

Does high-impact research come at the expense of quality? An automated analysis of the REF impact landscape

 blogs.lse.ac.uk/impactofsocialsciences/2017/03/02/does-high-impact-research-come-at-the-expense-of-quality-an-automated-analysis/

3/2/2017

*As part of the 2014 Research Excellence Framework, the impact of research was assessed for the first time. But how effective was the impact category in capturing the many diverse forms of interaction between academia and society? Were certain interpretations of impact more highly rewarded than others? And does high-impact research come at the expense of quality? **Martin Zaltz Austwick** and his research team used a text mining technique to analyse impact case study submissions and address some of these questions.*



I'm going to assume that readers of this blog are familiar with the essential aims and scope of REF2014 – to evaluate research activity generated by higher education institutions (HEIs) across the UK. Our [recent work](#) has focused on impact, and so here I'll talk exclusively of impact (effects outside academia) and quality (peer-reviewed evaluations of how good the research is). Environment, the third category, is not discussed here.

While quality was assessed directly based on academic publications, impact was judged based on a series of case studies submitted by HEIs; written cases that described the effects of particular research outside academia. HEFCE (the Higher Education Funding Council for England) has shared all impact case studies and also been kind enough to create [an API](#) for examining them, meaning we were able to access and use this corpus for text analysis (if you're so inclined, I've shared my [Python Code](#)). Our questions were around how impact was framed and reported, and how this was rewarded. The interaction of academia and society can take many different forms, such as entrepreneurship, activism, community engagement, policy, etc. We wanted to understand how the impact category captured this – did it reflect academic engagement in all its glory? Or narrow it towards specific forms? Perhaps just as importantly, does high-impact research come at the expense of quality?



Many routes to impact. Image credit: [Coloured Doors](#) via [public domain](#).

The nitty gritty

One of the unstated aims of the REF seems to have been to make the reporting framework and terminology as complex as possible, so I'll spend a little time unpacking that. Within the REF, there were four panels: Life Sciences (A), Physical Sciences (B), Social Sciences (C) and Arts and Humanities (D). Within these four panels were 36 Units of Assessment (UoAs), further subdividing the broad panels into more specific disciplinary areas (for example, UoA7, under Panel B, is "Earth Sciences and Environmental Systems", a subspeciality of Physical Sciences).

Institutions submitted academic outputs and impact case studies at UoA level. All selected researchers included their papers in their institution's submission for that UoA; meanwhile, the institution also decided on a number of impact case studies to be submitted too. For the sake of brevity, I'm going to describe submissions (publications, impact case studies, and anything else) by a specific *institution* to a specific *UoA* as a 'unit submission'.

Does impact come at the expense of quality?

One of the questions we considered was whether research impact and research quality are linked; does good research have high impact, or does impact arise from superficial but 'flashy' research? There are a couple of caveats to this: firstly, the impact cases submitted had to be linked to research outputs (papers), but did not have to be linked to the *same* outputs as were submitted to demonstrate research quality. So any analysis we do can't link the quality of research *papers* to their impact, only link the quality of a *unit submission* to its impact. We might frame this question: does a university's research in a particular area show a link between quality and impact?

The second caveat is that impact case studies had to be based on outputs that were judged to be of at least two-star quality, so that should exclude really poor research (such as the superficial, 'flashy' type described above). But as researchers did not know what rating their work would be given prior to submitting their impact case, this is no guarantee.

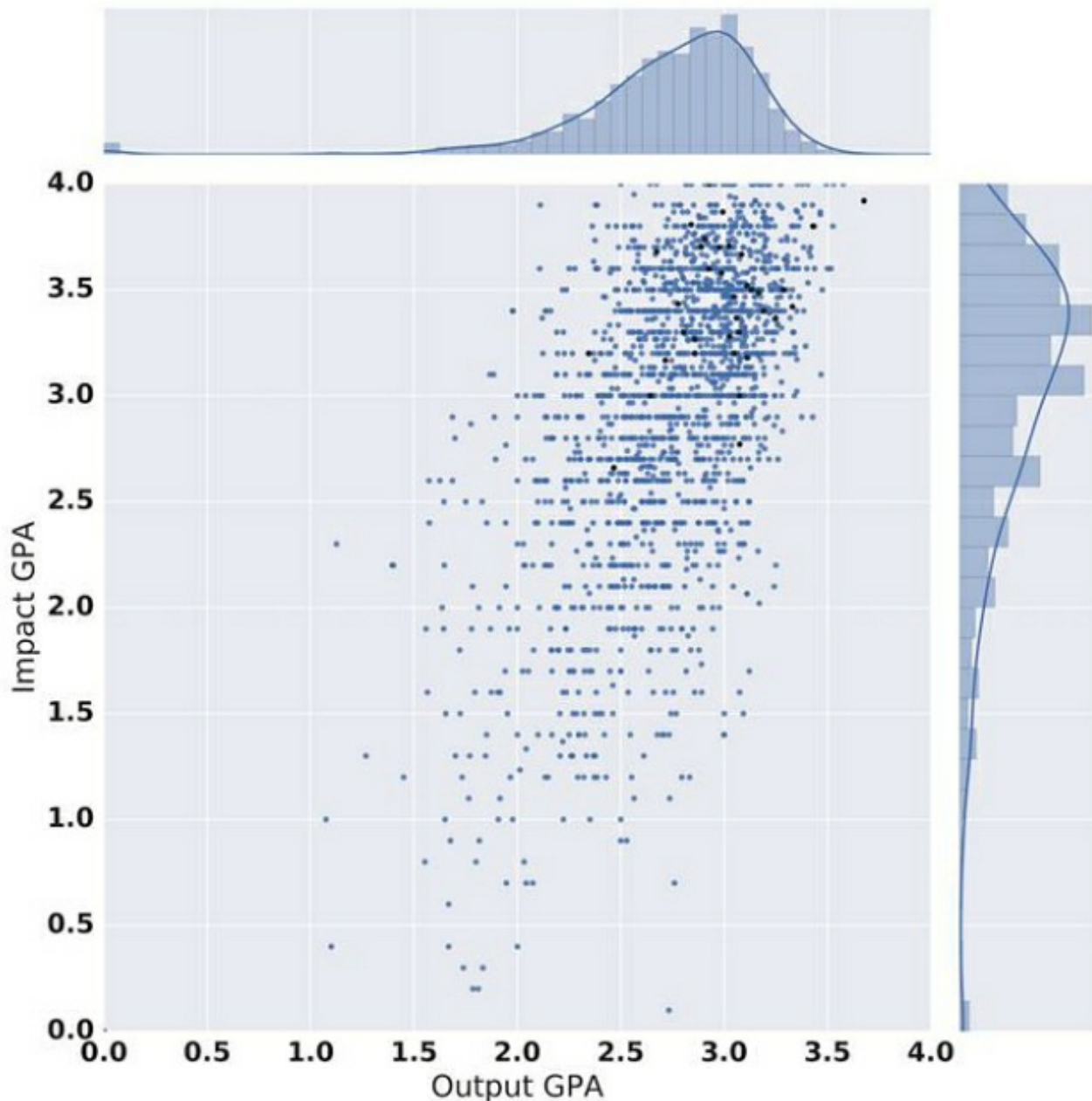


Figure 1: REF Impact vs REF Output (Quality). UCL submissions are highlighted in black. This figure was originally published in the article [‘Beyond Academia – Interrogating Research Impact in the Research Excellence Framework’](#) and is published under a [CC BY 4.0](#) license.

We used the Grade Point Average (GPA) for a unit submission – the simple mean of the scores across all submissions in that unit – and compared the Impact GPA to the Quality GPA. If we plot Quality vs Impact (Figure 1), we see what looks like a (very) fuzzy straight line – there is definitely a correlation between the two, but it is not the only determining factor. There are plenty of exceptions, but there’s a general trend towards high-quality work having high impact – the highest impact case studies (GPA = 4.0) scored GPAs of 2.5-3.6 in quality, and the highest quality unit submission (GPA = 3.7) scored very nearly the highest impact (GPA = 3.9).

What do institutions talk about when they talk about impact?

A [previous study by Kings College London](#) carried out topic modelling on the same text dataset, generating 60 “impact topics” and 3,709 “routes to impact”. But this didn’t quite get at the question we wanted to answer. We

wanted to understand what the impact cases had in common, and arrange them into a small number of classes of impact activity which were distinct and mutually exclusive – an approach amenable to cluster analysis. Without going into the technical detail, we used the proximity of words within sentence-like blocks of text to create six distinct clusters, and, through independent interpretation and consensus within the team, created labels to describe those clusters:

- Education (covering both pedagogical practice and education strategy)
- Public engagement (non-governmental and non-commercial beneficiaries through media, museums, public talks and workshops)
- Environment and energy solutions (both policy and technology-focussed)
- Enterprise (knowledge transfer and spin-out companies)
- Policy (impacts on local and central government and planning)
- Clinical uses (experimental medicine and clinical trials).

These activities should be familiar to academics involved in outreach, and many are quite closely linked to general research areas (some, such as public engagement, are much broader).

One of our questions was whether any of these language classes perform better (or more poorly) in the impact category. We sought to answer this by identifying every unit submission which belonged to a distinct class and examining its performance. Unfortunately, this was not possible for the whole-corpus analysis as many unit submissions were linked to more than one language class – we suspect this is because each HEI frames impact slightly differently, and so units didn't quite fit into these sector classes precisely.

We instead carried out an analysis on our own HEI, UCL, using just our submissions to create a new set of classes, similar but not identical to the whole-corpus analysis: education, public engagement, enterprise, policy, and clinical uses. Only three unit submissions showed links to more than one language class, so we were able to map language class onto performance. Class 1 (Policy) included some higher performing submissions, and Class 5 (Enterprise) showed a wider spread, but there was no consistent bias – both showed a wide overlap with other classes.

What next?

As commented, REF did not require that papers submitted for quality assessment be the same as those papers submitted in connection with impact cases – so an intrinsic link with quality could only be established (or refuted) at a unit submission level. In theory, a unit submission might contain talented but insular researchers as well as mediocre but outward-looking ones. While I think there's an argument that this allows for diverse academic careers and skillsets, by divorcing the outputs submitted for quality and those submitted for impact, there is a danger of creating a two-stream research landscape and career path; one which could, in turn, lead some to question the credibility of high-impact authors, perceived to be in some way lacking research quality. This presents an argument for submitting the same set of publications for both impact and quality assessment (although currently the reporting timescales for each are radically different, further muddying the waters).

We believe this presents an argument for HEIs to keep impact diverse in form and aim, and for HEFCE to continue to reward impact beyond what are perhaps the obvious and quantifiable pathways: enterprise, knowledge exchange, policy impact and clinical translation. The many academics working in community engagement, media, museums and events, crowdsourcing and citizen science, and the many other diverse forms in which academia interfaces with the wider world, should continue to be recognised for their work, and we'd be keen to see HEFCE, and its host institutions, continue to do so.

This blog post is based on the author's co-written article, ['Beyond Academia – Interrogating Research Impact in the Research Excellence Framework'](https://doi.org/10.1371/journal.pone.0168533), published in PLoS ONE (DOI: 10.1371/journal.pone.0168533).

Note: This article gives the views of the author, and not the position of the LSE Impact 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

Martin Zaltz Austwick is a Senior Lecturer in Spatial Data Science and Visualisation at the Bartlett Centre for Advanced Spatial Analysis (CASA) at UCL. His research interests include visualisation and analysis of cycling systems, freight transportation, sharing economies, communities of practice, and pedestrian behaviour; and Digital Humanities, including corpus text analysis and collaborative mapping. He has an active interest in public engagement through public talks, workshops, datavis, audio, blogging, and community engagement, and won the 2015 UCL Public Engagement Award for Institutional Leadership.

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