

We need informative metrics that will help, not hurt, the scientific endeavor – let's work to make metrics better.

 blogs.lse.ac.uk/impactofsocialsciences/2015/10/08/we-need-informative-metrics-how-to-make-metrics-better/

10/8/2015

Rather than expecting people to stop utilizing metrics altogether, we would be better off focusing on making sure the metrics are effective and accurate, argues [Brett Buttlere](#). By looking across a variety of indicators, supporting a centralised, interoperable metrics hub, and utilizing more theory in building metrics, scientists can better understand the diverse facets of research impact and research quality.



In recent years, it has become popular to bash [bibliometrics](#), and especially [journal level](#) metrics. Jane Tinkler, who was on the panel for the [Independent Review of the Role of Metrics in Research Assessment and Management](#), put it well when [she said](#), 'One of the most common concerns that colleagues discussed with us is that impact metrics focus on what is measurable at the expense of what is important.' Previously on the LSE Impact Blog, sociologist Les Back has suggested that [any attempt to measure academic thought is ill fated](#).

With such comments, I feel obliged to stress three small things, the first of which is that:

- 1. Metrics are necessary and good (but can be bad).**

While it is true, as Curt Rice says, that '[systems based on counting can be gamed](#),' science has always been about taking the vast complexity of the world and making it measurable and predictable. Yes, metrics and statistics sometimes miss nuance, but the goal of the metric is to maximally explain cases, *on average*, rather than at the individual level (e.g., regression).

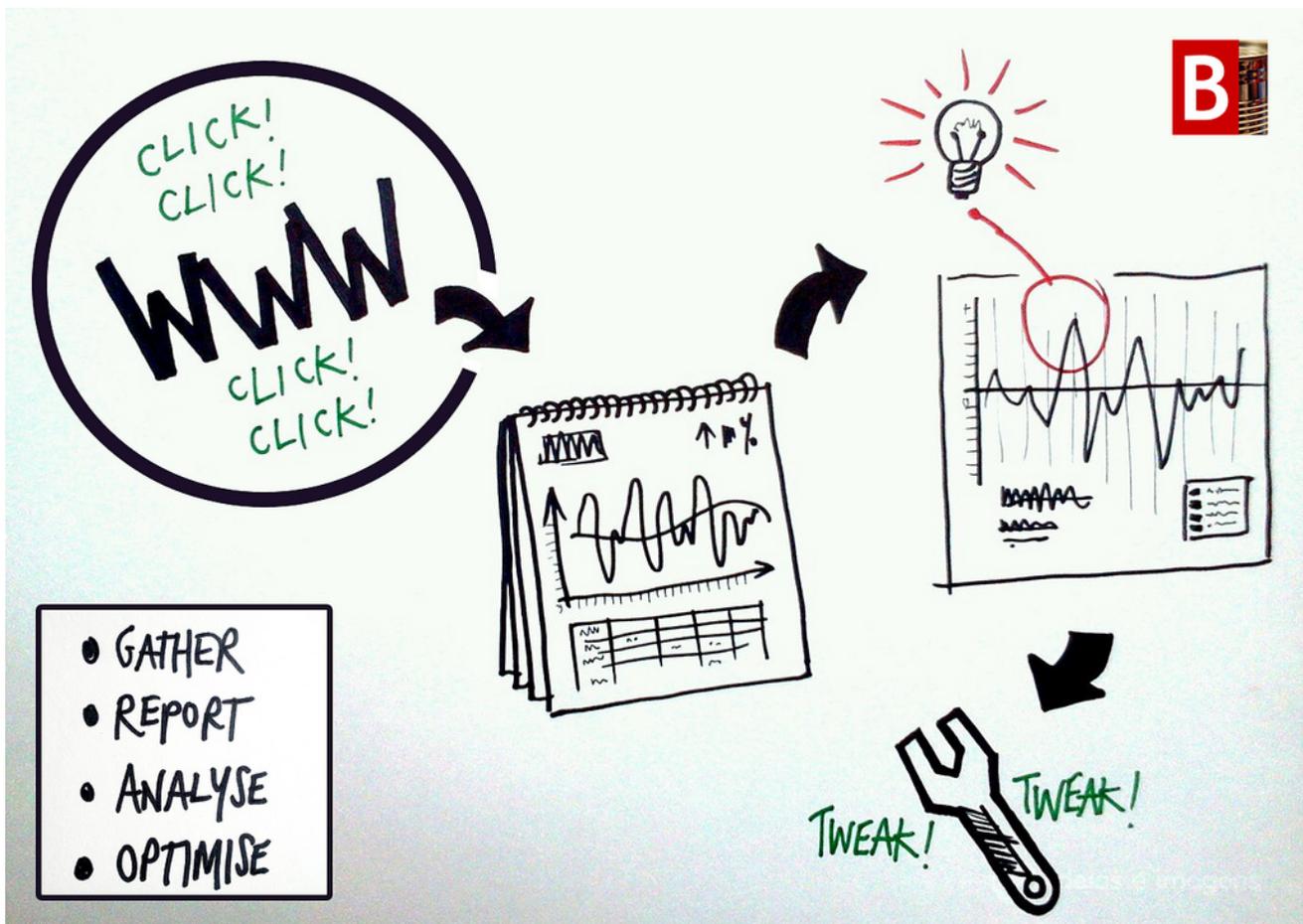


Image credit: [Web analytics framework](#) by James Royal-Lawson (CC BY-SA 2.0)

Understanding what is good science and who does it is an important variable to predict across a large range of contexts. More than 'metrics' simply being the way science works, their value and necessity is clearly evidenced by their explosion in usage and popularity. It is this importance, paired with the knowledge that humans will utilize any tool available to make their lives easier, which makes it so imperative to understand impact and what leads to it.

Blind use of bibliometrics is as inadvisable as blind use of any other data source. But the bottom line is that if we spent the time to read even one article from each application for every open faculty position to be filled, there would be little time for anything else.

Despite their importance, there is probably some merit to the notion that the metrics we currently use are easily hacked and flawed. It is important to also remember that the ones being measured are also human and have an interest in appearing as good as possible on these metrics; this is fundamentally the reason people clean their house before having someone over and the reason students study for (and cheat on) exams. There may be [fundamental problems with metrics](#), but I would suggest these are actually fundamental problems with the humans who use them, rather than the metrics themselves.

It is possible, even probably necessary, to build metrics that are better (unless we want the status to be the quo). Thus, I want to say something about:

2. Creating metrics that matter

Rather than expecting people to stop utilizing metrics altogether (which is unreasonable given the value they offer), we would be better off focusing on making sure the metrics are effective and accurate. They will never be problem

free, they will (probably) never be 'unhackable', but with a little coordinate effort they can get much better than they are now.

How to exactly achieve this goal was the topic of a recent [Quantifying and Analyzing Scholarly Communication on the Web](#) workshop at Websci15, where I gave a small presentation about [creating metrics that matter](#). This focus on creating more metrics echoes [several calls](#) on the LSE Impact blog, which suggest looking across metrics to better understand the facets of research impact, rather than striving for any single definition or measure of quality.

More than the need for metrics, the discussion was about how to make better use of the available data to construct the best metrics possible. Specifically, [Whalen et al](#) discussed the potential to better understand impact by looking at the 'topical distance between citing and cited papers'. [My response](#) focused more on the potential utility of social media discussions between scientists, and bringing in more theory from other fields (discussed briefly below).

Such a goal is easily achievable by integrating APIs into one hub by which to collect and analyze data about scientists (ORCID, Altmetric.com, and ContentMine are all moving toward somethings like this). This centralised data source would be helpful not only to bibliometrics, but could also be used for more 'fundamental' questions (e.g., what team or personal characteristics leads to optimal knowledge creation/ exchange?). As we learn more about the different facets of impact, it will probably become more difficult to cheat the system because it is difficult to change the multiple facets at once (and we can even analyse who cheats when).

More than simply creating many metrics, we need informative metrics that will help, rather than hurt, the scientific endeavour. This implies not only an ongoing effort to monitor the effects of metrics on the research enterprise, but also:

3. Including more science (theory!) in metrics

It is easy enough to build many metrics, but the best ones will probably utilize the accumulated scientific knowledge into their development. Scientists are humans (even if we pretend not to be), existing in cultures with histories and traditions. Many fields (e.g., Psychology, Philosophy of Science, Sociology, Marketing, Computer Science, Communication) can be utilized fruitfully to better understand scientific communication and impact.

Many of the metrics we utilize now are simple counts of things (e.g., citations, views, tweets, likes), but there is a wealth of other data available (e.g., keywords, review rating, comment sentiment). All of these data can be utilized to investigate hypotheses. For instance, in [my own work](#), we are examining how cognitive conflict in online discussion between scientists relates to productivity, generally speaking (e.g., length of discussions, view counts, citations).

More than just using science to make informative metrics, we should utilize it to ensure the changes we are implementing are actually helping. There are many initiatives right now to change the system, with [many of them lacking any data to indicate that they would help the situation](#). A little bit of empirical work demonstrating that proposed changes will help and are helping the system, along with an awareness of human tendencies to look to maximize their reward while minimizing their input, can go a long way toward avoiding later problems.

Ultimately, it is an empirical question how science operates most efficiently, and my main message is to say that we should [use science to improve science](#), not just to build better metrics, but to build a better functioning science and society, generally speaking.

Note: This article gives the views of the authors, and not the position of the Impact of Social Science blog, nor of the London School of Economics. Please review our [Comments Policy](#) if you have any concerns on posting a comment below.

About the Author

Brett Buttlere is a research assistant under Prof. Friedrich Hesse and Dr. Jürgen Buder at University of Tübingen

investigating knowledge exchange in the context of academia. Brett earned his Master's degree from Tilburg University under Dr. Jelte Wicherts and Dr. Marcel van Assen; the project focused on what psychologists feel are the most important problems to solve about the peer review system and how to best to solve them. Another project examined the motivations for and effects of 'extreme behavior' on individuals. His Bachelor's degree was received from Bradley University, where he worked on reducing prejudices toward Schizophrenia. Most generally, the goal of his research is to develop and apply fundamental psychological technologies to help individuals lead better lives (e.g., by trying to improve the efficiency of science, understanding and reducing extremism, elucidating the similarities and differences between humans and animals).

This is part of a series of pieces from the Quantifying and Analysing Scholarly Communication on the Web workshop. More from this series:

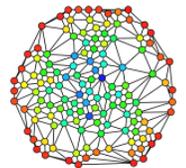
The ResearchGate Score: a good example of a bad metric

According to ResearchGate, the academic social networking site, their RG Score is "a new way to measure your scientific reputation". With such high aims, **Peter Kraker, Katy Jordan** and **Elisabeth Lex** take a closer look at the opaque metric. By reverse engineering the score, they find that a significant weight is linked to 'impact points' – a similar metric to the widely discredited journal impact factor. Transparency in metrics is the only way scholarly measures can be put into context and the only way biases – which are inherent in all socially created metrics – can be uncovered.



Bringing together bibliometrics research from different disciplines – what can we learn from each other?

Currently, there is little exchange between the different communities interested in the domain of bibliometrics. A recent conference aimed to bridge this gap. **Peter Kraker, Katrin Weller, Isabella Peters** and **Elisabeth Lex** report on the multitude of topics and viewpoints covered on the quantitative analysis of scientific research. A key theme was the strong need for more openness and transparency: transparency in research evaluation processes to avoid biases, transparency of algorithms that compute new scores and openness of useful technology.



Context is everything: Making the case for more nuanced citation impact measures.

Access to more and more publication and citation data offers the potential for more powerful impact measures than traditional bibliometrics. Accounting for more of the context in the relationship between the citing and cited publications could provide more subtle and nuanced impact measurement. **Ryan Whalen** looks at the different ways that scientific content are related, and how these relationships could be explored further to improve measures of scientific impact.



- Copyright © The Author (or The Authors) - Unless otherwise stated, this work is licensed under a Creative Commons Attribution Unported 3.0 License.