

Issues in Research Evaluation in the Light of Different Disciplines

Or

For Complexity

Dr Claire Donovan, Health Economics Research Group, Brunel University
Address to 3rd ESF MO Forum 'Evaluation of Publicly Funded Research'
Paris, 10 May 2011

Introduction

When we think about research evaluation, it falls into two main areas: 'quality' and 'impact', and today I want to focus on 'impact'.

I want to focus on 'impact' as this is a 'hot' topic internationally. And also because thinking in terms of disciplines can be quite problematic here, as 'impact' is often an interdisciplinary enterprise: evaluation is often conducted in disciplinary silos, or when efforts are made to have interdisciplinary, academics tend to lapse back into disciplinary-specific judgements. While more 'impact' is desired from research funding agencies, there are structural reasons why this is difficult – and perhaps the most significant barrier is that universities tend to remain to be organised on disciplinary grounds, as do research careers.

Demonstrating research impact is, on the face of it, easier for some academic disciplines than for others. The benefits of health and medical research, for example, can be quite obvious, and lend themselves to being demonstrated using simple metrics or quantification, but I want to encourage you to think about impact in terms of broader public value also.

Today I would like to talk about research 'impact' at the national assessment level in the light of the Australian experience and the forthcoming Research Excellence Framework (REF) in the UK, and also at the grant evaluation level in the light of the increasing importance of 'impact' in FP8, and for national research funding agencies (the UK again being a prime example). So I would like to talk about the importance of assessing the public value or the 'wider benefits' of university research. And the title of my presentation is 'Issues in Research Evaluation in the Light of Different Disciplines', but it also could go by another title, which is 'For Complexity' as I want to persuade you to want 'complexity' in research evaluation. When I say 'complexity', I am being provocative: what I mean is that, in my opinion, *if* it is

to be assessed, ‘impact’ needs to be assessed using a *sophisticated*, robust, contextual approach to assessing research impact (and research quality).

I will explain:

- what research ‘impact’ is and why it is important to assess it;
- how ‘impact’ may be assessed;
- what governments and funding agencies want from research assessment;
- what governments and funding agencies *should* want from research assessment;
- and will argue for complexity.

The point is that governments (or public research funding agencies) want to demonstrate that the research they have funded has made a difference. But it seems that all too often evaluators are fated not to get what they want because of the way they choose to get it. That is why I shall argue for complexity.

What is research ‘impact’?

To explain what research ‘impact’ is, we first need to make a distinction between what evaluators call research ‘quality’ and research ‘impact’.

Put simply, research ‘quality’ refers to academic publications or creative outputs, and citations between academic publications. Research ‘impact’ is the extra-academic influence of research: that is, the benefits that arise from academic research beyond the academic world. I guess a few of us might have encountered the ‘MIND THE GAP’ sign on the London Underground. Paul Nightingale and Alastair Scott (SPRU) talk of the ‘relevance gap’ between the research society most needs and that which is produced. They believe that a social contract exists between publicly funded researchers and the taxpayers who fund their activities, and so publicly funded research should address pressing social issues.

Governments are increasingly seeking to link research funding to the ‘impact’ of research, but are struggling with how best to do this. And I think that Nightingale and Scott are making a valid point, and this idea of a ‘relevance gap’ resonates (perhaps unfairly) within government, particularly with regard to the humanities, creative arts, and social sciences: not so much so with health and medical research and the ‘hard’ sciences.

Yet I think that this sets the debate about research ‘impact’ on the wrong foot because: (1) there is plenty of research that has ‘impact’ beyond academia – the problem is that this is rendered invisible by standard approaches to assessing research impact; and (2) it is essential to retain ‘blue skies’ research. Both (1) and (2) are of public value or benefit in themselves. There is also a tendency to separate ‘quality’ and ‘impact’ so that there is a misperception that somehow high impact equals low quality: but this misunderstands and undervalues methods by which academics target research to practitioners or ‘end-users’ (i.e. ‘grey literature’ or practitioner-oriented journals).

In Australia, I was Chair of a government committee tasked to find the most robust methodology for assessing research impact for Australia’s Research Quality Framework (RQF). As part of the RQF process, my committee sought to find what was best international practice in assessing research ‘impact’. My committee chose to argue that the Australian government needed to radically rethink its idea of what research ‘impact’ was.

In most industrialized countries, the innovation agenda has come to dominate science and technology policy, which has come to dominate and research policy, which has come to dominate Higher Education policy. This predicated thinking of ‘impact’ in economic terms: economic returns, and returns to business and industry, but not in terms of its broader public value or benefit. At heart, in terms of public value or benefit, there is a conflict between the agenda of innovation and the values that underpin Higher Education teaching and research. The ‘innovation agenda’, for example, has been criticized for marginalizing the role of the humanities, creative arts, and social sciences (HASS). I argue that the innovation agenda also undervalues all research fields in science, technology, engineering and medicine (STEM) also. The innovation agenda alienates STEM and HASS academics, and perpetuates a narrow view of what ‘impact’ is and what ‘impact’ can be.

Meaningful assessments of research impact must strive to measure the wider public value or benefit of research. This idea is embraced by post-New Public Management concept of ‘triple bottom line’ accounting which has been used in Australia to account for the public value of public services and of research not purely in economic terms, but by using a balanced approach to social, environmental, and economic gains. This alternative approach redefines the purpose of innovation policy, or science and technology policy. Yet, finding meaningful

public value eludes standard quantitative approaches to research assessment, and we find that the application of simple metrics tells us little, as there is a tendency to reduce data to the single bottom line of profit.

How ‘impact’ may be assessed

How we choose to measure ‘impact’ will determine the kind of impact we find. Historically, simple metrics fit with the aspirations of the innovation agenda, and contextual approaches lend themselves to public value agendas. Simple ‘impact’ metrics do not measure research impact. For example ‘technometrics’ such as number of patents and number of spin-off companies focus on economic returns, and only really gauge very low levels of impact or activity rather than wider public benefit, and also privilege private over public interests. And ‘sociometrics’ such as attempting to link macro-level social statistics (e.g. the crime rate or the divorce rate) to the efforts of particular researchers is just not credible (this was once described as ‘making alchemy look good’).

Simple metrics therefore ignore wider public value or benefit. More complex, holistic and contextual methods of assessment rely on peer (and ‘end-user’) judgements, and are necessarily complex but more meaningful.

There is a danger that over-attachment to simple metrics limits the imagination of what governments and evaluators want and how they seek to find it. This limits our imagination and our aspirations. Metrics-led thinking is like the analogy of searching for a dropped set of car keys on a dark night only in the circle of light provided by a nearby lamppost. To borrow another analogy, to allow simple metrics to shape research evaluation exercises because they exist and are cheap is like the tail wagging the dog: rather than seeking what is out there, simple metrics limit what can be found. And this applies not only to ‘impact’ metrics but ‘quality’ metrics also.

What do governments and research funders want?

Not unreasonably, in terms of research evaluation, governments want to link funding to high quality original research and to beneficial outcomes. Governments also want research evaluation exercises to be cost-efficient, simple, to be transparent and to have replicable evaluation methodologies, and to use remote technologies (i.e. databases that are independent

from researchers and so cannot be manipulated, and also do not involve taking up a lot of researchers' time in conducting assessments).

Yet, as we have seen, more simplistic metrics-based approaches do not deliver a sophisticated appreciation of research impact. Rather, this often yields disappointing results that do not credibly link the research produced to its outcomes.

What should governments and research funders want?

Governments should want to link funding to original high quality research to beneficial outcomes. Governments should seek complexity and the long way round rather than short-cuts. This complexity involves seeking a variety of types of public value, and mixed quantitative and qualitative assessment methods which are transparent, robust and replicable. Governments should also want to invest the necessary time and effort, and to use 'embedded' knowledge, interpretation and expertise (for example, the views of 'end-users').

The message today coming from best practice in impact evaluation is that the time of simplistic metrics has passed.

What do governments want? Governments want to demonstrate how research funding has made a difference. But it seems that governments will be fated not to get what they want because of the way they have so far chosen to get it.

What may change governments' minds? An example of a successful, large-scale contextually-driven, evaluation of research impact. Australia came tantalizingly close to this in the form of the Research Quality Framework, and HEFCE have piloted precisely that model and methodology of assessing research impact for the 2014 Research Excellence Framework. But if this is rejected by the academic community, will that then leave us with simple metrics instead?

For complexity

I will now conclude by arguing for complexity. Take the RQF approach which defined 'impact' as the social, economic, environmental and cultural benefits from research. It was found that metrics simply did not capture public value, and so a complex approach was adopted involving a minimal (but important) role for metrics, the evaluation of context

statements and case studies, and the use of ‘end-user’ testimony, if required. And this approach has been partially adopted by the UK REF.

This sought to reveal what was called ‘social benefit’:

“Improving quality of life; stimulating new approaches to social issues; changes in community attitudes, and influence upon developments or questions in society at large; informed public debate and improved policy-making; enhancing the knowledge and understanding of the nation; improved equity; and improvements in health, safety and security.”

‘economic benefit’:

“Improved productivity; adding to economic growth and wealth creation; enhancing the skills base; increased employment; reduced costs; increased innovation capability and global competitiveness; improvements in service delivery; and unquantified economic returns resulting from social and public policy adjustments.”

‘environmental benefit’:

“Improvements in environment and lifestyle; reduced waste and pollution; improved management of natural resources; reduced consumption of fossil fuels; uptake of recycling techniques; reduced environmental risk; preservation initiatives; conservation of biodiversity; enhancement of ecosystem services; improved plant and animal varieties; and adaptation to climate change.”

and ‘cultural benefit’:

“Supporting greater understanding of where we have come from, and who and what we are as a nation and society; understanding how we relate to other societies and cultures; stimulating creativity within the community; contributing to cultural preservation and enrichment; and bringing new ideas and new modes of experience to the nation.”

This combines to show the richness that contextual impact approaches bring to the understanding of research impact.

However, a change of government in Australia meant that ‘impact’ was abandoned because it was perceived to be too complex, and the new Excellence for Research in Australia (ERA) uses simple metrics in the form of:

- patents
- plant breeders’ rights
- registered designs
- registered commercialization income

That is the approach in total. What benefits may be captured here?

That is why, for ‘impact’ to work in the context of evaluating EU research, I argue for complexity which is sophisticated (uses contextual, qualitative and quantitative approaches), embraces meaningful public value (social, economic, environmental and cultural); values the humanities, creative arts and social sciences in their own terms; and appreciates the broader benefits of STEM and HASS for society.

To not think in complex terms is to limit government – and our own – imaginations about what research ‘impact’ actually is and can be. This is how we may gauge what, through their research function, universities give back to society.

The problem is selling complexity to government. As we found with the RQF and ERA, what governments think is best may not be what universities or researchers want, or may not coincide with best practice in the research evaluation community. The RQF approach to assessing research ‘impact’ is an example of the richness a sophisticated approach can bring when compared with simple metrics alone. Simple metrics cannot display public value.

Conclusion

My conclusions are quite simple:

- metrics widen the gap between academic research and research impact;

- public value closes the gap between academic research and research impact:
- this view of impact should be mainstream and (if you will pardon the pun) not just underground thinking.*

**For an account of the rise and demise of 'impact' in the RQF see: Donovan, C. (2008) 'The Australian Research Quality Framework: A live experiment in capturing the social, economic, environmental, and cultural returns of publicly funded research' New Directions for Evaluation, 118, 47–60.*