

What impact evidence was used in REF 2014? Disciplinary differences in how researchers demonstrate and assess impact

A new report produced by the Digital Science team explores the types of evidence used to demonstrate impact in REF2014 and pulls together guidance from leading professionals on good practice. Here [Tamar Loach](#) and [Martin Szomszor](#) present a broad look at the types of evidence in use in the REF impact case studies and reflect on the association between use of evidence in the various categories and the scores received by submissions in peer review.

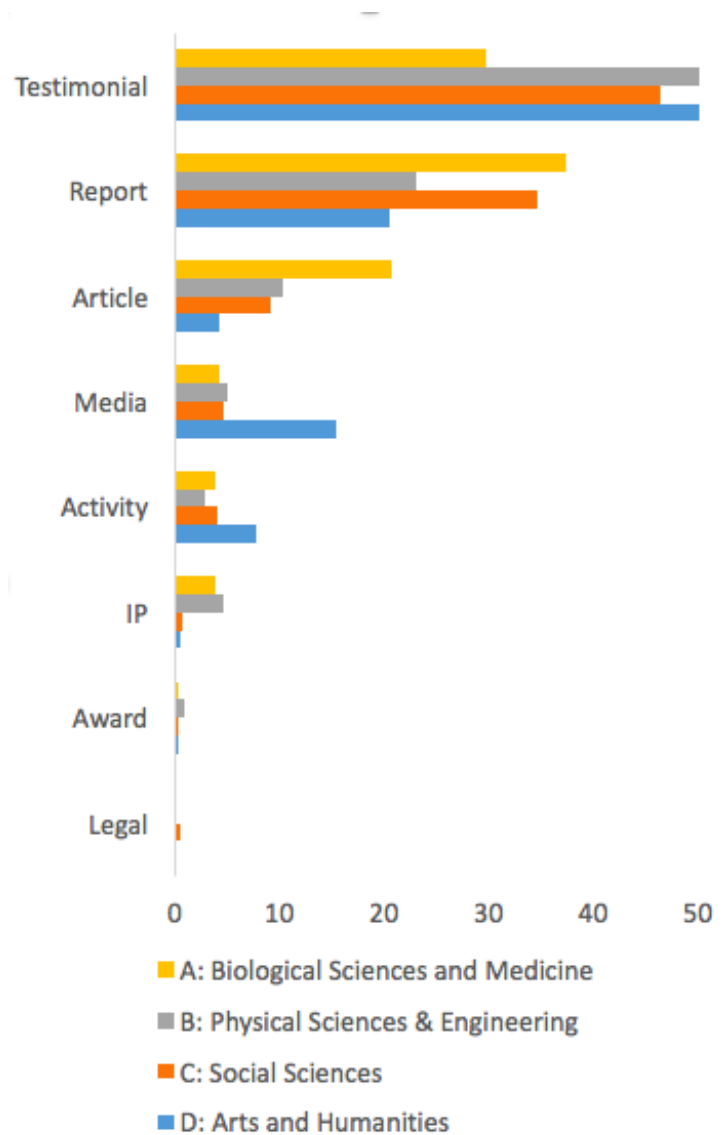


In our latest [Digital Research Report on impact evidence](#) we showed that corroborating documents included in impact case studies and submitted to REF 2014 varied in type by subject panel. The chart reproduced below shows how media items (like television programmes, youtube videos, news articles...) were used more in panel D (Arts and Humanities) than in any other panel, whilst reports appeared in panel A (Biological Sciences and Medicine) the most.

This kind of analysis is broad-brush and not particularly surprising – we have counted clinical guidelines as reports and media pieces are surely closer to the research carried out in panel D than in any other panel. What it does do is show that by allowing for flexibility in the way impact is reported (case studies could cite any evidence that corroborated impact so long as it was auditable) diversity in approach and content is encouraged and realised. Here we show this at the level of main panels, but the variety and multiplicity of impact stories in the [full set of case studies](#) is backed up by a huge array of impact evidence – evidence includes [technical documentation](#) on commercial websites, [social media and audience responses](#) to public activities and a huge number of URLs pointing to unknown resources.

Figure 1: The percentage of REF 2014 Impact Case Studies, by main panel, that contain categorised types of evidence.

Source: Loach, Tamar; Adams, Jonathan; Szomszor, Martin (2016): Digital Research Report: The Societal and Economic Impacts of Academic Research – International perspectives on good practice and managing evidence. Digital Science.



figshare. <https://dx.doi.org/10.6084/m9.figshare.3117928.v2>

Not all of these evidence types are represented in the aggregate numbers in the bar chart – we classified strings of text representing pieces of evidence using a collection of keywords and patterns to group evidence pieces into commonly occurring categories, but this inevitably misses out the less common evidence types. This obscures finer grained diversity, but with the evidence types we *are* able to categorise, we can look at the association between use of evidence in the various categories and the scores received by submissions in peer review. We are also limited by the public availability of scores, which are [available](#) at the Unit of Assessment/Institution level only. This means that we know the percentage of 4*,3*,2*,1* case studies that a university submitted to a particular panel, but we can't say what any particular case study scored.

At the high level we are able to access the association between scores and evidence types can be tested. The correlations between the amount of evidence of a given type, and the Grade Point Average (GPA) score for a set of case studies is shown in the coloured table below (again reproduced from our recent report). This looks at average behaviour and obscures outliers (which may themselves be more interesting than group trends) but we can see differences across panels, and use these differences to point to the benefit of flexibility in assessment systems. We saw that arts and humanities case studies were those most likely to include media as evidence of impact. But within Arts and Humanities, use of media does not affect the GPA associated with a case study set in a significant way. In contrast, the reports used in nearly 40% of Medical and Health case studies have a statistically significant positive

correlation with GPA. Although not a causal link, it does seem that use of reports as impact evidence is associated with highly scoring impact case studies across the panels. On the other hand testimonials seem to be associated with higher scores in Arts and Humanities but are negatively related to score in the Medical Sciences.

Figure 2: Spearman correlation between the indicative score and the amount of various types of evidence; there is a column for each subject panel.

	A: Biological Sciences & Medicine	B: Physical Sciences & Engineering	C: Social Sciences	D: Arts & Humanities
Activity	-0.03	-0.02	-0.04	-0.06
Article	0.19	0.09	0.02	-0.01
Award	-0.06	0.01	0.01	0
IP	0.05	0.05	-0.01	0
Legal	-0.03	0	0	0
Media	-0.01	0.07	-0.07	0
Report	0.19	0.11	0.15	0.08
Testimonial	-0.15	0.04	0.08	0.17

Source: Loach, Tamar; Adams, Jonathan; Szomszor, Martin (2016): Digital Research Report: The Societal and Economic Impacts of Academic Research – International perspectives on good practice and managing evidence. Digital Science. figshare. <https://dx.doi.org/10.6084/m9.figshare.3117928.v2>

A value of 1 implies maximal positive correlation, 0 no correlation, and -1 a maximal inverse correlation. The values in bold are significant (p value < 0.05, where the null hypothesis is that the indicative score and the amount of a given evidence type are uncorrelated.)

In our report we conclude that the case study format allowed a variety of evidence to be displayed, and that these types appeared – and were assessed – differently across research communities. The evidence ontology we use neglects less common forms of evidence, and the limited availability of scores means we’re really looking at averages and associating them with individual case studies. Nevertheless, we see different practices across broad subject categories. We also know that there is great variety in the impact case study database, and that there are multifarious differences across individual case studies. We use this variety to argue that the surprises and outliers in the impact case studies provide the most value and interest, and to encourage diversity in content, certainly not ruled by a formulaic impression of what good impact or good impact evidence looks like.

Featured image credit: Magnifying Glass [Loupe], 1963, Roy Lichtenstein[Gautier Poupeau \(CC BY\)](#) and [maps.unsplash.com](#) (Public Domain)

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 Authors

Tamar Loach is Research Metrics Analyst at Digital Science. Her background is in theoretical physics, and she has undertaken research projects both in this area and on alternative research metrics with the Complexity and Networks group, all at Imperial College London. She currently works with all kinds of research data, and helped develop the impact case study website and a synthetic analysis of the impact case studies for HEFCE after REF 2014.

Martin Szomszor is Head of Data Science at Digital Science. He was previously Deputy Head of Centre at the City eHealth Research Centre (2009-2011) where he led research on the use of social media in epidemic intelligence. He was also Chair of the 4th International Conference on Electronic Healthcare for the 21st Century in 2011. Martin was formerly a Research Fellow at the University of Southampton (2006-2009) where he worked on various Linked Data, Semantic Web, and Social Network analyses projects.

- Copyright 2015 LSE Impact of Social Sciences - Unless otherwise stated, this work is licensed under a Creative Commons Attribution Unported 3.0 License.