How to evaluate *ex ante* impact? An analysis of reviewers' comments on impact statements in grant applications

*Lai Ma[1]
Junwen Luo[1]
Thomas Feliciani[2]
Kalpana Shankar[1]

*Corresponding author – Email: lai.ma@ucd.ie
[1]School of Information and Communication Studies, University College Dublin, Ireland
[2]School of Sociology and Geary Institute for Public Policy, University College Dublin, Ireland

Abstract
Impact statements are increasingly required and assessed in grant applications. In this study, we used content analysis to examine the "comments on impact" section of the postal reviews and related documents of Science Foundation Ireland's Investigators' Programme to understand reviewers' *ex ante* impact assessment. We found three key patterns: (a) reviewers favoured short-term, tangible impacts, particularly commercial ones; (b) reviewers commented on process-oriented impact (formative) in a more concrete and elaborate manner than on outcome-oriented impact (summative); and (c) topics related to scientific impacts were widely discussed even though the impact section was to be used for evaluating economic and societal impacts. We conclude that for *ex ante* impact assessment to be effective, funding agencies should indicate the types of impact expected from research proposals clearly instead of a general 'wish list' and that more focus should be put on process-oriented impact than outcome-oriented impact.

Introduction
Impact is increasingly assessed in performance-based funding systems (PRFs) and funding agencies. In the context of grant evaluation, impact assessments can be implemented at various stages, including *ex ante* assessments, interim assessments, and *ex post* assessments. However, practices of *ex ante* impact assessment vary and why and how impact is evaluated is seldom stated explicitly. Many suggest that the need for evaluating impact is to account for the value for money of publicly funded research projects, while some also suggest that the requirement of impact is to nurture a research culture that engages with the society at large (Bozeman & Boardman 2009; Penfield et al. 2014; Spaapen & van Drooge 2011).

With the decrease in funding in public universities in many countries, funding agencies play an increasingly important role in allocating competitive funding, which can have effects on the creation of knowledge (Oppenheimer, et al. 2019) and the diversity of research fields (Whitley, Gläser & Laudel 2018). However, impact is an evolving and elusive concept in research evaluation and science policy. There are many definitions suggested by different funding agencies and research assessment exercises, not to mention the interpretations by different stakeholders (e.g. researchers, funding agencies, universities, reviewers, and so on). Nevertheless, if "impact" is a requirement for evaluating grant applications, it is important to be clear about what it is and how it would be assessed since the criteria and processes of impact assessments have not been standardised and are often ambiguous to both applicants and reviewers. A report on the United States National Science Foundation's Broader Impact

Initiatives (NABI 2018) states that the BI (Broader Impact) criterion is unclear and that random judgments on BI are common in the merit review process, among other issues. Studies have also shown that there are often discrepancies between how impact is defined by funding agencies and how it is understood by reviewers (de Jong et al. 2016). Furthermore, applicants and reviewers are often reluctant to address the societal impact of a proposal as they feel more competent in assessing its intellectual merits (Holbrook & Frodeman 2011).

As Brewer (2011) states, "Chance should play no role in allocating quality-related research funding" (p. 256). It is therefore pertinent to understand the current practices of *ex ante* impact assessment such that we can develop and improve evaluation criteria and procedures and, consequently, the transparency and fairness of impact assessments. Recent studies have analysed the criteria and processes of various funding agencies and funding schemes (Bozeman & Youtie 2017; Holbrook & Frodeman 2011; Langfeldt & Scordato 2015, 2016), as well as researchers' perceptions of impact assessments (Bozeman & Boardman, 2009, de Jong, Smit, & van Drooge 2016; de Jong & Muhonen 2018). Yet, we have little understanding of the actual practices of *ex ante* impact assessment in grant applications.

How do reviewers evaluate impact statements in grant applications without evidence or proof of impact? What do reviewers perceive as impact in grant applications? And do they evaluate impact statements based on *what* kinds of impact will be achieved or *how* impact will be achieved? In this paper, we conduct a content analysis of the comments on impact by peer reviewers, as well as the call document and guidelines for reviewers, of the Science Foundation Ireland Investigators Programme 2016 (hereafter "IvP2016") as an exploratory study to understand the actual practices of *ex ante* impact assessment in grant applications.

Literature review

> **Impact** (*Oxford Dictionary of English*)
> [1] the action of one object coming forcibly into contact with another
> [2] a marked effect or influence

The Impact Problematic

The impact of scientific research has been a topic of interest in bibliometrics and scientometrics for decades. Citation data is often used by research institutions and researchers to assess research performance of individuals, departments, universities, and countries, with the assumption that the more citations an article receives, the higher impact it has.[1] Since the 1990s, there has been a call to expand impact to include societal benefits of research (Bornmann & Marx 2014) although making a distinction between evaluation criteria of benefits internal to science, such as technological merit, and external to science such as social use, was proposed as early as the 1960s (Weinberg 1963). At the time of this writing, impact has been elaborated to encompass cultural, economic, environmental, health, political, societal impacts, and so on. The definitions and descriptions of impact have been mostly proposed by governments and funding agencies and the most prominent characteristic of the impact discourse is the necessity for showcasing the benefits of research beyond academia. Most notably, the definition by the Research Excellence Framework (REF) in the UK has

[1] Metrics as indicators of research impact have been extensively critiqued in the literature. See, for example, Gingras (2014), and also recent calls for responsible use of metrics in *The Metric Tide* (Wilsdon, et al., 2015) and the Leiden Manifesto (Hicks, et al., 2015).

been widely referred to and discussed: "an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia".[2] It is understood that the so-called "impact agenda" is a response to account for funding for research, as well as to advocate for continued or increased funding (Henshall 2011). In REF, for example, impact is evaluated in the performance-based funding scheme that determines the allocation of block grants to public universities in the UK, whereas in many funding agencies, impact is considered as part of the quality of grant applications, and as such, influences the distribution of grants in "worthwhile" research projects.

Some, however, have argued that the impact criterion in research evaluation can have negative consequences for research and knowledge production. The main reason is that impact as a criterion favours short-term, observable, and measurable impacts (see Greenhalgh & Fahy 2015; Martin 2011; Watermeyer 2016). These issues are particularly discernible in the social sciences and humanities because their impacts are often indirect and intangible. For instance, Muhonen et al. (2019) show that impact can be created and traced via many different pathways. Impact case studies in REF or other similar systems do not capture normal contributions of scholarship and research but emphasise 'extraordinary impacts' (Sivertsen and Meyer, 2020). These critiques have mainly been focusing on *ex post* impact, while the implications of *ex ante* impact assessment in grant applications, to the authors' knowledge, have not been sufficiently discussed.

Impact can be conceptualised as outcome-oriented (summative) and process-oriented (formative) (see European Science Foundation, 2012). Outcome-oriented impact emphasises tangible and measurable impacts where their attribution can be easily identified and traced. Hence, outcome-oriented impact is often discussed in a linear model (also called a "logic model" to show the intervention logic of a research programme) which implies causal relationships between funding (inputs), research activities (process) and impact (outputs).

Many have also discussed the differences between research outputs (e.g. publications), outcomes (e.g. new products and services, clinical trials), and impact (e.g. economic or social benefits resulted from the products/services). The Payback Framework, predominately used in health sciences research, presents a similar linear progression: the categorisation of 'paybacks' along with the complete research process includes knowledge production (e.g. journal articles, conference presentations), benefits to future research and research use (e.g. development of research skills, staff development), benefits from informing policy and product development (e.g. provision of information for political and executive decisions), health and health sector benefits (e.g., cost reduction in existing delivery), and broader economic benefits (Donovan & Hanney 2011).

Although the linear model, including the Payback Framework, can be useful as a structure to describe impact, it has been criticised for not corresponding well with reality. The dividing line between outputs, outcomes, and impact is not always clear-cut and it is not always possible to pinpoint when outputs and outcomes have transformed into impact (European Science Foundation, 2012). Furthermore, outcome-oriented impact has many unresolved tensions, including the ways by which impact can be captured, traced, and measured. The difficulties can be due to inadequate methodology and resources (Molas-Gallart & Tang 2011), but sometimes it is simply impossible to understand the pathways and the extent to

[2] We use the REF2014 definition in this paper since most studies have referred to it. We note that a new definition has been published by REF2021.

which any individual project has contributed to a certain 'impact' (Henshall 2011). Attribution, the counterfactual argument (what would have happened without the intervention), the time lag between output and impact, evidence of impact, and knowledge creep in understanding/interpreting impact are some of the most discussed challenges of impact assessment (Arnold et al. 2011; European Science Foundation 2012; Luukkonen 1998; Penfield, et al. 2014; Spaapen & van Drooge 2014).

Rather than identifying and measuring the outcomes of a research project, process-oriented impact focuses on the *processes* involved in achieving impact (Molas-Gallart & Tang 2011; Ramberg & Knall 2012); in other words, how impact can/will be achieved. In particular, the 'productive interactions' between researchers and stakeholders proposed by Spaapen and van Drooge (2011) has been widely discussed by other researchers in their studies. Their proposal of 'productive interactions' uses evaluation as a *learning tool* rather than an accounting apparatus, to understand the interactions between researchers and stakeholders so as to make recommendations for the improvement necessary for achieving impact.

In sum, depending on the purpose of the assessment, the institutional setting, and the socio-political context, impact can be interpreted as outputs, outcomes, and broader impacts of research in the evaluation process. What constitutes impact hence varies across space and time. For these and other reasons noted earlier, challenges of impact assessment persist. These challenges may be tackled by novel methodologies and infrastructures; or they may be addressed by shifting the focus from outcome-oriented impact to process-oriented impact. Since most discussions about impact are concerned with *ex post* impact assessments, this raises an interesting question: if there are so many uncertainties and inconsistencies in *ex post* impact assessments, how can we evaluate impact *ex ante* fairly and objectively in the context of grant applications?

Practices and Challenges of Impact Assessment in Funding Agencies

In a review of the methods and practices of impact assessments of four funding agencies, Langfeldt & Scordato (2015) show that impacts are assessed differently in terms of scoring and weighting: some assign a specific score/weight to impact, some consider impact in an aggregated score but no individual score/weighting is given (see also Holbrook & Frodeman 2011). In some funding schemes such as the Research Council of Norway's Researcher Projects Scheme, impact is not assessed at all. The composition of evaluators of impact is also different: most involve peers (i.e. researchers), while some also involve users from industry, public agencies, and so on (Langfeldt & Scordato 2015).

The Challenges of Impact Assessment Report published by the European Science Foundation (2012) recommends that it is more important first to identify what is to be assessed before considering how to measure impact. Understandably, impact can be defined more broadly or narrowly in different funding agencies and funding schemes. For example, the criterion of impact in Horizon 2020 is defined as 'the extent to which the outputs of the project should contribute at the European and/or international level', whereas the Knowledge-building Project for Industry of Research Council of Norway specifies impact relevant to trade industry where user participation is emphasised (Langfeldt & Scordato 2015). Nevertheless, funding agencies usually only require a description of the potential impact in grant applications, leaving applicants flexibility to describe the expected outcomes and specific plans/pathways of impact. Chubb and Watermeyer (2017) have reported that, however, the perceptions and experiences of writing pathways to impact statements are often described as

"lies, stories, disguise, hoodwink, game-playing, distorting fear, distrust, over-engineering, flower-up, bull-dust, disconnected, narrowing" and so on by applicants (p. 2365).

Ex ante impact assessment are judgments based on prediction, because reviewers evaluate the outputs and outcomes of a research study which has yet to take place, and because serendipity is an important element in realising impact (Henshall 2011; Meagher, Lyall & Nutley 2008). Studies of impact assessment in grant applications have shown that both applicants and reviewers prefer economic impact over societal impact in grant proposals. One of the reasons is that they perceive impact in public sectors is less favoured by funders compared to tangible outputs in commercial sectors (de Jong, Smit, & van Drooge 2016). Bozeman & Boardman (2009) argue that potential societal impact requires a notion of 'social good', for which scientists and reviewers of grant proposals may have less expertise and relevant experience to make claims about or judge. Furthermore, the preference for economic impact can be affected by methodological issues: while there are widely applied and effective methods for evaluating economic impact, methods for social impact assessments have limited theoretical guidance (Bozeman & Youtie 2017).

A related challenge in impact assessment is academic independence, involving the objectivity and credibility of researchers. While some funding agencies encourage and favour partnership with stakeholders in order to create impact, these types of interactions can be regarded as negative since researchers perceive decreased academic independence with increased partnership (de Jong & Muhonen 2018). Similar issues have been discussed in scientific assessment in environmental policy, where many scientists are aware of the potential resistance to their findings from governments, industry, and broader society. These external forces not only shape scientific assessment, but also affect the objectivity of their work and the credibility of the researchers (Oppenheimer, et al. 2019). Bozeman and Boardman (2009) have suggested that scientists perceive 'facts' as epistemic objects that are amenable to peer evaluation, whereas 'values' as subjective expressions of personal preference.

While the relationship between the impact requirement and academic independence demands further studies, it is evident that institutional support plays an important role in achieving impact. Bozeman & Youtie (2017) maintain that the institutional driving force behind funding programmes is of 'more than incidental interest to evaluators of social impacts'. de Jong & Muhonen (2018) suggest the 'societal impact capacity' of academic researchers and their research environment are critical for producing impact. Studies have shown that policies and institution-building play significant roles in societal impact and innovation (Bozeman & Boardman 2009; Hintz 2019; Roberts 2009). In other words, scientific research cannot be expected to achieve impact 'naturally' without institutional support.

While there are many discussions about *ex post* impact assessment, particularly in the context of the UK REF, studies about *ex ante* impact assessment have been rare, possibly due to the limited availability of data such as reports of peer review. Yet, impact has become one of the required components in many funding schemes and its assessment can deter the kinds of research proposals funded and hence the kinds of knowledge to be pursued and produced now and in the future. Thus, impact assessment in grant applications is an emerging area of study in research evaluation and science policy.

Background: Science Foundation Ireland and the Investigators Programme

Science Foundation Ireland (SFI) is the largest science funding agency in Ireland. It was established in 2000 as a sub-board of now dissolved Forfás[3] and became a separate legal entity in 2003. Since 2013, SFI's remit has been widened to include both oriented basic and applied research in STEM fields (Science, Technology, Engineering and Mathematics), with the goal of driving Ireland a global leader in both scientific research excellence and impact in strategic areas that concern the development and competitiveness of industry and enterprise (SFI 2012).

SFI describes impact as 'the demonstrable contribution that excellent research makes to society and the economy', including economic, societal, international engagement, policy and public service, health and wellbeing, environmental, professional services, and human capacity impacts (SFI 2017a). Funding awardees are required to report their impact in the annual report with a list of 11 categories of impact statements where at least one and up to the most relevant five statements must be made, including, but not limited to, knowledge production and transfer, public engagement, and commercialisation activities (2017b).

The Investigators Programme (IvP)[4] of SFI was one of its largest-scale awards to individual scientists. IvP supported the development of world-class research capability and human capital in the 14 priority research areas identified by the National Research Prioritisation (NRP) Steering Group (O'Hara et al. 2011) and deemed likely to yield greatest economic and societal impact in Ireland (SFI 2017c). The specific objectives of IvP2016 are stated as follows:

- To support excellent scientific research that has potential economic and societal impact aligned to Innovation 2020 enterprise themes
- To build capacity, expertise and relationships that will allow researchers based in Ireland to lead consortia and to win further support through various non-Exchequer funding schemes, such as Horizon 2020
- To support relevant collaborations and partnerships between academia and industry
- To maintain Ireland's top-20 position in international bibliometric rankings through an increase in the number and quality of journal publications
- To allow Ireland-based researchers to win top-tier international prizes (e.g. the Nobel Prize, the European Science Prize, the Lasker Award, etc.)
- To facilitate partnerships with other agencies
- To support researchers returning to active academic research after a prolonged absence through the Investigator Career Advancement (ICA) component of the call

The review process of IvP varied slightly from 2012 to 2016. IvP2016, examined in this study, awarded 32 out of 186 proposals, with the total funding amount over 34 million euros. The IvP2016 awards were made after postal and sitting panel reviews. The two stages of review were carried out by international peer reviewers. At the postal review stage, all eligible proposals were reviewed by 3-5 individual reviewers independent from each other. At the sitting panel review stage, each proposal was assigned three panel reviewers who jointly made funding recommendations to SFI. The number of panel reviewers, the number of

[3] Forfás was the national policy advisory board for enterprise, trade, science, technology and innovation in Ireland.

[4] In 2019, the Investigators Programme has been replaced by the Frontiers for the Future Programme.

proposals under panel review and the percentage of the proposals recommended for funding were consistent amongst the four disciplinary panels:

- Panel 1 – Communications, Engineering, Computer Science, Mathematics, and Geoscience
- Panel 2 – Materials and Chemistry
- Panel 3 – Life Sciences: Human Health
- Panel 4 – Life Sciences: Microbiology, Ecology, Biomarkers of Disease

The peer review criteria of IvP were the same for both review stages: the *quality, significance, and relevance* of a) the proposed investigators; b) the proposed research; and c) the impact statement. The same rating scales and weightings were assigned to the three sections. SFI's award decisions mostly follow the recommendations from the panels (approximately 91% of cases).

In the call document and the evaluation guidelines, there is a section dedicated to explaining economic and societal impact and the requirements of the impact statement. It is clearly stated that SFI recognises some research projects may have immediate impact, while others may take longer to achieve impact. The applicants were asked to address the following questions in the impact statement: Who will benefit from this research? What plans will be put in place to increase the chances of economic and societal impacts from the proposed research? And, over what timeframe might the benefits from the research be realised? The reviewers were asked to assess the potential impact of the proposed research and the likelihood of the delivery of impact.

In the guidelines for reviewers, SFI lists common characteristics of high quality and poor impact statements. They also include information about national priorities and activities, as well as the involvement of beneficiaries and end users. There is no specific structure or checklist for the reviewers to comment on the impact statements and there is no rubric to allocate points for specific aspects and characteristics of impact.

Data and Method

We obtained a corpus of 261 IvP2016 postal reviews from SFI. These reviews were partially redacted by SFI to ensure the anonymity of applicants and reviewers. Each review includes the following sections: comments on the applicant(s), comments on research programme, comments on potential for impact, comments on ethical issues, comments on budget, team and award duration, and conclusion. The reviews also include the application status of the proposal reviewed ('awarded', 'reserve list–declined', and 'declined').

Using a bottom-up approach, we extracted specific terms the reviewers used to describe and discuss impact. The extracted terms included, for example, training of scientists, licensing and so on. However, broad and abstract descriptions such as 'economic and societal impact' were excluded. Since the comments were partially redacted, we were not able to identify impact on specific research fields, industries, innovation, or products.

The codebook was developed by analysing a stratified sample from equal numbers of awarded and declined proposals, amounting to 22.5% of the corpus which was sufficient to reach code saturation. When developing the codebook, it became clear that the comments on

impact were clustered around topics related to science, education, and commercial potentials. On the basis that both applicants and reviewers were informed about the types of impact that can be discussed and evaluated in a proposal, we decided to use the types of impact indicated in the call document to develop our coding scheme. We added two extra types and split a mixed one into two types based on the initial analysis of the topics. The final coding scheme consists of the following 11 types of impacts:

- Research, science and publications*
- Powering an innovative and enterprising economy
- Creating high-value jobs
- Industry collaboration/links*
- Attracting, developing and nurturing businesses#
- Attracting, developing and nurturing scientists and talented people#
- Increasing the effectiveness of public services and policy
- Enhancing quality of life, health and creative output
- Developing the country's international reputation
- Educating and training the population
- Solving major national and global problems and challenges

*Types of impact added
"Attracting, developing and nurturing businesses, scientists and talented people" split into two categories

The extracted terms and descriptions from each review were coded under the eleven types of impact. The coding result of all the reviews was then analysed based on the four panels and application status ('awarded', 'reserve list–declined', and 'declined'). All coding was done manually. The extraction of terms and descriptions and their categorisation into impact types was initially done by one researcher who read and coded the impact section of the reviews, then crossed-checked by a second researcher who read the full-text of the reviews. The results were largely consistent between the two researchers, and disagreements were resolved by reading and discussing the original comments.

Finally, we calculated and visualised: (a) the likelihood that each type of impact appears in reviews on impact statements, i.e. relative frequency of each code; (b) for each code, the probability for proposals in which the code appears to be eventually awarded; and (c) for each pair of codes, the probability that they both appear in a review (code correlation) and the award probability of proposals in which the pair of codes appear (probability of award for each pair of codes).

Findings and Analysis

Our analysis shows that reviewers were looking for relatively short-term impacts that could be predicted or foreseen as a result/consequence of the proposed research project. Reviewers had more reservations about evaluating longer-term impact which they expressed by using conditional statements such as 'In case of a successful project…', 'If the given estimates of this work…', 'If the project achieves its stated aims…' This is because the likelihood of long-term impacts is dependent upon the completion and success of the proposed research, which itself is uncertain. One reviewer stated explicitly: 'The time frame is admitted to be hard to quantify by the proposers, though they do cite realistic short term and long term objectives. The short-term benefits seem likely to be realized, while the longer-term benefits

appear more dependent on project success'. Reviewers did not differentiate outputs, outcomes, and broader impacts as in the linear model (see Penfield, et al., 2014). For example, publications and patents (outputs), new businesses and products (outcomes) and quality of life (impact) were all interpreted as 'impact' in the comments.

Table 1 presents some examples of terms and descriptions extracted from reviews for each impact type. The comments are similar in formats and vocabularies. Some of the most commonly used terms are 'education and training', 'science', 'industrial partners', 'IP', 'commercialisation', and 'exploitation'. Most comments did not provide explanations or elaborations as to whether certain types of impact are desirable or achievable; rather, they were simple, direct statements to confirm or reject the impact statement. The feasibility of achieving impacts within a certain timeframe was often commented upon, but again, explanations or elaborations are often lacking. A small number of comments were very brief and cursory. For example, a reviewer would comment that "I think this project will/will not have economic and societal impact" but did not provide reasons for their comments.

Impact Types	**Examples of Terms and Descriptions in Reviews**
Research, science and publications	'increase of scientific and technical know how', 'disseminate their research outputs', 'novel method', 'generate novel data', 'publications', 'science'
Powering an innovative and enterprising economy	'exploitation', 'intellectual property', 'translation', 'generating a commercial income stream'
Creating high-value jobs	'job creation in Ireland', 'occupational possibilities'
Industry collaboration/links	'forming new industry partnerships', 'collaboration with industry', 'substantial involvement of commercial entities'
Attracting, developing and nurturing businesses	'spin-off companies', 'new industry or business', 'creation of industry advisory board'
Attracting, developing and nurturing scientists and talented people	'education', 'training of PhD students and postdocs', 'training next generation of scientists'
Increasing the effectiveness of public services and policy	'policies and public service'
Enhancing quality of life, health and creative output	'cultural life', 'improving the quality of life in Ireland's citizens'
Developing the country's international reputation	'international visibility', 'enhancing Ireland's international reputation in research'
Educating and training the population	'public education and engagement', 'engaging with society', 'outreach'
Solving major national and global problems and challenges	'priority area'

Table 1 Examples of terms and descriptions in reviews for each impact type

The reviewers' comments on impact align with many of the objectives of IvP2016 (see Figure 1). While impact types, including 'Increasing the effectiveness of public services and policy', 'Enhancing quality of life, health and creative output', 'Educating and training the population' and 'Solving major national and global problems', are expected impacts according to the call document and the guidelines for reviewers, they do not correspond with the objectives of IvP2016 and they were also less discussed by reviewers. We postulate that the reviewers did not perceive these types of impact as important as the shorter-term, more tangible impacts in the commercial sector, and/or that the impact statements did not mention

them as the applicants perceived impacts that are more aligned with the objectives of the programme would be more favourable. However, further studies are needed to understand the reasons behind the lack of mentions and discussions.

Figure 1 Alignment between IvP objectives and types of impact

There was a strong emphasis on economic development and innovation in the reviewers' comments. The category 'Powering an innovative and enterprising economy' was discussed in terms of commercialisation, exploitation, and intellectual property, whereas the category 'Attracting, developing and nurturing businesses' was commented upon based on the possibilities of spin-off companies, licensing, and other business opportunities. Although the two categories are both concerned with economic prospects, they are different in nature: the former describes the impact on innovation with more abstract and open-ended terms, whereas the latter uses more concrete plans for developing specific businesses and/or products. Both are relevant to a third category: industry collaboration/links (see also Figure 2). They were often mentioned together in the comments as researchers who already have contacts and/or collaborative experiences with industry partners were perceived as having more potential for creating economic impacts. However, few comments mentioned creating high-value jobs or equivalent.

Interestingly, few reviews commented on 'Solving major national and global problems and challenges'. One reason could be that this is such a strong claim that neither applicants nor reviewers could state or comment confidently. It could also be due to the redaction of the comments in the reviews since the research fields, industries, company names, products, and hence specific national and/or global problems could not be identified. Relatedly, few reviews discussed longer-term impacts such as 'Enhancing life, health, and creative outputs' and 'Increasing the effectiveness of public services and policy', likely because they are more uncertain and difficult to predict *ex ante.* The reasons include the diverse pathways of knowledge dissemination and the broad scope of research beneficiaries, not to mention the time needed to produce 'long-term impact' is difficult to predict. Put another way, these types

of impact often cannot be planned *ex ante*—as they are the ultimate impacts in a linear model, i.e. after outputs and outcomes have been produced and manifested.

Although the impact statement of the grant application was reserved for articulating economic and societal impacts according to the call document and guidelines for reviewers, most reviewers commented on scientific and research impact, for example, the potential for publications and for leveraging research funding. Their descriptions involved terms such as 'knowledge', 'dissemination', 'novel data', 'fundamental and applied facets of academic research'. Relatedly, many comments mentioned the values of education and training of the next generation of scientists, particularly at the postgraduate level. There was a strong emphasis on the contributions to knowledge and the development of science and scientists, as shown in Figure 2 by the big node size of two impact types of 'Research, science and publication' and 'Scientists and talented people'.

In addition, Figure 2 illustrates in visual form some of the relationships among impact types. The co-occurrence of impact types is represented by the edge width: there is a strong indication that the most frequently mentioned impact types are often discussed together in a review, as the most frequent impact types have strong links with each other. However, the edge shades of these stronger/wider links are darker than those narrow links, which indicates that the award probability of frequent co-occurrence codes is lower. It should be noted that comments involving 'public services and policy' or 'enhancing the quality of life', although being very few (appearing in only 2% of the reviews in our dataset), are associated with a high probability of award.

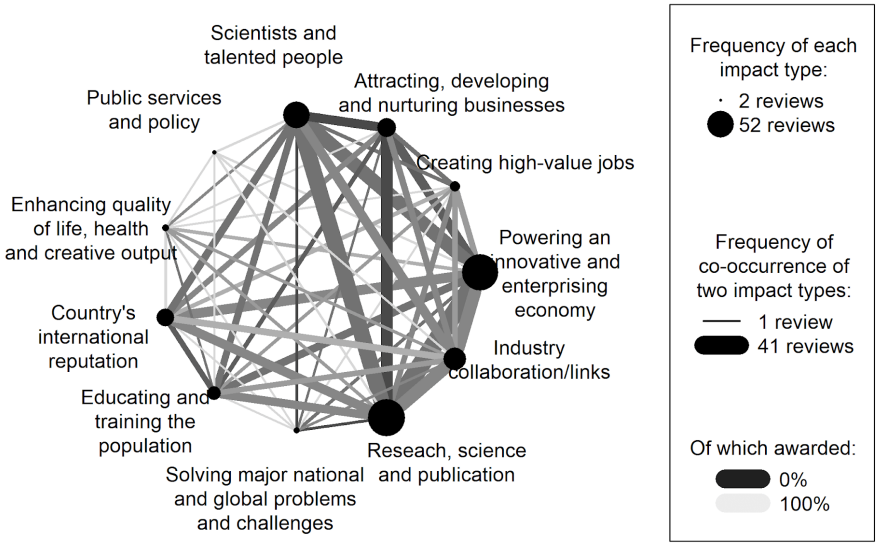

Figure 2 The co-occurrence network of impact types and the associated probability of award. Node size represents the frequency of each impact type in the reviews. Edge width represents the frequency of co-occurrence of any two impact types (i.e. are both discussed in a review). Edge shade represents the percentage of proposals that were awarded, out of the set of proposals with co-occurrence of any two impact types(where brighter shades indicate higher proportion of awarded proposals).

When reviewers commented on the perceived impacts based on the merits of the research programmes, they particularly favoured existing and future industrial partnerships. One reason could be that the SFI call document and evaluation guidelines present a strong emphasis in areas that concern the development and competitiveness of industry and enterprise in Ireland, which consequently steers more focus on industrial partnerships by both applicants and reviewers. Although reviewers commented on the timeframe of the plans or pathways to achieve impacts, they did not question the evidence required for demonstrating impacts after the application has been awarded, nor did they evaluate components of the plans. There were no noteworthy differences between the four different panels (see Figure 3) [5] as to what impact means and how it is assessed, although this could also be due to the data redaction.

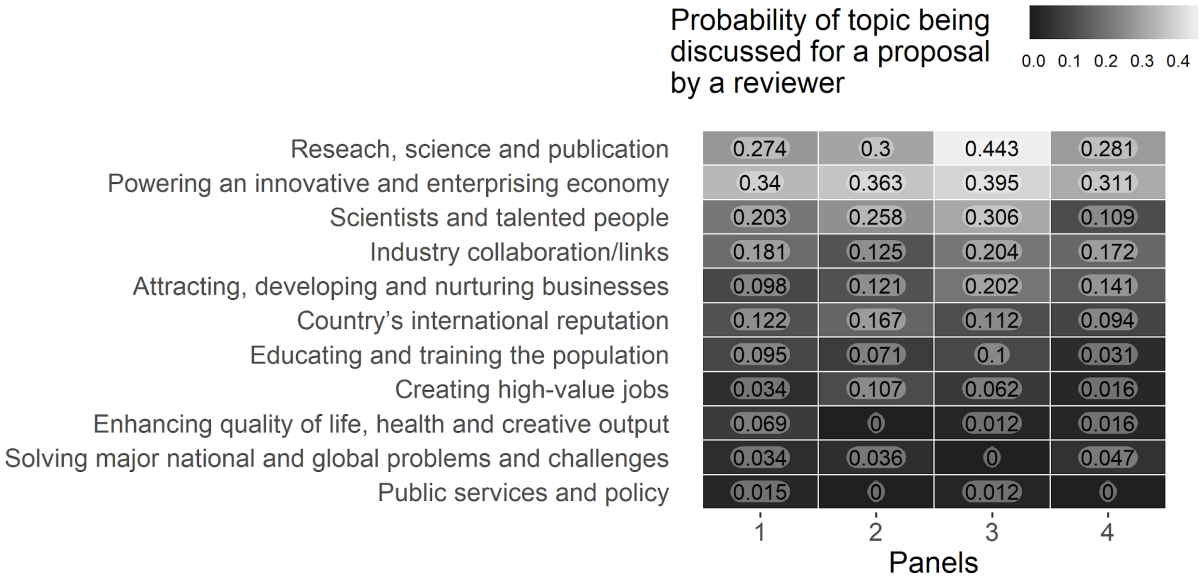

Figure 3 The relative frequency and award probability of impact types of four review panels

Overall, the reviewers addressed both outcome-oriented impact and process-oriented impact, with more elaborate comments on the latter. Process-oriented impact, in the form of 'productive interactions' with postgraduate students, industry partners, scientists and other research institutions worldwide were discussed in more detail. This is likely because claims

[5] Our comment that there are no noteworthy differences between panels is based on the observation that between-panels differences are really small (entirely negligible in most cases), whereas differences between topics are much larger. To clarify with an example from Figure 3: here we see that the differences in topic selection between the panels are generally very small (in most cases, < 0.1). There are only few exceptions, most notably the differences between panels 3 and 4 with regard to the three topics at the top, see e.g. "Research, science and publication", which is discussed 44% of the times by reviewers in panel 3, but only 28% of the times by reviewers in panel 4. Considering that our data only contains 261 reviews spread across four panels, and considering that only a fraction of these reviews cover the topic "Research, science and publication", it is clear that this difference between panels 3 and 4 is based on a very limited number of reviews.

about outcome-oriented impact require more precise predictions and evidence, as such, they were commented upon in a more ambiguous way.

Discussion

What kinds of impact do reviewers comment on in grant applications? What criteria are used to evaluate statements of potential impact? Although we cannot generalise the findings to all types of *ex ante* impact assessment, this study highlights some issues important for future design of impact assessment in grants evaluation and the development of methodology. Furthermore, it also raises theoretical questions with regard to accountability and responsibility in research culture.

The study shows that impact in SFI's Investigators Programme, in both the call document and the reviewers' comments on impact, encompasses but does not differentiate outputs, outcomes, and broader impacts. While it is understood that there is no clear-cut difference between 'impacts' in different stages, it is evident that SFI made no attempts to differentiate tangible/intangible, measurable/immeasurable, short-/long-term or outcome-/process-oriented impacts. The inclusive approach is similar to the Broader Impacts Initiative of the National Science Foundation in the U.S. As such, some of the same issues of impact assessment emerge in this study: the criterion of impact is unclear and vague and can lead to random or cursory reviews (NABI 2018). For example, even though the guidelines for reviewers and applicants did not indicate a preference for long-term impacts or short-term outcomes, the reviewers tended to comment favourably on short-term, tangible outcomes. In contrast, many reviewers had doubts about the potential impacts of basic research—not because of their scientific merits, but the possibility of impacts to be achieved within the timeframe of the grant period. One reviewer commented (emphasis added):

> Since long it is the ultimate goal of many such research efforts to bring *basic research* on [redacted] into the realm of applications but [redacted] remains a challenging step. Nevertheless, history suggests that one can expect successful conclusions for subsets of the proposed [redacted]. However, it is challenging to judge at this point what the likelihood, scale and value of the societal and/or economic effect the proposed research will have on Ireland.

The reviews also show a focus on economic impacts, particularly in the commercial sectors, mirroring similar results in previous studies (Bozeman & Boardman 2009; Bozeman & Youtie 2018; de Jong, Smit & van Drooge 2016). One of the possible reasons is that societal impacts are less concrete, more uncertain and difficult to predict than outcomes in the commercial sectors, not only for reviewers but probably also for funding agencies and applicants to describe their expected impacts. In other words, outcomes such as intellectual property, patents, and commercialisation activities are more observable and tangible, and are hence easier to describe and evaluate. In contrast, societal impacts can be very broad and intangible; their evaluation would be based on normative judgments about 'social good' or 'common good' which can be more subjective and arbitrary.

The preference for short-term economic impacts in the study presents an important issue in *ex ante* impact assessments: Funding agencies should clearly indicate what types of impact are *actually* desired for a funding programme. In other words, the criteria as to what constitutes impact in a funding programme should be clear to both reviewers and applicants for the assessment to be fair. Statements about potential impact are essentially predictions; in contrast to *ex post* impact assessment, very little objective evidence can be provided in

impact statements of research proposals. Therefore, it would be useful if a typology of impact, including the types of expected impact, timeframe, beneficiaries, and other evaluative entities are included in the guidelines for preparing and evaluating impact statements.

In addition to the clarification of impact types in terms of outputs, outcomes and broader impacts, it is also important to consider how outcome-oriented (summative) impact or process-oriented impact (formative) can be evaluated and aligned with programme objectives. As many have discussed, the criterion of impact, or the 'impact agenda', has been a response to, inasmuch as a requirement in, the new public management regime (e.g. Bozeman & Boardman 2009; Penfield, et al. 2014; Watermeyer 2016). Hence, outcomes and impacts are highlighted as evidence to account for value-for-money. There are, however, many methodological challenges due to the fluid nature of impact: in *ex post* assessments, there are questions concerning tracing and measuring direct and indirect impacts, whereas in *ex ante* impact assessments, potential impacts can only be described and justified but cannot be verified or proved. As such, it is more challenging to evaluate impact statements in an objective manner *ex ante* because the judgments are based on the reviewers' knowledge and experiences rather than evidence. The focus on outcome-oriented impacts in *ex ante* impact assessments can probably explain the preference for shorter-term, tangible economic impacts compared to societal impacts.

If we consider process-oriented impacts in terms of productive interactions (Spaapen & van Drooge 2011), there is much room to improve the fairness and objectivity of *ex ante* impact assessments. Impact statements can be used to articulate the plans and pathways the researchers intend to interact directly and indirectly with the stakeholders; they can also provide context to show how their research can fulfil the needs of the society and economy. The evaluation of impact can then emphasise *processes* to achieve impact. This move could resolve some of the issues related to the difficulties of prediction signified by the various issues including attribution and the counterfactual argument in *ex post* impact assessment. Shifting the focus to processes-oriented impacts can also be considered as a shift to using evaluation as a learning tool (Spaapen & van Drooge 2011) for nurturing a research culture that is responsive to and responsible for society. Practically, the shift can also motivate and guide researchers to consider and plan engagement and interaction activities that can be evaluated in a less ambiguous and abstract manner. That said, *ex ante* impact assessments can never be entirely objective, for reviewers' comments on economic and societal impacts, broadly construed, are largely based on normative understanding of social values which are always contestable and debatable.

The alignment between the types of impacts and programme objectives in this study, as shown in Figure 1, can be another explanation for the lack of comments on societal impacts described in a more abstract manner (e.g. 'enhancing quality of life, health and creative output'). The alignment corresponds to the high number of comments on scientific excellence and the training of scientists. Hence, it can be postulated that, first, reviewers made comments and decisions based on the objectives of funding programme, and less on the wide range of impacts in the call documents and the guidelines for reviewers; and, second, that they commented on the kinds of impact perceived as appropriate for the funding programme and desired by the funding agency. In other words, their comments reflect the specific needs of the funding programme, not necessarily whether the research proposals would lead to the highest and widest impacts. If the impact assessment was to select the best proposals of the highest potential of long-term impacts, the assessments did not fulfil the objective of the evaluation. The future design of impact assessment should consider the purpose of the impact

assessment and delineate the kinds of impact appropriate for the funding programme as a wish list of impacts may not be the most useful for reviewers.

While it is important to improve the criteria and process of *ex ante* impact assessment, questions concerning academic freedom and knowledge production pertaining to the notion of impact should also be considered by researchers, funding agencies, and policy-makers. For instance, the recent announcement of the removal of the Pathway to Impact statement in UKRI (2020) proposals 2020 is not without its controversy and generated heated debate on social media and in the higher education press. Some research policy makers and academics welcomed this move as a nod to reducing bureaucracy, paperwork, and thus promoting more investment in doing the actual research (or at least proposing it) while others lamented it as a de-emphasis on engaged scholarship.

Hence, it is important to examine whether impact is a goal in itself for the sake of auditability, or a process in which researchers advance knowledge and benefit humanity sustainably. It is essential to consider how to motivate researchers to reflect on impact more carefully and seriously in their work—when scientists perceive science policy focuses narrowly on commercial impacts (de Jong et al. 2016), when there is low reward for impact compared to publications (de Jong & Muhonen 2018), and when academic independence is put into question as commercial and political interests are prioritised (de Jong & Muhonen 2018; Oppenheimer et al. 2019). Lam (2011) has found that the main reason for scientists to pursue commercial activities is not because of financial rewards, but rather, reputational and intrinsic reasons. If 'impact' is to be encouraged and evaluated, there have to be structural changes in the landscape of scholarly communication and the academic reward system.

Conclusion

Impact assessment is not unique to research evaluation and science policy. It has been widely used in the policy domain to assess the impact of policy intervention. The European Commission, for example, provides guidelines for impact assessments for various programmes. There are, however, different interpretations of impact assessment, and so different approaches to assess impact (see, for example, White 2010) in different domains of public policy. Impact assessments in the context of performance-based funding systems and grant evaluations in funding agencies, however, are mainly to account for the value for money of publicly-funded research. Previous studies have shown that impact of scientific research is challenging to describe and assess *ex post* because the pathways can be diverse, indirect, and undetectable. It is not surprising that it is even more difficult to ascertain impact *ex ante*. While there is increasing pressure to evaluate impact in grant applications, there is little guidance as to what types of impact can be assessed prospectively, and how they can/should be evaluated.

Using content analysis of peer reviewers' comments on impact in SFI IvP 2016, the findings of this study are aligned with those in previous studies that there is a preference for economic impacts related to the commercial sectors and that reviewers' comments were abstract on outcome-oriented impact compared to the more elaborate discussion on process-oriented impact. We also find that reviewers did not distinguish outputs, outcomes, and broader impacts in their reviews as in the linear model of impact (Penfield et al. 2014). Hence, we argue that a wish list of potential impacts does not necessarily induce imagination and innovation, but rather, they can lead to uncertainties and randomness in assessing and scoring

impact statements (NABI 2018) and ambiguous and abstract reviews (Hellström & Hellström 2017).

We conclude that it is essential to clarify the purposes and criteria for evaluating potential impact, especially when funding allocations can have significant implications for knowledge production and solving important problems such as poverty and the climate crisis. Process-oriented impacts are more appropriate in *ex ante* impact assessment for they reduce uncertainties and randomness in evaluation, on the one hand, and they prompt plans and activities to achieve impacts, on the other. In sum, *ex ante* impact assessment plays an important role in the allocation of funding, further analyses of peer reviews of different types of funding programmes and the decision-making process of peer reviewers are needed to improve the fairness, transparency, and efficiency of impact assessments.

References

Arnold et al. (2011) Understanding the Long Term Impact of the Framework Programme, European Policy Evaluation Consortium (EPEC) Report. Brussel: European Policy Evaluation Consortium. Accessed October 9, 2019, from https://ec.europa.eu/research/evaluations/pdf/archive/other_reports_studies_and_documents/long_term_impact_of_the_fp.pdf

Bornmann, L., and Marx, W. (2014) 'How Should the Societal Impact of Research be Generated and Measured? A Proposal for a Simple and Practicable Approach to Allow Interdisciplinary Comparisons', *Scientometrics*, 98/1: 211–219. https://doi.org/10.1007/s11192-013-1020-x

Bozeman, B., and Boardman, C. (2009) 'Broad Impacts and Narrow Perspectives: Passing the Buck on Science and Social Impacts', *Social Epistemology*, 23/3-4: 183–198. https://doi.org/10.1080/02691720903364019

Bozeman, B., and Youtie, J. (2017) 'Socio-economic impacts and public value of government-funded research: Lessons from four US National Science Foundation initiatives', *Research Policy*, 46: 1387-1398.

Brewer, J. D. (2011) 'The impact of impact', *Research Evaluation*, 20/3: 255-256.

Chubb, J., and Watermeyer, R. (2017) 'Artifice or Integrity in the Marketization of Research Impact? Investigating the Moral Economy of (Pathways to) Impact Statements within Research Funding Proposals in the UK and Australia', *Studies in Higher Education*, 42/12: 2360-2372.

de Jong, S., Smit, J., and van Drooge, L. (2016) '"Scientists" Response to Societal Impact Policies: A policy Paradox', *Science and Public Policy*, 43/1: 102-114.

de Jong, S. P. L., and Muhonen, R. (2020) 'Who Benefits from *Ex Ante* Societal Impact Evaluation in the European Funding Arena? A Cross-Country Comparison of Societal Impact Capacity in the Social Sciences and Humanities', *Research Evaluation*, 29/1: 22-33. https://doi.org/10.1093/reseval/rvy036

Donovan, C., and Hanney, S. (2011) 'The "Payback Framework" Explained', *Research Evaluation*, 20/3: 181-183.

European Commission. (2009) 'Impact Assessment Guidelines'. Accessed October 9, 2019, from https://ec.europa.eu/smart-regulation/impact/commission_guidelines/docs/iag_2009_en.pdf

European Science Foundation. (2012) 'The Challenges of Impact Assessment.' Accessed October 9, 2019, from http://archives.esf.org/coordinating-research/mo-fora/evaluation-of-publicly-funded-research.html

Gingras, Yves. (2014) *Bibliometrics and Research Evaluation: Uses and Abuses*. Cambridge, Massachusetts: The MIT Press.
Greenhalgh, T., and Fahy, N. (2015) 'Research Impact in the Community-Based Health Sciences: An Analysis of 162 Case Studies from the 2014 UK Research Excellence Framework', *BMC Medicine*, 13/1: 232.
Hellström, T., and Hellström, C. (2017) 'Achieving Impact: Impact Evaluations and Narrative Simplication', *Prometheus*, 35/3: 215-230.
Henshall, C. (2011) 'The impact of Payback Research: Developing and Using Evidence in Policy', *Research Evaluation*, 20/3: 257-258.
Hicks, D., Wouters, P., Woltman, L., de Rijcke, S. and Rafols, I. (2015) 'The Leiden Manifesto for Research Metrics', *Nature*, 520: 429-431.
Hintz, E. S. (2019) 'Failed Inventor Initiatives, from the Franklin Institute to Quirky'. In: M. Wisnioski, E. S. Hintz, & M. S. Kleine (Eds.), *Does America Need More Innovators?* pp. 165-190. Cambridge, Massachusetts, The MIT Press.
Holbrook, J. B., and Frodeman, R. (2011) 'Peer Review and the *Ex Ante* Assessment of Societal Impacts', *Research Evaluation*, 20/3: 239-246. DOI: 10.3152/095820211X12941371876788
Lam, A. (2011) 'What Motivates Academic Scientists to Engage in Research Commercialization: 'Gold', 'Ribbon' or 'Puzzle'?', *Research Policy*, 40/10: 1354–1368. https://doi.org/10.1016/J.RESPOL.2011.09.002
Langfeldt, L., and Scordato, L. (2015) 'Assessing the Broader Impacts of Research: A Review of Methods and Practices', *NIFU Working Paper 8/2015*. Oslo: Nordic Institute for Studies in Innovation, Research and Education.
Langfeldt, L., and Scordato, L. (2016) 'Efficiency and Flexibility in Research Funding: A Comparative Study of Funding Instruments and Review Criteria', *NIFU Report 2016.9.* Oslo: Nordic Institute for Studies in Innovation, Research and Education.
Luukkonen, T. (1998) 'The Difficulties in Assessing the Impact of EU Framework Programmes', *Research Policy*, 27/6: 599–610. https://doi.org/10.1016/S0048-7333(98)00058-4
Martin, B. R. (2011) 'The Research Excellence Framework and the 'Impact Agenda': Are We Creating a Frankenstein Monster?', *Research Evaluation*, 20/3: 247-254.
Meagher, L., Lyall, C., and Nutley, S. (2008) 'Flows of Knowledge, Expertise and Influence: A Method for Assessing Policy and Practice Impacts from Social Science Research', *Research Evaluation*, *17*/3: 163–173. https://doi.org/10.3152/095820208X331720
Molas-Gallart, J. and Tang, P. (2011) 'Tracing 'Productive Interactions' to Identify Social Impacts: An Example from the Social Sciences', *Research Evaluation*, 20/3: 219-226. DOI: 10.3152/095820211X12941371876706
Muhonen, R., Benneworth, P., and Olmos-Peñuela, J. (2020) 'From Productive Interactions to Impact Pathways: Understanding the Key Dimensions in Developing SSH Research Societal Impact', *Research Evaluation*, 29/1: 34-47. https://doi.org/10.1093/reseval/rvz003
NABI (National Alliance for Broader Impacts). (2018) The Current State of Broader Impacts: Advancing Science and Benefiting Society. Accessed October 9, 2019, from https://broaderimpacts.net/wp-content/uploads/2018/01/nabi-current-state-of-bi-011118.pdf
O'Hara, J. et al. (2011). Report of the Research Prioritisation Steering Group. Accessed October 9, 2019, from https://dbei.gov.ie/en/Publications/Publication-files/Research-Prioritisation.pdf
Oppenheimer, Michael, et al. (2019) *Discerning Experts: The Practices of Scientific Assessment for Environmental Policy*. Chicago: University of Chicago Press.

Penfield, T., et al. (2014) 'Assessment, Evaluations, and Definitions of Research Impact: A Review', *Research Evaluation*, 23/1: 21-32.
Power, Michael. (1997) *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.
Ramberg, I. and Knall, M. (2012) *Challenges measuring effects of research and innovation policy interventions in ex-post impact evaluations. A synthesis report.* 4 March 2012. Oslo: NIFU.
Roberts, M. R. (2009) 'Realizing Societal Benefit from Academic Research: Analysis of the National Science Foundation's Broader Impacts Criterion', *Social Epistemology*, 23/3-4: 199–219. https://doi.org/10.1080/02691720903364035
SFI (Science Foundation Ireland). (2012). Science Foundation Ireland Agenda 2020: Excellence and Impact. Accessed October 9, 2019, from https://www.sfi.ie/funding/sfi-policies-and-guidance/research-impact/AGENDA-2020.pdf
SFI (Science Foundation Ireland). (2017a) Research Impact. Accessed October 9, 2019, from https://www.sfi.ie/funding/award-management/research-impact/
SFI (Science Foundation Ireland). (2017b) Reporting Procedures. Accessed October 9, 2019, from https://www.sfi.ie/funding/award-management/reporting-procedures/
SFI (Science Foundation Ireland). (2017c). SFI Investigators Programme. Accessed October 9, 2019, from https://www.sfi.ie/funding/funding-calls/sfi-investigators-programme/index.xml
Sivertsen, G., & Meijer, I. (2020) 'Normal versus Extraordinary Societal Impact: How to Understand, Evaluate, and Improve Research Activities in Their Relations to Society?', *Research Evaluation*, 29/1: 66–70. https://doi.org/10.1093/reseval/rvz032
Spaapen, J., and van Drooge, L. (2011) 'Introducing 'Productive Interactions' in Social Impact Assessment', *Research Evaluation*, 20/3: 211-218. DOI: 10.3152/095820211X12941371876742
UKRI (2020). Pathways to Impact: Impact core to the UK Research and Innovation application process. Accessed on January 28, 2020, at https://www.ukri.org/news/pathways-to-impact-impact-core-to-the-uk-research-and-innovation-application-process/
Watermeyer, R. (2016) 'Impact in the REF: Issues and Obstacles', *Studies in Higher Education*, 41/2: 199–214. https://doi.org/10.1080/03075079.2014.915303
Weinberg, A. M. (1963/2000) 'Criteria for scientific choice', Minerva 1/2: 159–171.
Wilsdon, J., et al. (2015) *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management.* DOI: 10.13140/RG.2.1.4929.1363
White, H. (2010) 'A Contribution to Current Debates in Impact Evaluation', *Evaluation*, 16/2: 153–164. https://doi.org/10.1177/1356389010361562
Whitley, R., Gläser, J., and Laudel, G. (2018) 'The Impact of Changing Funding and Authority Relationships on Scientific Innovations', *Minerva*, 56: 109-134.