Metrics vs. Peer review: time to bury the hatchet?

Prof. Ele Belfiore
Centre for Research in Communication and Culture
Department of Social Sciences, Loughborough University
e.belfiore@lboro.ac.uk
@elebelfiore
A brief introduction

- An interest in the politics of cultural policy making through the analysis of the discursive formation around justifications of public funding for the arts and culture

- Years spent exploring the central role of the notion of the ‘socio-economic impact’ of the arts in policy rhetoric and the problem of impact assessment

- The rise of the impact agenda in HE

- The problem of seeing ‘impact’ as a proxy for ‘value’
On the importance of ‘judgements-between’

“There will never be enough money”.
“Choices will always have to be made, judgments-between”

(Richard Hoggart)
The context for the HEFCE Review and the Metrics Tide report

- Could the REF be cheaper to run and less burdensome for institutions with a more heavy reliance on metrics, yet without compromising the integrity of the process?
- And what about metrics for impact?
So, what about metrics as a way to assess research quality?

Tempting, as definitely cheaper and easy. However... The conclusions we reached in *The Metric Tide* still stand:

- **Finding:** Peer review, despite its flaws and limitations, continues to command widespread support across disciplines. Metrics should support, not supplant, expert judgement.

- **Recommendation:** The research community should develop a more sophisticated and nuanced approach to the contribution and limitations of quantitative indicators.
The problems with metrics as proxy for quality

- Limitations in coverage of bibliographic databases
- Self-citation (and different attitudes in different disciplines)
- Difficulties with interdisciplinary and collaborative (multi-authored) research
- The JIF question (especially when Journal-level indicators are used as a substitute for publication-level citation statistics)
- The challenge of applying bibliometrics methods to books
- Can’t deal with ‘unusual’ research outputs, although these are eligible for submission to the REF
- Can’t deal effectively with publications in languages other than English
- Variation in citations practices across disciplines
- Gender and diversity issues
Citation is problematic per se, but it is especially so for some

Focus on gender and diversity issues:

Sugimoto et al. (2013): a bibliometric analysis of the relationship between gender and research output (5,483,841 research papers and articles; 27,329,915 authorships).

Their conclusion: “in the most productive countries, all articles with women in dominant author positions receive fewer citations than those with men in the same positions. And this citation disadvantage is accentuated by the fact that women’s publication portfolios are more domestic than their male colleagues – they profit less from the extra citations that international collaborations accrue. Given that citations now play a central part in the evaluation of researchers, this situation can only worsen gender disparities.”
The politics of citation practices

- **Implicit bias** – it may not a plot to hold academic women and minorities back, but it is not less damaging for that

  ‘Although explicit gender bias still exists, implicit (unconscious) gender bias is particularly troubling because it can lead well-intentioned people to unwittingly perpetuate inequalities’.

  (Kaatz, Gutierrez & Carnes 2014:371)

- If this bias makes some scholars’ work invisible, what are we really assessing?
The Journal of Political Philosophy and the invisible (black) scholars

May 2017:
A section of the journal was devoted to papers discussing the Black Lives Matter movement
None of the authors were black.

Nature’s editors in 2012: ‘We believe that in commissioning articles or in thinking about who is doing interesting or relevant work […] men most readily come to editorial minds.’
The ‘real’ problem with metrics is not a ‘technical’ one...

Citation is a social practice, and as such reflect social norms and hierarchies

- In the economy of reputation that rules the REF, prestige, citations, influence are precious currency. Therefore the disadvantage faced by women and ethnic minorities is real: it affects career prospects, employability, access to research funding, etc.

- This is an issue not limited to research quality assessment. Boring, Ottoboni & Stark (2016) study of SET shows that:
  - SET are biased against female instructors by an amount that is large and statistically significant.
  - It is not possible to adjust for the bias, because it depends on so many factors.
  - Gender biases can be large enough to cause more effective instructors to get lower SET than less effective instructors.
(Unsurprisingly), unconscious bias also affects judgments about the quality of research.

Knobloch-Westerwick, S., Glynn, C. J., & Huge, M. (2013): “Participants rated conference abstracts ostensibly authored by females or males, with author associations rotated. The abstracts fell into research areas perceived as gender-typed or gender-neutral to ascertain impacts from gender typing of topics. Publications from male authors were associated with greater scientific quality, in particular if the topic was male-typed. Collaboration interest was highest for male authors working on male-typed topics. Respondent sex did not influence these patterns.”
And yet, peer review is also flawed...

- Double blind peer review increases female authors’ chances of acceptance.
- Making an effort to have more women peer review is a desirable move, but won’t automatically solve the problem: plenty of evidence shows that those groups directly affected by bias, also exercise it themselves.
- Bornmann, Mutz & Daniel (2007) conclude “in all, among grant applicants men have statistically significant greater odds of receiving grants than women by about 7%.”
- But, whilst this might work for peer review of publications, it does not work for REF nor for peer review of funding applications.
What then? Going beyond the metrics vs. peer review stand-off

- The conclusion of *The Metric Tide* still stands: metrics as support of peer judgment of quality, not a substitute

- What is the **real** purpose of research quality assessment?

- Is fetishizing either research quality assessment approach the best way to ensure we create an environment conducive to the production of excellent scholarship? (NO!)