



Governing by narratives: REF impact case studies and restrictive storytelling in performance measurement

Justyna Bandola-Gill ^a and Katherine E. Smith ^b

^aSchool of Social and Political Science, University of Edinburgh, Edinburgh, UK; ^bSchool of Social Work and Social Policy, University of Strathclyde, Glasgow, UK

ABSTRACT


Performance assessment is permeating increasingly diverse domains of higher education, even in areas previously perceived to be too complex and idiosyncratic to quantify. The UK's attempts to assess 'research impact' within the Research Excellence Framework (REF) are illustrative of this trend and are being closely monitored by several other countries. A fundamental rationale for employing narrative case studies to assess impact within REF, rather than taking a (less resource intensive) quantified approach, was that this would allow for the variation, complexity and idiosyncrasy inherent in research impact. This paper considers whether this promise of narrative flexibility has been realised, by analysing a combination of REF impact case study reports and interviews and focus group discussions with actors involved in case study production. Informed by this analysis, our central argument is that the very quality which allows narratives to govern is their ability to standardise performance (albeit whilst retaining a degree of flexibility). The paper proposes that REF impact case studies position narratives of impact as technologies of governance in ways that restrict the 'plot line' and belie the far more complex accounts held by those working to achieve research impact. This is partly because, as research impact becomes institutionalised within universities' measurement infrastructures, higher education institutions become impact gatekeepers, filtering out narratives that are deemed overly complex or insufficiently persuasive, while perpetuating particular approaches to recounting tales of impact that are deemed likely to perform well. Crucially, these narratives not only describe impact but actively construct it as an auditable phenomenon.

KEYWORDS

Research impact; academic engagement; REF(the UK Research Excellence Framework); research evaluation; impact agenda

1. Introduction

As performance measurement had expanded into new areas of governance, we have witnessed a rise of quantification as a key technology of governance; a trend Hoggett (1996, 22) calls 'measurement-fever'. The dominant logic within public policy and administration now often appears to be that the central way for states to govern problems is to measure them, with higher education leading the way (Shore and Wright 1999). This raises the question: are all higher education domains quantifiable? This question is of central importance to university governance in a paradigm that increasingly appears to imply, 'if you cannot measure it, you cannot manage it' (Feller, 2009,

CONTACT Justyna Bandola-Gill  Justyna.bandola-gill@ed.ac.uk  School of Social and Political Science, University of Edinburgh, 22 George Square, Edinburgh EH8 9LD, UK

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

323). We explore this problem by focusing on a field that previously has not been subject to audit – research impact and its assessment through case studies within the UK Research Excellence Framework (REF). This design, intended to enable the ‘commensuration of incommensurables’ (Espeland 1998), providing a first step to incentivising and managing impact within universities.

The genesis of the REF as an evaluative framework dates back to the 1980s, when the UK government was fundamentally questioning the value of academic research and proposing significant budgetary cuts (Bulmer 1987). Christopher Ball (then Warden of Keble College, Oxford University 2014b) recalled the following dinner conversation with Peter Swinnerton-Dyer (then Chair of the University Grants Committee) and David Phillips (then Chair of the Advisory Board for the Research Councils):

[W]e used to have dinner together and plan our strategy. One evening Peter said: ‘I can no longer defend the funding of universities [...] without real accountability to government’ [...] So we discussed it and I suppose at that dinner we invented the research selectivity exercise. (Attributed to Christopher Ball in Kogan and Hanney 2000, 97–98)

The Research Assessment Exercise that ensued was a periodic assessment (reporting in 1986, 1989, 1992, 1996, 2001 and 2008) of the research outputs and activities of higher education institutions (mainly universities). A new requirement for grant applicants to UK research councils to articulate impact plans soon followed (Payne-Gifford 2013) and, in 2009, the Higher Education Funding Council of England (HEFCE) announced that the successor to the Research Assessment Exercise, the REF, would incorporate impact as an assessment criterion, focusing on the contribution of science to public policy making and to public engagement (Williams and Grant, 2018).

It was clear from the start that measuring ‘impact’ was no easy task. The perceived problems with the measurement of research impact are even greater than those associated with measuring the quality of research outputs since impact is elusive, complex and context-dependent (Smith et al. 2020). A range of options for approaching this challenge were considered, informed by a RAND report that had been commissioned by HEFCE (Grant et al. 2010). The report recommended that an approach combining quantitative and qualitative measures should be employed and argued that a case study approach could mitigate the fact that ‘the issue of attribution is complex’ (Grant et al. 2010, xi). The UK appeared to follow this advice, employing impact case studies that allowed for a mixture of data forms.

This paper aims to assess how REF impact case studies have functioned as performance assessment tools of research impact, enabling the governance of this area within UK universities. It draws on a combination of documentary analysis (of REF impact case studies) and interviews and focus group discussions with academics and senior knowledge exchange professionals. Theoretically, we combine insights from the sociology of quantification (demonstrating how and why audits create objects through commensuration) with discursive institutionalism (explaining how and why particular ideas about high scoring REF case studies become institutionalised in ways which increasingly restrict the creative space supposedly afforded by these non-quantitative tools). Our analysis, which intentionally looks across diverse disciplines, highlights the importance of persuasive storytelling and identifies just four common ‘plot-lines’ which, while perceived by our participants to tell impact stories within the required REF format (see Box 1), sit uneasily with their much more complex accounts of the various ways in which academic work shapes life beyond the academy in practice.

2. Theorising assessment regimes

2.1. Performance measurement and governance

The Research Excellence Framework is the recent manifestation of the move towards quantified accountability and the rise of ‘regimes of measurement’ in higher education (Espeland and Stevens 1998, 402). In order to achieve this, such phenomena as ‘research excellence’ and (most recently) ‘impact’ have to be transformed into ‘auditable objects’ (Power 2015). They are reduced

in complexity and heterogeneity through the process of commensuration: ‘the transformation of qualities into quantities that share a metric’ (Espeland and Sauder 2007, 14). Calculative devices (forms and templates) act as material ‘inscriptions’ (Latour 1987) through which real-life phenomenon is translated into the physical form (such as a map, a document or, in our case, an impact case study).

The case study template is a tool of theorisation and sense-making (Power 2015) about the relationship between the academics, their organisations and their broader socio-economic environments. To achieve their purpose, impact case studies must act as boundary objects (Star and Griesemer 1989); they have to be stable enough to carry the meaning of ‘impact’ but at the same time be malleable to allow for multiple interpretations (in our case: interpretation of impact by different disciplinary communities and external impact assessors from policy, practice and industry). Impact case studies translate abstract objectives into workable, specific actions and practices. In so doing, they not only describe the practices of impact but ‘enact’ particular ideas of impact (Law and Urry 2004), which then (as we discuss) become embedded within institutions.

The performance assessment literature tells us that calculative devices quickly became tools of governing; as such, we must expect REF impact case studies to follow this trajectory within the universities. And indeed, evaluation methods implicitly involve defining what ‘good research’ is and, consequently, steer scientific practices (Rijcke et al. 2016). Universities are reactive to these pressures – they adapt their strategies, communication styles and internal structures to respond to external performance assessment systems (Woelert and McKenzie 2018; Espeland and Sauder 2007). Unsurprisingly, therefore, existing research documents various ways in which REF impact has become embedded within university governance, including via the broadening of career progression criteria (Bandola-Gill 2019), changes to internal managerial structures to provide impact leadership and lines of accountability (Power 2015), and dedicated knowledge exchange and research impact experts, who are helping write impact into university strategies, missions and (for REF2021) REF environment statements (Smith et al. 2020).

2.2. Narratives in performance measurement

Despite a rich scholarship on the role of metrics in constructing and performing social realities, less attention has been paid to qualitative formats within this. This omission is significant as narratives have been shown to have powerful governing effects. Decision-makers ‘think in terms of stories’ (Kaplan 1986, 771) – narratives shape the predominant framings of problems in ways that point us towards particular solutions and stakeholders (and away from others) (Roe 1994). They employ an interpretative angle (Czarniawska 1997), thereby imbuing particular problems and their solutions with normative value (Feldman et al. 2004). This paper explores a specific genre of narratives: case studies – described as analytic narratives (Alexandrova 2009) since they aim to achieve learning as well description. This function is enabled through ‘narrative explanation’ (Morgan 2019), oriented towards identifying causal mechanisms within presented stories.

Narratives – and their governing effects – are shaped by their constitutive elements. A narrative is ‘a sequence of events, experiences of actions with a plot that ties together different parts into a meaningful whole’ (Feldman et al. 2004, 148; see also Czarniawska 1997). Scholars working with narrative approaches have identified key structures as the plot, the heroes and the moral (McBeth et al. 2007). A particular significance is attached to the plot as an organising structure that attaches meaning to a series of events (Roe 1994). As McBeth et al. (2007, 540) argue: ‘plots serve to link characters to settings, assign the roles of the characters, and, most importantly, assign blame through some assertion of causation (while usually assigning intent as well)’. Characters help to identify the key groups affected by the problem (Roe 1994), whereas the moral of the story implies a solution (McBeth et al. 2007).

Despite this rich (and growing) literature on the role of narratives in policy and organisations, less is known about their role as ‘tools’ of governing. This is important as these formats are increasingly

used in research evaluation systems, without a closer reflection on *how* they perform their goals. We begin addressing this gap by examining how impact case study narratives construct and perform ideas of impact, exploring the role these particular narrative-based performance assessment tools are playing in the governance of higher education in the UK.

3. Impact in REF

REF2014 defined impact as: 'an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia' (REF 2011, 31), a definition which remains in place for REF2021 (REF 2019). The format of impact case studies (for both REF2014 and REF2021) is set out in [Box 1](#).

Box 1. The case study template

- (1) Summary of the impact (100 words)
- (2) Underpinning research (500 words)
- (3) References to the research (maximum of six references)
- (4) Details of the impact (indicative maximum 750 words)
- (5) Sources to corroborate the impact (indicative maximum of ten references)

Source: REF (2019).

The REF is divided into units of assessment that map onto specific disciplines or research areas. The number of impact case studies to be submitted is determined by the number of full-time research-active staff who are returned by each institution, with a requirement for two case studies per 0-19.99 full-time academics returned to each unit of assessment (REF 2019). Case studies are assessed by peer-review, conducted by a panel of experts (both academic and non-academic), based on two criteria: reach and significance (REF 2011). In REF2014, impact case studies accounted for 20% of the total score but this has been increased to 25% for REF2021 (REF 2019).

The results are translated into metrics via a ranking system that runs from 4* (the highest) to 0 (the lowest). A pilot study using impact case studies to assess research impact in 29 UK higher education institutions was broadly positive about the flexibility provided by narrative case studies (HEFCE, 2010). Since then, comparative assessments have continued to provide support for the UK's decision to employ a narrative approach to assessing impact on the basis that this allows for greater variation and complexity than simpler and cheaper bibliometric (or 'altmetric') approaches (see Penfield et al. 2014). Key reports, including the government-commissioned Stern Review, have recommended the continued use of narrative based impact case studies for this reason (Stern 2016).

4. Methods

This research draws on two main sources of data: interviews/focus groups and document analysis. The first source comprises two focus groups (with 7 and 5 participants respectively) and 12 individual interviews (see [Table 1](#)) at a major research-intensive university in the UK. Participants were selected based on their disciplinary background (with the aim of achieving sampling diversity) and their experience with, or interest in, the UK's approach to research impact. The interviews and focus groups were semi-structured and covered four main topics: definitions of impact, characteristics of a good impact case study, barriers and facilitators of the process of writing an impact case study, and institutional and organisational support for impact evaluation. Conversations were digitally recorded and transcribed. The data were coded using thematic analysis, involving an iterative process of coding and re-coding.

Table 1. List of the study participants.

Data collection	Total no. of participants	Disciplinary background
Focus Group 1	7	Informatics (1) Knowledge Exchange professional (1) Languages (1) Nursing (1) Public health medicine (1) Political science (1) Psychology (1) Sociology (1)
Focus Group 2	5	Animal Science (1) Biomedical science (1) Dentistry (1) Neurology (1) Psychiatry (1)
Individual Interviews	12	Chemistry (1) History (1) Informatics (1) Law (1) Maths (1) Physics (3) Knowledge exchange professional (1) Education (1)
Total	24	18 disciplines plus knowledge exchange professionals

The selection of REF case studies was guided by our assessment of interview and focus group data against the widely-cited Becher-Biglan typology of academic disciplines categorising disciplines along two dimensions: soft–hard and pure–applied (Becher 1989). Looking at our data, we found that a distinction between applied and pure science was, unsurprisingly (given our focus on impact), strongly apparent. However, we could see no obvious distinction in our data between ‘hard’ and ‘soft’ disciplines (i.e. we could see no evidence in our data of a hard/soft disciplinary distinction playing a role in different views about, or experiences of, research impact). We also noted that our data implied a distinction between ‘applied’ and ‘soft’ that was more of a spectrum than a binary categorisation. Hence, in identifying disciplines to be included in our case studies, we used our data to place disciplines along a pure–applied spectrum. At one end, we identified strongly applied disciplines that tended to involve training people to work in particular settings or directly producing guidance for such settings (e.g. computing, medicine, social work, law, education). At the other end, we placed disciplines that were more commonly focused on producing ‘pure’ knowledge involving the construction of new ideas (e.g. history, theoretical physics and mathematics). These disciplines tended to focus on publics as their main audience and were the disciplines that seemed to struggle most with REF impact case studies. In the middle, we identified a cluster of social science disciplines producing ‘critical’ knowledge. These disciplines are concerned with issues that are already being discussed in the public/policy domain but they tend to produce knowledge that pushes against current trajectories (e.g. sociology, geography and politics). Although work in these disciplines often has a clear policy focus, our participants reported struggling to achieve impact in ways that were documentable for REF.

The choice of the specific units of assessment within these three clusters was guided by the principles of case study selection (Seawright and Gerring 2008), highlighting the value of selecting cases that are similar enough to support identification of common themes and different enough to enable analytical value of comparison. We selected one discipline from each of the three clusters discussed above, focusing on disciplines that were similar in terms of their focus on socio-political issues (rather than economic/technological or cultural ones). We chose three units of assessment: (1) Public Health, Health Services and Primary Care; (2) Sociology; and (3) History.

To analyse how each approached REF2014 impact case studies, and to gain more insights as to how different approaches fared in this system of performance assessment, we analysed the impact

case studies produced by the three highest and the three lowest-scoring universities for each of these disciplinary units of assessment, based on the publically available results of REF2014 (<https://results.ref.ac.uk/>). For Public Health, three universities were equally positioned as the third-highest scoring university so we selected one randomly (using a random number generator). The summary of the 66 case studies analysed for this paper is presented in Table 2 and the original data are available at <https://impact.ref.ac.uk/casestudies/>.

We employed thematic analysis to assess these case studies. First, we drew on our interview/focus group data and a review of existing literature (Smith et al. 2020) to develop a matrix that asked six questions of the REF impact case studies (see Table 3). We piloted this matrix on six randomly selected case studies (ensuring these covered all three units of assessment) as well as presenting it (via seminars) to academic audiences with an interest in research impact, where we also sought feedback. The piloting experience and the feedback we received led us to simplify the matrix for data extraction, to focus on examining the plot (i.e. the description of the baseline situation and underpinning research as well as reports of strategies/actions to achieve impact), the heroes (the central actors in the case study) and the moral (the outcome or overarching message, which often employed quantification), as well as noting observations around the narrative type. Having first organised our sample impact case studies into similar plot clusters – ones with similar strategies, target groups, outcomes, etc., we identified four clear archetypal narrative types. We then grouped the REF impact case studies into these narrative types and undertook further analysis in which we sought to identify the key characteristics of each narrative type and to confirm the fit of each case study into these categories.

5. Findings – REF impact case studies as narrative tools

5.1. Restricting impact storytelling

The case study template to which all REF impact case studies must adhere (Box 1) incorporates section word limits and a five-page maximum, all of which functions to restrict the ways in which stories of research impact can be told. It was notable that several of the academics we interviewed differentiated between meaningful research impact and ‘REF impact’, which was positioned as impact narrated for a bureaucratic exercise. For example, in the following extract the interviewee

Table 2. Sample of impact case studies analysed.

Unit of assessment	High scoring		Low scoring	
	University	Number of case studies	University	Number of case studies
Public Health, Health Services and Primary Care	University of Bristol	7	Queen’s University Belfast	4
	University of Oxford	6	Leeds Beckett University	3
	University of Sheffield	9	University of Manchester	4
Sociology	University of York	3	University of Winchester	2
	University of Lancaster	3	University of Abertay Dundee	2
	University of Cardiff	4	Queen’s University Belfast	3
History	University of Hertfordshire	3	University of Sunderland	2
	University of Aberdeen	3	University of Westminster	2
	King’s College London	4	Anglia Ruskin University	2
		42		24

Table 3. The original and revised analytic matrix used to analyse our sample REF impact case studies.

Questions in our original matrix (used during piloting)	Mapped to categories in our revised matrix (used to analyse the final set of 66 REF impact case studies)
What is the problem the case study addresses? What strategies were employed?	Plot
What is the link between research and impact? What is the outcome of the impact activities?	Moral
How is the benefit to society evidenced? Who are the key actors/stakeholders involved in the case study?	Heroes
What is the main narrative presented in the case study?	Narrative type

suggests that universities are interpreting REF requirements for impact case studies in ways that mean it becomes little more than a ‘box-ticking’ exercise:

You’ve got this general word of ‘impact’, which is how well you serve society. But then there’s something very specific about how well you tick the right boxes for the REF impact case studies. And it seems like other people are making decisions about that somewhere else in the university. I don’t know, there’s something slightly odd about that, that this word carries two different meanings: a more general and very specific sense. (Academic, Social Science)

Our participants also included two knowledge exchange (KE) professionals and, although they did not dismiss the REF approach to impact as ‘box-ticking’, they nonetheless acknowledged a similar distinction between the messy realities of impact and the much neater accounts required for impact case studies. Across disciplines and professions, participants suggested that it was essential to create a persuasive narrative to attract a high REF impact score:

The key question would be how do you express it [impact] in a convincing way? Because it could be quite hard to come up with really definitive evidence. I think the advice I always gave to people when I was advising on putting case studies together was that ... what people need to do is be honest with themselves about whether they have a convincing argument. (Academic, Math)

Our data suggest that universities, as institutions, are helping to craft what is understood to represent ‘convincing arguments’ within impact case studies, via training, sharing of ‘best practice’ examples and processes of internal review and assessment. For example:

[We] had meetings and workshops and whatever where you could go along to learn how to write your impact assessment, and then they selected from the submitted impact assessments. [...] It was used as a lever of power and influence by our administrators. (Academic, Medicine)

Therefore, the art of creating successful impact narratives was shaped by institutional processes. The following sections explore the three composite elements of narratives discussed in the literature (the plot, the moral, the heroes) and how they ought to be articulated, according to the participants, in order to succeed in REF terms. We then turn to our analysis of the four archetypal storylines we identified in our REF2014 sample.

5.2. The plot: defining and shaping performance

One of the key characteristics of impact case studies is their linearity. The REF case study format (Box 1) consists of sections requiring reporting on research publications followed by sections on the impact(s) achieved, which implicitly assumes an ordered move from research to impact. This organises the case studies in an inherently plot-like structure, implying an expectation of what one interviewee called ‘a clear flow going through [the case study]’ (KE Professional).

This assumption of causality was in striking contradiction to the way participants perceived realities of achieving impact. Regardless of the target for impact (e.g. policy, industry, culture or society), routes to achieving impact were consistently articulated as being more complex, unpredictable and serendipitous than REF case studies suggest. For example:

A Nobel physics prize winner [...] accidentally discovered magnetoresistance [which meant] you could suddenly buy really big hard disc drives for your laptop. This chap [...] discovered it entirely accidentally and it was research that has been funded by a company in Germany and he told them, 'I've seen this really weird effect, are you interested in patenting it?' They said, 'no, no interest at all'. And ten years down the line, it was picked up and turned into this product. So it was entirely an accidental series of events and I don't think government understands the impact of accident. You know, it's more complicated than they think. (Academic, Physics)

Participants' accounts suggest that measuring impact performance via case studies simplified the complex reality by focusing only on specific aspects of the monitored phenomena. Unlike quantitative formats, which achieve this goal via commensuration (Espeland and Sauder 2007), the case study narratives achieved this via the cohesiveness of the plotline. This is reflected in the fact participants across disciplines repeatedly suggested that the parsimony of narratives differentiated high and low scoring case studies. For example:

[The impact case study] has to be lean and mean. It shouldn't be too broad. They have to have a very clear path from initiation to the point at which they have a clear impact either in a policy sense or in a product or in some kind of change in the world in some way or another. (Academic, Informatics)

The task of articulating a clear storyline appeared to be challenging as it required case study authors to organise a set of multiple, at times disconnected, events and projects into a story that was cohesive, concise and organised according to REF performance assessment criteria. Perhaps unsurprisingly, our analysis of REF2014 case studies suggests this resulted in some consistent narrative structure traits. One common approach, for example, was for case studies to first identify a baseline situation, in order to present the state of the problem, then to clearly outline the change (or 'impact') against this baseline (71% of high scoring and 67% out of low scoring cases employed this structure).

The preparation of case studies for assessment thus required any 'plot holes' in reported cases to be 'filled' and participants suggested that linking research to impact was a challenge across disciplines. Yet, the precise nature of these 'plot holes' seemed to vary with disciplinary context. The more applied researchers (those working in Education, Medicine or Public Health) appeared to find it relatively easy to identify the 'impact' of their work and the stakeholders with whom they were working, and rarely struggled to provide documentary sources to corroborate these accounts. However, a tendency to develop proposals for policy and practice based on broad portfolios of knowledge and experience (e.g. research syntheses) meant it was sometimes difficult to link the main characters in these impact case studies with original research outputs:

I think there were potential cases in health where people had helped with an intervention to change or inform practice in health. But where it was more based on ... not common knowledge but not their own research. Perhaps research from other areas [...]. The material is not based on the individual's own research, it's more a compilation of or knowledge from elsewhere. (Academic, Medical Science)

In contrast, researchers working in less applied areas appeared to more commonly struggle with plot holes relating to the outcomes and the possibility of claiming a link between research and impact. For example:

We [as a discipline] say: 'oh we've invented the internet', people kind of laugh at you. It may be true but the consequences of that development are just so enormous that it's difficult to comprehend. And ... a lot of impact that we have, especially the economic impact, is by chance ... it's the sort of lucky, side benefits of what we do. Our research projects are not to develop a better lightbulb, although it may turn out that the technology that you've developed generated the LED lightbulb, which is now ubiquitous across all of society. That would not have happened without physics and yet we can't make an impact case saying well here's the lightbulb that we generated because that's not the way it works. We developed the technology which then others will take an impact against. (Academic, Particle Physics)

As the above extract illustrates, the challenges here seemed to relate to the complexity of research-based innovation and policy change, plus the multifaceted approach to translating

research. In both cases, what is made clear is that there are often far more ‘characters’ involved in achieving impact than the short format of an impact case study allows (a point we return to in section 5.4).

5.3. Moral of the story – defining impact

The next element (intersecting with the plot) is the ‘lesson’ that brings the narrative to its conclusion: the ‘moral’ of the story. In impact case studies, this seemed to require proof of achieving a specific (always positive) change on the basis of research. Formally, the specific, disciplinary forms of impact were not predefined beyond presenting illustrative instances of possible impacts within the units of assessment (REF 2010). Regardless of the target groups and impact types, the vast majority of interviewees perceived the specificity of the outcome to be the most important characteristic of highly scoring impact case studies. To work well, participants consistently suggested that the case study narrative had to have a clear ending: a well-evidenced change when contrasted with the pre-intervention situation:

I think that for all of [the disciplines] a good case is when you have a body of research which needs to be evidenced. But from that, you have clear insights and findings. So it’s not woolly – I did research on de, de, de. But this research found de, de, de and the insights were de. So it’s very clear what came out of it. (KE Professional)

It was notable that participants seemed to assume (as the REF criteria also assume – Smith et al. 2020), that this change would be positive. The vast majority of participants suggested that this change had to be presented in concrete terms, both in terms of the affected groups and in terms of the outcomes. For example:

It’s no longer enough to, on the basis of well-designed research, to suggest impact, and this is societally relevant and obvious, this is a big societal problem that we’re doing a little bit of work on. You have to get very specific in terms of who are the non-academic people that you’re going to interface with, and how are you going to do it and what difference is it going to make to them. (Academic, Sociology)

The expectation of specificity regarding change resulted in a variety of discursive strategies employed by case study authors. One strategy (employed in 90% of the high scoring and only 54% of the low scoring case studies) was to strengthen the end of the story by incorporating material that would enhance the sense of ‘objectivity’ – notably by adding quantifiable results as ‘proof’. The apparent attraction of incorporating quantifiable indicators into impact narratives steers the performance of impact (and perhaps the case study selection) towards instrumental (rather than broader, conceptual) impacts since the latter are recognised as being harder to document (Smith et al. 2020). This trend was observable across disciplines and has been documented in earlier discipline-specific analysis (e.g. Meagher and Martin 2017; Smith and Stewart 2017) and in the government-commissioned Stern Review (2016).

A large number of participants discussed using various forms of metrics to show that research resulted in economic, social and legal benefits and it was notable that impacts that could be monetised were seen as particularly attractive by participants. For example:

A good impact case has a very clear connection between the science paper that generated the advance and a dollar sign at the end of it, of this is what it’s delivered. (Academic, Physics)

This was a subtle yet powerful expression of the belief (commonly articulated by our participants) that the ‘public value’ of research signifies the economic value. This reflects a broader trend of marketisation in academia, where knowledge is conceived as a form of capital (Olssen and Peters 2005) and universities are framed as ‘engines’ of economic growth (Berman 2011). This understanding of impact – reflected in both participants’ accounts and the case studies we analysed – was grounded in broader assumptions that REF was a process for illustrating the value for money of UK university research, with one focus group participant framing this as a need to demonstrate how academics are contributing to ‘UK plc’.

Hence, despite the qualitative nature of the case study format, the concise reporting format appears to invite the metricisation of impact (across disciplines) to concisely convey the ‘objectivity’ of impact claims, providing narratives with a clear and persuasive ending.

5.4. Heroes of the story – individualising impact

The final element of impact narratives is the heroes: the individuals whose experiences are at the centre of the story. Case studies individualised impact performances by focusing on personal responsibility for achievements (Dunlop 2018). These ‘heroic impact narratives’ (Thomson 2013) were consistently described by participants as being inherently inadequate for capturing the complex and diffuse ways in which research contributes to society; contributions almost always described as collaborative (see section 5.2). Yet, REF impact case studies were interpreted as requiring a transferral of responsibility for achieving impact onto individual academics in ways that enabled them to feature as the story’s main characters. This positioning of individuals at the centre of impact case studies was reflected in the following extract, in which an interviewee recalled their team shifting their impact strategy from a team to an individual focus:

I think the idea that you can take an individual and slice them to separate impact cases is something that was a revelation to people here, because they kind of thought it had to be like a group. (Academic, Informatics)

In order for the academic characters of impact stories to appear ‘heroic’, other characters – notably the recipients of research-based advice (but sometimes, as above, also research team members) – were necessarily present but often incongruously passive. Indeed, impact targets were rarely described as active contributors to research impact other than, for example, as occasional research commissioners (see also Smith et al. 2020).

Having been rendered passive, there was a clear hierarchy of preferred audience type for impact case studies. While academics in the social sciences and public health more commonly focused on policy and practice audiences, and historians on publics, the latter were consistently described as a less desirable impact audience. Our data suggest this was because it is difficult to conclude case studies with ‘objective’ (e.g. quantified) indicators of impact for large and varied public audiences. This reflects earlier research (e.g. Watermeyer and Chubb 2019; Wilkinson 2019; Derrick 2018) which has identified widespread uncertainty among academics and impact assessors about the relationship between public engagement and research impact, and a sense that (with some discipline specific exceptions) focusing on public engagement is a risky route to impact.

5.5. Narrative archetypes and making narratives comparable

Taken together, these three consistent structural elements of case studies (the plot, the moral and the heroes) seemed to reduce the complexities of impact stories. This narrowing appears, in turn, to have limited the range of narrative types, resulting in relatively homogeneous approaches to reporting impact stories. Our thematic analysis of impact case studies identified only four narrative types: Reframing, Problem-solving, Tool-making, and Public Engagement (Table 4). These four types reflect existing literature on knowledge utilisation (e.g. Weiss 1979), incorporating both direct, instrumental impacts and indirect, conceptual impacts. However, the fact we identified only four narrative types seems surprising, given both the complexity of participants’ accounts of ‘real world’ impacts and the fact that a key rationale for employing case studies in REF was to enable variety. In the remainder of this section, we consider each narrative type in more detail.

5.5.1. Problem-solving

Problem-solving, the most common narrative, started by outlining existing problems, then positioned research as providing solutions, ending with an account of how the situation had been improved or resolved by the research. Such narratives entailed, for example, informing service

Table 4. The summary of impact narratives.

Type of narrative	Summary	Prevalence	
		High scoring cases	Low scoring cases
Problem-solving	Research is offering solutions to existing policy and practice problems.	23	8
Tool building	Research is providing specific tools to support policy and practice.	8	5
Reframing	Research is reframing existing account of problems and/or solutions.	5	3
Public engagement	Research is facilitating learning and deliberation across various stakeholders.	6	8

provision (Cardiff University 2014a), providing evidence to approach problems, including cot deaths (Bristol University 2014a), suicide (Bristol University 2014b) and various health problems (e.g. Oxford University 2014a), or changing official guidelines (Sheffield University 2014a). These types of narratives often involved specific forms of knowledge production, going beyond traditional scientific activities, including advisory work, contracted research or evaluations and participatory formats (e.g. Lancaster University 2014b). This type of narrative was particularly popular in public health (77% of high scoring and 45% of low scoring cases) and sociology (60% of high scoring and 43% of low scoring cases). It also appeared to perform well in REF (Table 4). Both its popularity and success seem likely to relate to the linearity of this narrative type, and the clear relationship that it provides between the plot and the outcome.

5.5.2. Tool-building

A subset of the problem-solving narrative was tool-building. The core of the narrative structure was similar to problem-solving but was unique in the specificity of the solution, with research presented as the basis of evidence ‘tools’ which supported decision-makers to reach evidence-informed choices. These ‘tools’ included economic models (Sheffield University 2014b), indicator sets (Sheffield University 2014c) and historical databases (University of Hertfordshire 2014a). This type of a narrative was particularly popular in public health (23% of high scoring and 36% of low scoring).

5.5.3. Public engagement

The next most popular narrative type was public engagement, which involved collaboration with various stakeholders to either produce or disseminate knowledge. It involved presenting situations in which issues or ideas were presented as relatively unknown (in general or among particular audiences) prior to the actions of researchers, who had worked to enable various audiences to gain new knowledge or awareness. This type of narrative was most common in history (60% of high scoring and 63% of low scoring cases), aligning with our participants’ accounts of the importance of public audiences for humanities and basic sciences. This approach often had an element of ‘enlightenment’ (Weiss 1979) as the outcome of the story, particularly in high scoring cases. For example, shedding light on lesser-known topics which have led to improvements in economic performance of heritage sites (University of Aberdeen 2014) or informing and raising awareness of specific historical events and problems (University of Hertfordshire 2014b). These narratives were often educational (e.g. providing learning via exhibitions and media appearances: University of Westminster 2014) rather than dialogical.

Even though this narrative was most commonly employed in History, it was also identified in lower scoring cases in Sociology (43%) and Public Health (9%). Amongst all the narrative types, it was the only one with a higher number of low-scoring cases in our analysis (Table 3). In the lower scoring cases that employed this narrative type across all three disciplines, the narrative involved participation in debates and/or forms of dissemination but lacked clear, positive outcomes (i.e. a ‘moral’ to the story).

5.5.4. Reframing

The least common narrative type involved reframing. Like public engagement, this narrative focused on the role of research in shedding new light on an issue; however, here research was used to transform or adapt existing conceptualisations of problems, leading to, for example, new policy framings. This narrative type was particularly evident in Sociology (20% of high scoring and 14% of low scoring) and History (respectively 30% and 17%). Examples of such narratives included changing approaches to policing (Cardiff University, 2014b), reframing climate change policies from individual attitudes and behaviours to social practices (Lancaster University 2014a) and introducing 'historical thinking' to policymaking (e.g. King's College London 2014). It is worth noting that this type of a narrative was described by our participants as one that was most often excluded from submission by the universities, as it was perceived to be risky and not fitting the institutionalised ideas of a 'good' case study.

6. Concluding discussion – governing by narratives

This paper explores impact case studies as technologies of governance that – akin to any calculative instruments – achieve specific governing effects. Nominally, the REF case studies were employed as reporting devices, aimed at recording social and economic benefits of produced research. But, like other formats of performance measurement (Power 2004), the case studies entailed an inherently normative core, prescribing future behaviours on the basis of the reported ones. In the concluding section we unpack two specific implications of 'governing by narratives': first, impact case studies achieved governing effects by defining an idea of what 'performance' means in this context; second, impact case studies were institutionalised within the broader (competitive) performance assessment infrastructures of UK universities. The significant financial implications of impact case study performance further steered behaviour towards specific forms of practices of storytelling deemed most likely to perform well.

Impact case studies, as reflected in our analysis, defined the idea of impact (previously seen as uncommensurable) and as such established a bounded field of practice – one that could be monitored and hence governed. The scholarship in the sociology of quantification has shown that translating complex and idiosyncratic phenomena – such as 'impact' – into auditable objects occurs through commensuration (Espeland and Stevens 1998). In the case of impact narratives, we demonstrate that impact has been constructed via a particular narrative structure (a clear plot line, a moral and a small number of heroes), standardising *format* rather than *content* within the performance measurement. Within the case studies, the heroes were the academic researchers helping other stakeholders achieve change, the moral involved recounting the 'effects' of impact activities, and the plot imposed a causal structure on the performance (a beginning, a middle and a clear and convincing 'end'). This standardisation resulted in just four narrative types within our sample of 66 case studies: Problem-solving, Tool-making, Public Engagement, and Reframing. Even though the case studies escaped some of the perils of quantification by allowing relatively flexible case study content, they also narrowed the scope of storytelling. This was particularly evident in the way high performing case studies used quantification (especially numbers expressing an economic value) to help present the culmination of the moral with a convincing plot ending.

This focus on economic value was not merely a matter of narrative construction. Our research has shown that narratives used as technologies of assessment not only presented depictions of impact but actively define – and therefore 'enact' (Law and Urry 2004) – the idea of impact, guiding future impact practices. The UK's REF impact case studies are performative (MacKenzie, Muniesa, and Siu 2007); they not only describe but also construct a very specific vision of impact, as streamlined, linear and preferably economic. The high value of impact case studies that perform well in REF means this idea becomes incentivised and institutionalised within the universities. This narrowing

effect is perhaps particularly evident in a context that has not previously been subject to performance assessment which now transformed into an established – and highly governable field.

Building on this, our second insight relates to the positioning of universities as gatekeepers of impact narratives. As our interviewees described, universities invest in mechanisms to support the production of successful impact case studies (e.g. training courses and internal management and assessment processes). This transformation of ‘impact’ into an object of an audit renders stories of impact more easily comparable and, here, it is notable that not all of the narrative types we identified performed equally well. Specifically, in our sample, ‘public engagement’ narratives performed less well in REF than the other three types (see also: Watermeyer and Chubb 2019, Watermeyer 2015). This is despite the fact that HEFCE’s (2009) stated ambition for incorporating impact into REF included incentivising academics to undertake public engagement (see Introduction).

While it is beyond the scope of this article to assess these mechanisms, our interviewees noted that these processes often function to institutionalise restricted ideas about what ‘successful’ REF impact case studies look like, filtering out case studies that do not conform (Schmidt 2008). In effect, certain ideas about what REF ‘counts’ as success for impact case studies have gained traction in the UK and universities have institutionalised these perceptions, encouraging case study authors to employ ‘successful’ narrative tropes. All this helps explain why we identified just four narrative types, despite examining case studies from very different disciplines and topics. Such accounts sit uneasily with the complex ways in which participants described impacts occurring in practice. Indeed, several participants distinguished ‘REF impact’ from meaningful research impact, with the former often being derided as a ‘box-ticking’ exercise that produces ‘fairy tale’ versions of impact.

Taken together these two effects – of performativity of the tools of assessment and of institutional filtering of successful case studies – point to the emergence of the (relatively narrow) impact field in a very specific form within higher education in the UK. The history of REF (outlined in the Introduction) suggests that the UK’s approach to performance assessing impact via case studies, aimed to achieve audit without the restrictions of quantification. Nevertheless, our analysis suggests that the way in which REF case studies are being produced and assessed has resulted in significant standardisation, with effects that are not dissimilar to quantified forms of performance measurement (Power 2004). There are, of course, pragmatic benefits to this effect; in a performance assessment system designed to inform resource allocation decisions, it is desirable to be able to compare and contrast – a process that standardisation facilitates. However, the risk is that the significant narrowing down of the practices considered to constitute ‘impact’ informs the emergence of monoculture of impact where only the types of activities that are believed to fit the audit criteria become incentivised and legitimised by the universities. This, in turn, risks hindering experimentation and creativity and, our findings suggest, squeezing resources to support wider public engagement practices.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by University of Edinburgh.

ORCID

Justyna Bandola-Gill  <http://orcid.org/0000-0002-3418-4085>

Katherine E. Smith  <http://orcid.org/0000-0002-1060-4102>

References

- Alexandrova, A. 2009. "When Analytic Narratives Explain." *Journal of the Philosophy of History* 3 (1): 1–24.
- Bandola-Gill, J. 2019. "Between Relevance and Excellence? Research Impact Agenda and the Production of Policy Knowledge." *Science and Public Policy* 46 (6): 895–905.
- Becher, T. 1989. *Academic Tribes and Territories*. Buckingham: Society for Research into Higher Education & Open University Press.
- Berman, E. P. 2011. *Creating the Market University: How Academic Science Became an Economic Engine*. Princeton: Princeton University Press.
- Bristol University. 2014a. *Bristol Research Leads to a Worldwide Fall in the Number of Cot Deaths*.
- Bristol University. 2014b. *Fewer Suicides Worldwide Following Changes in Policy and Practice Influenced by University of Bristol Research*.
- Bulmer, M., ed. 1987. *Social Science Research and Government: Comparative Essays on Britain and the United States*. Cambridge: Cambridge University Press.
- Cardiff University. 2014a. *Improving the Response to Victims of Violence*.
- Cardiff University. 2014b. *Improving Community Policing*.
- Czarniawska, B. 1997. *Narrating the Organization: Dramas of Institutional Identity*. Chicago: University of Chicago Press.
- Derrick, G. 2018. *The Evaluators' Eye: Impact Assessment and Academic Peer Review*. London: Palgrave Macmillan.
- Dunlop, C. A. 2018. "The Political Economy of Politics and International Studies Impact: REF2014 Case Analysis." *British Politics* 13 (3): 270–94.
- Espeland, W. N. 1998. *The Struggle for Water: Politics, Rationality, and Identity in the American Southwest*. Chicago: University of Chicago Press.
- Espeland, W. N., and M. Sauder. 2007. "Rankings and Reactivity: How Public Measures Recreate Social Worlds." *American Journal of Sociology* 113 (1): 1–40.
- Espeland, W. N., and M. L. Stevens. 1998. "Commensuration as a Social Process." *Annual Review of Sociology* 24 (1): 313–43.
- Feldman, M. S., K. Sköldbberg, R. N. Brown, and D. Horner. 2004. "Making sense of stories: A rhetorical approach to narrative analysis." *Journal of public administration research and theory* 14 (2): 147–170.
- Feller, I. 2009. "Performance measurement and the governance of American academic science." *Minerva* 47 (3): 323–344.
- Grant, J., P. B. Brutscher, S. E. Kirk, L. Butler, and S. Wooding. 2010. *Capturing Research Impacts: A Review of International Practice. Documented Briefing*. Rand Corporation. https://www.rand.org/pubs/documented_briefings/DB578.html
- HEFCE. 2009. *Research Excellence Framework: Second Consultation on the Assessment and Funding of Research*.
- HEFCE. 2010. *Research Excellence Framework impact pilot exercise: Findings of the expert panels. A report to the UK higher education funding bodies by the chairs of the impact pilot panels*. https://www.ref.ac.uk/2014/media/ref/content/pub/researchexcellenceframeworkimpactpiloteexercisefindingsoftheexpertpanels/re01_10.pdf.
- Hoggett, P. 1996. "New Modes of Control in the Public Service." *Public Administration* 74 (1): 9–32.
- Kaplan, T. J. 1986. "The Narrative Structure of Policy Analysis." *Journal of Policy Analysis and Management* 5 (4): 761–78.
- King's College London. 2014. *Happy Families? Embedding History in Policy Making at the Department for Education*.
- Kogan, M., and S. Hanney. 2000. *Reforming Higher Education*. London: Jessica Kingsley Publishers.
- Lancaster University. 2014a. *From Attitudes to Practices: New Approaches to Climate Change Policy*.
- Lancaster University. 2014b. *Shaping Ethical Provision, Care Quality and Design Sensitivity in New Health Technologies*.
- Latour, B. 1987. *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, MA: Harvard University Press.
- Law, J., and J. Urry. 2004. "Enacting the Social." *Economy and Society* 33 (3): 390–410.
- MacKenzie, D. A., F. Muniesa, and L. Siu, eds. 2007. *Do Economists Make Markets? On the Performativity of Economics*. Princeton: Princeton University Press.
- McBeth, M. K., E. A. Shanahan, R. J. Arnell, and P. L. Hathaway. 2007. "The Intersection of Narrative Policy Analysis and Policy Change Theory." *Policy Studies Journal* 35 (1): 87–108.
- Meagher, L. R., and U. Martin. 2017. "Slightly Dirty Maths: The Richly Textured Mechanisms of Impact." *Research Evaluation* 26 (1): 15–27.
- Morgan, M. S. 2019. "Exemplification and the Use-Values of Cases and Case Studies." *Studies in History and Philosophy of Science Part A* 78: 5–13.
- Olssen, M., and M. A. Peters. 2005. "Neoliberalism, Higher Education and the Knowledge Economy: From the Free Market to Knowledge Capitalism." *Journal of Education Policy* 20 (3): 313–45.
- Oxford University. 2014a. *Hormone Replacement Therapy and Cancer Risk: The Million Women Study*.
- Oxford University. 2014b. *Shaping International and UK Tobacco Policy and Practice*.
- Payne-Gifford, S.. 2013. "What is the Meaning of the Impact Agenda: Is it a Repackaged or a New Entity? Views from Inside the Research Councils." In *Achieving Impact in Research*, edited by Pam Denicolo, 10–19. London: SAGE.
- Penfield, T., M. J. Baker, R. Scoble, and M. C. Wykes. 2014. "Assessment, Evaluations, and Definitions of Research Impact: A Review." *Research Evaluation* 23 (1): 21–32.

- Power, M. 2004. "Counting, Control and Calculation: Reflections on Measuring and Management." *Human Relations* 57 (6): 765–83.
- Power, M. 2015. "How Accounting Begins: Object Formation and the Accretion of Infrastructure." *Accounting, Organizations and Society* 47: 43–55.
- REF. 2010. *Research Excellence Framework Impact Pilot Exercise: Findings of the Expert Panels*.
- REF. 2011. *Research Excellence Framework: Assessment Framework and Guidance on Submissions*.
- REF. 2019. *Guidance on Submissions (2019/01)*.
- Rijcke, S. D., P. F. Wouters, A. D. Rushforth, T. P. Franssen, and B. Hammarfelt. 2016. "Evaluation Practices and Effects of Indicator Use – A Literature Review." *Research Evaluation* 25 (2): 161–69.
- Roe, E. 1994. *Narrative Policy Analysis*. Durham, NC: Duke University Press.
- Schmidt, V. A. 2008. "Discursive Institutionalism: The Explanatory Power of Ideas and Discourse." *Annual Review of Political Science* 11, 303–326.
- Seawright, J., and J. Gerring. 2008. "Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options." *Political Research Quarterly* 61 (2): 294–308.
- Sheffield University. 2014a. *Changed Clinical Guidelines to Improve the Diagnosis of Deep Vein Thrombosis (DVT)*.
- Sheffield University. 2014b. *Identifying Failing Hospitals: A New Measure Implemented by the NHS*.
- Sheffield University. 2014c. *The Impact of the Sheffield Alcohol Policy Model on Alcohol Policy*.
- Shore, C., and S. Wright. 1999. "Audit Culture and Anthropology: Neo-Liberalism in British Higher Education." *The Journal of the Royal Anthropological Institute* 5 (4): 557.
- Smith, K., J. Bandola-Gill, N. Meer, E. Stewart, and R. Watermeyer. 2020. *The Impact Agenda: Constructing, Debating and Challenging the Assessments of Research Impact in the UK*. Polity Press.
- Smith, K. E., and E. Stewart. 2017. "We need to talk about impact: Why social policy academics need to engage with the UK's research impact agenda." *Journal of Social Policy* 46 (1): 109–127.
- Star, S. L., and J. R. Griesemer. 1989. "Institutional Ecology, Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39." *Social Studies of Science* 19 (3): 387–420.
- Stern, N. 2016. *Building on success and learning from experience: An independent review of the Research Excellence Framework*.
- Thomson, P. 2013. "Heroic Impact Narratives Create a Dangerous Divide Between the Researcher and the Local Context." LSE Impact of Social Sciences Blog. <http://blogs.lse.ac.uk/impactofsocialsciences/2013/04/26/heroic-impact-narratives/>.
- University of Aberdeen. 2014. *Ironwork and Dover Castle: Generating Economic and Cultural Impacts Through Collaboration with English Heritage*.
- University of Hertfordshire. 2014a. *The Old Bailey Online: Democratising Access to Social History*.
- University of Hertfordshire. 2014b. *Threads of Feeling: Exhibiting the Foundling Textiles*.
- University of Westminster. 2014. *The Response of Navies to Operational and Political Needs*.
- Watermeyer, R. 2015. "Lost in the 'third space': the impact of public engagement in higher education on academic identity, research practice and career progression." *European Journal of Higher Education* 5 (3): 331–347.
- Watermeyer, R., and J. Chubb. 2019. "Evaluating 'Impact' in the UK's Research Excellence Framework (REF): liminality, looseness and new modalities of scholarly distinction." *Studies in Higher Education* 44 (9): 1554–66.
- Weiss, C. H. 1979. "The Many Meanings of Research Utilization." *Public Administration Review* 39 (5): 426–31.
- Wilkinson, C. 2019. "Evidencing impact: A Case Study of UK Academic Perspectives on Evidencing Research Impact." *Studies in Higher Education* 44 (1): 72–85.
- Woelert, P., and L. McKenzie. 2018. "Follow the Money? How Australian Universities Replicate National Performance-Based Funding Mechanisms." *Research Evaluation* 27 (3): 184–95.