideal academic: Gender, excellence and changing academia. Doctoral dissertations 98/2015. Aalto University, Helsinki.

Myklebust, J. P. (2017a) In search of a new form of university governance. University World News 450. 10.03.2017. <http://www.universityworldnews.com/article.php?story=20170309180941360> 7.9.2017.

Myklebust, J. P. (2017b) Should universities be run like businesses? University World News 473. 8.9.2017. <http://www.universityworldnews.com/article.php?story=20170908102945748> 7.9.2017

NSD [Norsk senter for forskningsdata] (2017) Register over vitenskapelige publiseringskanaler: kriterier for godkjenning av publiseringskanaler. NSD – Norsk senter for forskningsdata, Bergen. <https://dbh.nsd.uib.no/publiseringskanaler/OmKriterier> 2.11.2017.

Ochsner, M., Hug, S. E. & Daniel, H-D. (2016) Research assessment in the humanities: Introduction. In Ochsner, M., Hug, S. E. & Daniel, H-D. (eds.) *Research assessment in the humanities: towards criteria and procedures*, 1–10. Springer Open. https://doi.org/10.1007/978-3-319-29016-4_1

Østerud, Ø. (2009) Forord. In Østerud, Ø. (ed.) *Hvordan måle vitenskap? Søkelys på bibliometriske metoder*, 5–7. Det Norske Videnskaps-Akademi – Novus forlag, Oslo. <http://www.dnva.no/binfil/download.php?tid=41358> 2.11.2017.

Paasi, A. (2013). Fennia: positioning a 'peripheral' but international journal under conditions of academic capitalism. *Fennia* 191(1) 1–13. https://doi.org/10.11143/7787

van Reekum, R. (ed.) (2015) The new university: a special issue on the future of the university. Krisis 2015(2). <*http://krisis.eu/the-new-university/*> 2.11.2017.

Roll-Hansen, N. (2009) Om å "måle" kvalitet av forskning. In Østerud, Ø. (ed.) Hvordan måle vitenskap? Søkelys på bibliometriske metoder, 71–80. Det Norske Videnskaps-Akademi – Novus forlag, Oslo. <*http://www.dnva.no/binfil/download.php?tid=41358*> 2.11.2017.

Sandnes, F. E. (2016) Hvordan melke nye tellekanter. Khrono 27.4.2016. <*https://khrono.no/debatt/hvordan-melke-tellekanter-i-2016*> 2.11.2017.

Seglen, P. O. (1997) Why the impact factor of journals should not be used for evaluating research. *BMJ* 314, 498–502. https://doi.org/10.1136/bmj.314.7079.497

Seglen, P. O. (2009) Er tidsskrift-renommé og artikkeltelling adekvate mål for vitenskapelig kvalitet og kvantitet? In Østerud, Ø. (ed.) *Hvordan måle vitenskap? Søkelys på bibliometriske metoder*, 39–70. Det Norske Videnskaps-Akademi – Novus forlag, Oslo. <*http://www.dnva.no/binfil/download.php?tid=41358*> 2.11.2017.

Sivertsen, G. (2009) Publiseringsindikatoren. In Østerud, Ø. (ed.) *Hvordan måle vitenskap? Søkelys på bibliometriske metoder*, 11–37. Det Norske Videnskaps-Akademi – Novus forlag, Oslo. <*http://www.dnva.no/binfil/download.php?tid=41358*> 2.11.2017.

Sivertsen, G. (2016) Publication-based funding: The Norwegian model. In Ochsner, M., Hug, S. E. & Daniel, H-D. (eds.) *Research assessment in the humanities: towards criteria and procedures*, 79–90. Springer Open. https://doi.org/10.1007/978-3-319-29016-4_7

Sjøberg, S. (2017) Null poeng til Aarebrot? Khrono. 19.9.2017. <*https://khrono.no/debatt/null-poeng-til-frank-aarebrot*> 2.11.2017.

UHR [Universitets- og høgskolerådet] (n.d.) Publiseringskanaler. Universitets- og høgskolerådet, Oslo. <*http://www.uhr.no/rad_og_utvalg/utvalg/det_nasjonale_publiseringsutvalget/publiseringskanaler*> 2.11.2017.

Waarenperä, U. (ed.) (2011) *Universitetsrankning och bibliometriska mätningar: konsekvenser för forskning och kunskapsutveckling*. Konferenser 74. Kungl. Vitterhets Historie och Antikvitets Akademien, Stockholm. <*https://vitterhetsakad.bokorder.se/sv-SE/article/2103/universitetsrankning-och-bibliometriska-matni*>