Academic sell-out: how an obsession with metrics and rankings is damaging academia

This item was submitted to Loughborough University's Institutional Repository by the/an author.


Additional Information:

- This is an Accepted Manuscript of an article published in the Journal of Marketing for Higher Education on 23 Oct 2014, available online: http://www.tandfonline.com/10.1080/08841241.2014.970248

Metadata Record: https://dspace.lboro.ac.uk/2134/16336

Version: Accepted for publication

Publisher: © Taylor and Francis

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Academic sell-out: how an obsession with metrics and rankings is damaging academia

by

Thorsten Gruber
Centre for Service Management (CSM)
Loughborough University

Final version published in
Journal of Marketing for Higher Education, Vol. 24 No 2, pp. 165-177
Academic sell-out: how an obsession with metrics and rankings is damaging academia

Increasingly, academics have to demonstrate that their research has academic impact. Universities normally use journal rankings and journal impact factors to assess the research impact of individual academics. More recently, citation counts for individual articles and the $h$-index have also been used to measure the academic impact of academics. There are, however, several serious problems with relying on journal rankings, journal impact factors and citation counts. For example, articles without any impact may be published in highly ranked journals or journals with high impact factor, whereas articles with high impact could be published in lower ranked journals or journals with low impact factor. Citation counts can also be easily gamed and manipulated and the $h$-index disadvantages early career academics. This paper discusses these and several other problems and suggests alternatives such as post-publication peer review and open-access journals.

Keywords: metrics; journal impact factor; journal ranking; citations; open access; altmetrics; $h$-index; academia

“Science is of course about discovery, about digging to discover the truth. But it is also communication, persuasion, marketing. I am a salesman.” (Bhattacharjee, 2013)

For the last three years, I have been an Associate Editor (AE) for the Journal of Marketing for Higher Education (JMHE) and I am very grateful to the two co-editors, Anthony Lowrie and Jane Hemsley-Brown, for giving me this wonderful opportunity. To mark the end of my three-year term as AE, I would like to share some thoughts with you. I always thought the aim of researchers was to discover new things and to create knowledge by presenting and publishing quality research. For some time, however, I have been observing an, in my view, unhealthy trend of institutions (and individuals) becoming increasingly obsessed with journal metrics and, recently, article-level metrics. I do not know where this obsession with measuring, the urge to quantify everything comes from. Authors such as De Angelis and Harvie (2009) compare it to the times of Frederick Taylor who used clipboards and stopwatches to measure performance in US factories. Maybe it has something to do with having a strong neo-liberal public policy regime as suggested by Holmwood (2013) or is a sign of the increasing business orientation of
higher education (Gruber, Reppel, & Voss, 2010). As early as 2003, the JMHE editorial board member Hugh Willmott wrote about the commercialization of higher education.

I have therefore chosen the term “academic sell-out” as the title of my commentary paper: Like music bands that change their musical direction and give up their values to pursue commercial success, academics are increasingly “selling” their publications not only to disseminate research findings but to attain other goals set by their institutions as well. In my view, the recent focus on article citation counts and article download numbers will increase this tendency even more. Like music artists who have to sell singles and albums (nowadays mp3 downloads) to keep their record labels happy, academics will have to “promote” their work to attract viewers and citations to keep their institutions happy. Publishing in highly ranked journals and/or journals with high impact factors and accumulating as many citations as possible are important “means” for academics to improve their own and their institution’s reputation and to advance their career (“ends”). Following these incentives is therefore totally rational, but probably not in the best interest of academia, society and humanity (Schekman, 2013). Of course, I have to admit that I have done the same and “played the game”. Maybe at this point of my career it is a good time to use the opportunity of a commentary paper to make readers aware that this game has some serious drawbacks as well.

The role of academic impact

In the past, academics would focus on conducting studies in the research areas they were interested in, publish results and then move on to the next project. Increasingly, academics here in the UK, but presumably in other countries as well, have to demonstrate that their research has academic impact as well as economical, societal and/or cultural impact. My commentary paper will focus on the former. Research is said to have an academic impact “when the influence is upon another academic author, researcher, or university organization” (LSE Public Policy Group, 2011, p. 11).

Citation indicators are usually used to measure academic impact. As a consequence, in the new metrics reality, academics have to carefully plan projects and actively influence citations of their good work. Citations may soon play an important role in recruitment, academic probation/tenure and promotion decisions. They are already important for league tables here in the UK and some sub-panels already used citation data in the current 2014 Research Excellence Framework (REF), which evaluates the quality of research conducted in UK universities. For this purpose, each academic had to make four submissions (in most cases, published journal articles) for the time-period 2008-2013. During this year, an expert panel has been reviewing
and grading (from 1 to 4, with 4 being the highest) these submissions. The results will be announced at the end of the year and will inform the selective research funding allocation for the participating institutions. The results will also be used to develop league tables of UK higher education institutions. It is likely that citation data will be used in REF 2020 for the unit of assessment “business and management” as well.

At the moment, universities normally use journal rankings and journal impact factors to assess the research impact of individual academics. Common journal rankings are the Academic Journal Quality Guide (Association of Business Schools (ABS), UK), the Australian Business Deans Council list (ABDC, Australia) and the VHB-Journal (Verband der Hochschullehrer für Betriebswirtschaft, Germany). The best known and most widely used journal impact factor data are provided by Thomson Reuters (Journal Citation Report®) and Elsevier (SCImago Journal Rank).

In the following, I will discuss how an obsession with journal rankings, journal impact factors and citation counts is damaging academia. Following that, I will then suggest some alternatives.

**Obsession with journal rankings and luxury journals**

There is an obsession with journal rankings, at least here in the UK, and Cluley (2014) and Willmott (2011) go so far as to call it “journal list fetishism”. Even Nobel prize winners such as Randy Schekman (2013) point to the problem of taking the “place” of publication as a proxy for the “quality” of the conducted research. Schekman (2013) especially draws attention to the fact that relying on top journals, which he calls “luxury-journals”, is not enough as they are not the only ones that publish outstanding research. They are “brands” that, similarly to fashion designers who sometimes sell limited editions of suits or handbags, artificially keep journal space low as they know that “scarcity stokes demand”. The same phenomenon can be observed in other areas of life too: for example, whenever Apple® launches the newest version of its iPhone®, long queues can be witnessed outside their stores. Of course Apple® is aware of the popularity of its phone and could stock more products, but the strong visual power of people queuing (and sometimes even camping) outside shops sends a strong message to other customers: come and join the queue as the phone must clearly be worth the long wait.

Parker (2014) describes a European business school (he does not reveal the name but it becomes clear that he must be referring to a well-known UK business school) that puts a strong focus on the ABS Academic Journal Quality Guide (latest version at the time of writing dates from 2010) that ranks journals in four categories (grades 1 to 4, with 4 being the best). Parker
(2014, 285-286) explains how appointments were made “by counting the ABS scores on publications”, focusing only on publications graded 3 and 4 and excluding books and book chapters. He then remarks that “there was seemingly no interest in evidence of skill in teaching, or of collegiality and experience in holding responsibility, and combative performance management interviews with the Dean underlined the point that ABS was all that mattered”. The problem is that all journal rankings have their flaws, the scores for individual journals can differ significantly between ranking (e.g., *Industrial Marketing Management* is a grade 3 in the ABS list (second highest category), an A* in the ABDC list (highest category) but only rated C (fourth highest category) in the VHB ranking), and that there are no agreed journal rankings. Further, highly cited and potentially influential articles could be published in low ranked journals, and articles without any citations and impact on the scientific society at all could be published in highly ranked journals.

**Obsession with journal impact factors**

Recently, Thomson Reuters released the latest version of its Journal Citation Reports® that lists and ranks academic journals according to their impact factors. Every year in June/July, journal editors become anxious around that time as the new report shows whether their journals went up or down this year. A few days later, the new impact factors would then already appear on the journal websites.

Curry (2012), in his excellent blog called “Reciprocal Space”, discusses the abuse of journal impact factors to evaluate individual articles and academics. He points out that the annual impact factor, defined as the “mean number of citations to articles published in any given journal in the two preceding years” originally had a good purpose (to help librarians decide which journals to subscribe to). However, as early as the 1990s, authors such as Seglen (1997) showed that citation rates are significantly skewed (i.e., 85% of published articles have fewer citations than average). In Curry’s (2012) view, the situation got worse when this “statistically indefensible indicator” of the performance of journals, this “mis-measure”, started to be applied to individual academics and their publications. Curry (2012) is clearly irritated by what he calls “malady” and the apparent addiction to a statistically worthless measure. He complains about academics being dependent on a valuation system that is relying on a false indicator: “We spend our lives fretting about how high an impact factor we can attach to our published research because it has become such an important determinant in the award of the grants and promotions needed to advance a career. Instead of relying on journal impact factors, that Curry (2012) wants
to stigmatize with a smear campaign in the same way cigarettes have been in the past, he suggests that academics should value articles that are used and cited.

Similarly, Schekman (2013) stresses the fact that the impact factor is an “average” and therefore cannot tell us anything about the quality of an individual article. He even goes so far as to say that the focus on impact factors “is as damaging to science as the bonus culture is to banking”. The former Editor-in-Chief of the highly regarded *Science*, Bruce Alberts (2013), shares Curry’s and Schekman’s concerns and regards the misuse of the journal impact factor as “highly destructive”.

Finally, Lawrence (2007, R. 584) believes that the focus on getting papers published in journals with high impact is “turning our thoughts and efforts away from scientific problems and solutions, and towards the process of submission, reviewing and publication”. For him, possible negative consequences are that academics follow fashions and work in already well-established areas so that they know that there are sufficient colleagues to notice their work (and cite it). Venturing into new areas, however, would be risky as chances are that no fellow colleagues would have a shared interest and would not cite their work. Similarly, Alberts (2013) believes that academics would be discouraged from carrying out risky but possibly ground-breaking studies. Academics would rather stay in already highly populated research areas, which would then lead to more “me-too science”.

Lawrence (2007) also mentions the negative impact the focus on metrics has on the behavior of academics who would feel tempted to hype their work, slice up the findings as much as possible (publish several small papers instead of one big one), come up with simple conclusions but at the same time complexify the material (to make it difficult for reviewers to find faults in it) and may even ignore findings that do not fit the story they are trying to “sell”.

**Obsession with citation counts**

More recently, there seems to have been a shift from journal to article level metrics. Citations of *individual* articles are increasingly being used to measure the impact of an academic’s work (Woodside, 2009; Li, Sividas, & Johnson, 2014). Woodside (2009, p. 4) wants to “redirect the focus from relying on journal impact factors in evaluating the quality and quantity of an academic’s scholarly contribution and to focus thinking on the candidate’s scholarly impact”. The idea is that an article has impact if other academics find it valuable enough to cite it in their work. Google Scholar already creates rankings of academics based on their citation scores. For each user profile, Google Scholar allows several tags (e.g., in my case, Marketing; Service Management; Service Marketing; Service Quality; and Service Innovation). Clicking on a tag...
leads then to a ranking of academics (based on total number of citations) who used the same tag on their profile.

Stremersch, Verniers, and Verhoef (2007), however, show that the numbers of citations an article attracts does not only depend on its quality but on a number of others factors as well: the article’s domain subject area (e.g., articles on e-commerce, relationship marketing, and services marketing tend to be cited more than other articles. By contrast, articles on advertising, consumer knowledge, and sales tend to be cited less than other articles), the author’s publication record, editorial board membership, business school ranking, and personal promotion.

Further, Li et al.’s (2014) recent study reveals that citations seem to attract more citations. They show the existence of the so-called Matthew effect not only for famous authors but also famous papers. The term Matthew effect, which refers to the Gospel of Matthew, the first book of the New Testament, was introduced by the sociologist Robert Merton in 1968, who showed that already well-known scientists will often get more credit for their work than their not-so-famous colleagues. Li et al. (2014) also point to the importance of attracting citations quickly to papers: “if a paper is not cited early, while it may not go uncited in the future, it may result in low relational in scholar’s mind, and hence its long-term influence is in great doubt”.

Moreover, Colquhoun (2014) points to several issues with relying on citation counts:

- High citations could be the result of a good or a bad paper. The number itself does not tell readers anything about the quality of the paper. Fellow academics may cite an article because they appreciate the content or because they criticize the content. Articles may also be cited to please reviewers or editors (Hofacker, Gleim, & Lawson, 2009).
- Citation counts should not be used for evaluating recent work as it takes a long time for citations to accumulate, which is especially damaging for young researchers and academics taking career breaks.
- Citation scores differ significantly between the different providers (Scopus®, Web of Science and Google Scholar).
- Citation scores can be gamed. For example, review articles attract more citations than original work. In this connection, McPeek (2012) who was the Editor-in-Chief of the American Naturalist, shows how journals can game the system by publishing not only review papers but also methods papers. If the article describes a method that a lot of academics use, it will attract a lot of citations. For example, the second most cited article in the Journal of the Academy of Marketing Science for the period 2009-2013 is a paper on Partial Least Squares Structural Equation Modeling (Hair, Sarstedt, Ringle, & Mena, 2012) that was published in 2012 and already has 321 citations according to Google
Scholar. Another paper by the same authors published a year earlier (Hair, Ringle, & Sarstedt, 2011) is the most cited paper in *Journal of Marketing Theory and Practice* with currently 664 citations (by comparison, the next most cited paper in the journal in the same year has 33 citations).

- Citation scores can easily be manipulated. For example, Delgado López-Cózar, Robinson-García, and Torres Salinas (2012) show how academics can easily manipulate Google Scholar by uploading false documents that contain citations to their own work. The authors show how they were able to generate 774 citations in 129 papers and thereby increased the journals’ and the authors’ *h*-index significantly.

Another problem is that the person who cites an article may not have read it but may have just taken the citation from another article (e.g., a review article). For example, Simkin and Roychowdhury (2003) used stochastic modeling of the citation process to estimate that only about 20% of academics who cite articles actually read the original. Further, Wright and Armstrong (2008) analyzed citations to one of the most cited articles in marketing (Armstrong & Overton, 1977) and found that in 49 of the 50 studies they examined, the findings were reported incorrectly, indicating that the citers did not read the original paper (or did not understand it properly).

Finally, Schekman (2013) points to a few other problems of relying on citation counts: papers may get cited because they are covering sexy topics, are eye-catching, provocative or just plainly wrong.

**Using the *h*-index to compare the performance of academics**

The *h*-index is used to compare the impact of individual academics quantitatively. The metric is named after Jorge E. Hirsch who introduced it in 2005. Hirsch (2005, 16569) suggests that a “scientist 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 ≤ *h* citations each”. In simpler terms, it is also defined as “the largest number *h* such that *h* publications have at least *h* citations” (Google Scholar). For example, Google Scholar automatically calculates the *h*-index for academics with a Google Scholar profile. However, all the problems mentioned before with regard to citations also apply to the *h*-index as it is based on citations.

Ball (2012) also points to the danger of creating “league tables” based on the *h*-index. For example, Sir Harold Kroto, whose research clearly had significant impact (he won the Nobel Prize in Chemistry in 1996) is only in 264th position in Chemistry on the *h*-index. One of the
main problems I have with this metric is that it disadvantages early career researchers, who have not had the time to publish a large number of articles that could then have accumulated a number of citations (Colquhoun, 2014). As a consequence, they would not rank highly in any h-index based league tables.

A new obsession – the rise of altmetrics
There is also an increased use of more comprehensive data sources, altmetrics, that include views (HTML views and PDF downloads), discussions (journal comments, science blogs, Wikipedia, Twitter, Facebook and other social media) and bookmarks (e.g., Mendeley).

However, altmetric scores do not necessarily measure the impact of one’s work but attention (Crotty, 2013). For examples, articles with trendy titles or hyped-up topics attract attention (Colquhoun, 2014) but they are not necessarily quality articles. Moreover, altmetrics offer several opportunities for gaming, as for example retweets and Facebook Likes can be purchased for just a few dollars (Davis, 2012; Colquhoun, 2014). Colquhoun and Plested (2014) regard them as “one of the silliest metrics so far proposed”. The authors go on to say that “Altmetrics is the latest buzzword in the vocabulary of bibliometricians. It attempts to measure the “impact” of a piece of research by counting the number of times that it’s mentioned in tweets, Facebook pages, blogs, YouTube and news media. That sounds childish, and it is” (Colquhoun & Plested, 2014). Especially the use of altmetrics for hiring purposes is a terrifying idea for them.

Similarly, Crotty (2013) is concerned that a “Silicon Valley/internet startup mindset” is entering academia where “attention and popularity of a product seems more important than the actual value it generates”. I agree with Crotty (2013) that this “Silicon Valley/internet startup mindset” is indeed dangerous for academia as attention and popularity are not necessarily the same as quality research.

The dark side of metrics – fraudsters and sell-outs
For Lawrence (2007), metrics distort behavior and decide careers. With the increasing reliance on metrics, he especially expects “citation-fishing and citation-bartering” (R.583) to become major activities of academics. Metrics also harm science by encouraging researchers to focus on eye-catching research. As outlined before, metrics are also prone to gaming. Worst of all, however, metrics may even encourage fraud (Colquhoun & Plested, 2014). I agree with Colquhoun (2014) and Lawrence (2007) that metrics change individuals’ behavior in a negative way and therefore become unfit for their stated purpose. In this context, Colquhoun (2014) refers to Goodhart’s law ("When a measure becomes a target, it ceases to be a good measure"),
named after the banker Charles Goodhart who coined it, and uses it to show that metrics should not be extended to science.

The quotation at the beginning of this commentary paper is from Diederik Stapel, the former Dean of the School of Social and Behavioral Sciences at Tilburg University, Netherlands. Stapel was a social psychologist and an academic superstar until it was revealed that he had manipulated and fabricated data for more than 50 of his publications, several of them in top journals and 10 PhD dissertations as well. In an interview with the New York Times, Stapel “described his behavior as an addiction that drove him to carry out acts of increasingly daring fraud, like a junkie seeking a bigger and better high.” (Bhattacharjee, 2013). His ambition, his need to perform again and again and publish in the top journals seemed to have driven him to such behavior. Schekman (2013) also points to the lure of luxury journals that “encourage the cutting of corners, and contribute to the escalating number of papers that are retracted as flawed or fraudulent”.

Even if academics are not engaging in fraudulent activities to increase their number of publications and/or citations scores, I still think that one major drawback of metrics is the increasing pressure to “sell” our work. As mentioned before, it used to be that academics would conduct studies, publish the results and then move on to the next project. However, in the new reality, having published an article is not the “end” but the “beginning” of an extensive marketing campaign: tweeting and/or blogging about the latest publication, putting a link to the article in the email signature, shooting a video or recording a podcast highlighting the key findings of the article, and putting the article (pre-copyright form version) in university repositories and on platforms such as researchgate and academia.edu. All done in the hope of creating awareness that then (hopefully) will lead to downloads and citations. Of course I am aware of the fact that there was always a selling component in academia, but the recent metrics development is increasing this component too much, leading to a “sell-out”. The quotation from the beginning of this commentary paper continues with Stapel saying: “I am on the road. People are on the road with their talk. With the same talk. It’s like a circus.” […] “They give a talk in Berlin, two days later they give the same talk in Amsterdam, then they go to London. They are traveling salesmen selling their story.” (Bhattacharjee, 2013).

Even though that kind of behavior should come naturally to a marketing professor like myself, I still believe that it turns us academics into full-time marketers and the time it takes to do all the promotion could better be invested in research and/or teaching activities instead. I am concerned that like Krüss’ (1962) novel character Timm Thaler who sold his unique gift, his highly desirable laughter, to the joyless Baron Lefuet (German for “Devil”, spelled backwards)
in exchange for the ability to win any bet he places, we academics are in danger of “selling” our virtues in exchange for our obsession with metrics.

**The way forward**

First of all, I strongly agree with Li et al. (2014) that journal impact factors or journal rankings should *not* be used to measure the impact of individual articles, as articles without any impact may be published in highly ranked journals or journals with high impact factor, whereas articles with high impact could be published in lower ranked journals or journals with low impact factors.

I also share the view of authors such as Fenner and Lin (2013) that the quality of research cannot be measured or fully captured in numbers. The same applies to all creative work. For example, coming back to my music analogy from the beginning of this article, it would not be wise to say that the more records artists sell, the higher the “quality” of their music is. We all know that a lot of records that sold millions of copies are of rather dubious quality, to say the least. I therefore agree with Colquhoun and Plested (2014) that “if you want to know about the quality of a paper, you have to read it”. Consequently, we should be very cautious about journal rankings, impact factors and altmetrics and read articles instead. That is what the REF panels here in the UK will be doing: “No sub-panel will make any use of journal impact factors, rankings, lists or the perceived standing of publishers in assessing the quality of research outputs” (REF 2014).

My only recommendation to REF panels that goes beyond current practice would be to read articles without *any* author, journal and affiliation information, as this could bias their judgments (consciously or subconsciously). The same should apply to journal submissions that should be completely anonymous as well. Not even the editor should know who the author is and at which institution (s)he is. Then only the quality (or the lack of it) of the written word would matter to make quality judgments. I know that the JMHE co-editor Anthony Lowrie feels strongly about this issue too.

Further, in order to resist the lure (and the temptation to “cut corners”) of having to publish in prestigious journals, Schekman (2013) recommends publishing in open-access journals as they do not have artificial space limitations like “luxury journals” and also do not have the pressure to sell expensive subscriptions to libraries. I think that open-access journals could indeed be a feasible alternative to traditional journals, as long as they have an adequate review system in place. However, publishing in freely available open-access journals is only then a feasible option if recruitment and promotion panels at universities indeed read the published
articles to judge their quality instead of just relying on journal rankings and/or journal impact factors.

To help panels with making quality judgments (not all panel members may be familiar with the research area of the job applicant or the academic going for promotion) but also to provide fellow academics (especially early career academics and PhD students) with an overview of good publications, a service such as “Faculty of 1000” could be introduced in business and management: Faculty of 1000 Prime (http://f1000.com/prime) is a post-publication peer review service offered by the Science Navigation Group that consists of 5,000 faculty members (senior scientists, leading experts and associates) covering 40 disciplines and approximately 3,500 journals in all areas of biology and medicine. The Faculty recommends the most important articles, they rate them and provide short explanations for their selections. The Faculty rates publications on the articles’ merits rather than relying on journal impact factors or journal rankings. The Science Navigation Group also offers Faculty of 1000 Research, which is an Open Science journal that offers fast, open, post-publication peer-reviewed research, with underlying datasets. Finally, there is Faculty of 1000 Posters, an open access repository for free deposition of slide presentations and conference posters.

I think that introducing such services to the business and management disciplines would be a way forward and clearly more helpful than relying on simplistic metrics. Otherwise,

we are guilty of facilitating a culture which will be detrimental to the future of research and the individuals sucked into the whirlpool of impact factors, citations and claims of ‘my h factor is bigger than yours’. We haven’t yet quite reached the position where we are solely judged on such dodgy numbers. Let’s make sure we never do (Donald, 2013).

Finally, I strongly believe that senior academics in particular should create awareness of the dangerous developments with metrics and rankings in academia. We cannot expect early career academics to do that as they are under increasing pressure to “perform” and have to “play the game” to secure their jobs and to advance their careers. It is only rational and they cannot be blamed for that.

I genuinely hope that my commentary paper will make a small contribution and will raise awareness in our business and management community. I am also planning to give a seminar on this topic at universities to create further awareness, start a critical discourse about the danger of metrics and rankings, and to critically discuss the way forward. I am fully aware that it will probably take a long time before a change of attitude can be observed in universities (if ever at all) but I hope that we will eventually be able to break the dangerous obsession with these unhealthy metrics and rankings.
References


REF (2014). All FAQs about the Research Excellence Framework (REF) - How will journal impact factors, rankings or lists, or the perceived standing of publishers be used to inform the assessment of research outputs? Retrieved from: http://www.ref.ac.uk/faq/all/


