Research Evaluation, 30(2), 2021, 169–178
doi: 10.1093/reseval/rvaa031
Advance Access Publication Date: 25 December 2020
Article

Gathering evidence of impact from research support services: Examining impact in the context of the Centre for Environmental Data Analysis

Poppy Townsend[1] and Clare Wilkinson [2,*]

[1]Centre for Environmental Data Analysis, RAL Space, STFC Rutherford Appleton Laboratory, Harwell Campus, Didcot OX11 0QX, UK and [2]Science Communication Unit, Department of Applied Sciences, Coldharbour Lane, UWE Bristol, Bristol BS16 1QY, UK

*Corresponding author. Email: Clare.Wilkinson@uwe.ac.uk.

Abstract

The Centre for Environmental Data Analysis (CEDA) is a provider of two major services to the environmental science community; JASMIN and the CEDA Archive. CEDA is frequently required to evidence the impact it has on researchers and wider society. However, this is challenging as there are currently no formal or standard processes for collecting impact information. To understand how CEDA could collect impact information, and to allow its users to shape this monitoring, over 500 users provided their opinions, preferences and suggestions as to how to share impact, via six focus groups and an online survey. The results suggest that whilst there was a high degree of willingness to provide impact information to CEDA there remains confusion around what 'impact' is. Users are keen to share impact in ways which utilize existing processes, and at times which make sense to both the research and the impact, whilst also understanding the need and purpose for sharing that information.

Key words: impact; data; archives; environmental research; data centre; big data

1. Introduction

Funding organizations in the UK increasingly require evidence as to how the research or facilities that they support provide benefits to wider society, as illustrated by various grant application guidelines from UK Research Councils (e.g., ESRC 2017; NERC 2018; STFC 2018) and via the Research Excellence Framework (REF) utilized by UK Universities (REF2021 2018). This desire to map and evidence the longer-term influence of research is commonly known as measuring 'impact'.

One facility that requires evidence of impact is CEDA (www.ceda.ac.uk). Funded predominantly by UKRI, CEDA provides various services to help support the environmental science community (CEDA 2018a). These services include firstly, the CEDA Archive (www.archive.ceda.ac.uk); a collection of UK environmental data which is archived for long-term use. Types of data include, but are not limited to; satellite images, climate models, and atmospheric chemistry measurements (CEDA 2018b). Secondly, JASMIN (www.jasmin.ac.uk); a data-intensive supercomputer co-located with the archive, allowing users to bring their data for analysis and also utilize archived data (JASMIN 2018). Despite the use of these services by large numbers of environmental scientists, with over 17,000 active users, there has been minimal dialogue between the operators of the services and users, with most exchanges typically limited to user queries asking CEDA for functional support. CEDA therefore provides a service that is well utilized, but little is known about what users then do with those resources.

Like many centrally supported services, CEDA has been increasingly encouraged to provide its funders, and ultimately the UK taxpayer, with evidence of its impact to wider society outside of academia (Lawrence and Townsend 2016). When users sign up to the CEDA Archive services they are asked for what purpose they intend to use any restricted data. However, this does not happen for datasets which are open access. For JASMIN, service users have to apply for access and are again asked about the purpose of their use.

However, there has not been extensive analysis of these data, due to limited staff resources and the sheer scale of applications received. Previously, any consideration of impact has been achieved on an adhoc basis, for example by directly asking key users what their research impacts are, but in recent years, there has been a desire to formalize this process. This has occurred in parallel to increasing requests of researchers to report on research impact through a number of other mechanisms, including research applications, REF, and databases such as ResearchFish (Reed 2016; Solans-Domènech et al. 2019).

Where CEDA differs from some of these other research impact reporting infrastructures is that its work does not directly lead to impact; instead, it enables scientists to do their impactful research, and it is therefore akin to other sorts of research support that might exist for researchers, such as libraries, museums, and archives, or other forms of technical support. Demonstrating impact is already difficult (Watermeyer 2011; Penfield et al. 2014; McCowan 2018), but evidencing impact when it is once (or more) removed, and collecting evidence from researchers who may already be feeding in evidence across multiple mechanisms, is perhaps, even more challenging.

To further explore this context, this research aimed to investigate how information on impact could be collected from CEDA users and to propose recommendations as to how an impact gathering process could be implemented.

2. The context of research impact

2.1 Definitions of research impact

Many governments now expect researchers, whose research is publicly funded, to demonstrate how their research is relevant to the world outside academia (Morgan 2014) and measuring impacts from research has become increasingly integral in grant applications and assessments of research excellence (Reed 2016). Although reasons for wanting to evidence impact are complex, Penfield et al. (2014) describe four primary purposes for assessing impact. These include the need to have an overview of higher education, for functions of accountability (to government, stakeholders, and wider public), to inform ongoing funding, and finally, to understand better ways of delivering impact (Penfield et al. 2014).

Despite this increased focus on impact, many organizations are operating with subtly different definitions of impact (Williams 2020). In the UK, the Higher Education Funding Council (HEFCE 2016) defines 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', whilst UKRI (2018a) defines it as 'the demonstrable contribution that excellent research makes to society and the economy'. Despite some variations, definitions of impact tend to have the sense of extension outside of academia in common, as well as an implicit expectation that such impacts will be of benefit.

2.2.Processes for reporting impact

The UK's main assessment of research impact, introduced in 2014 by HEFCE, is the REF (HEFCE 2016). The REF's purpose is to provide accountability and evidence of benefits for publicly funded research, to allow comparisons between higher education organizations, and to inform funding allocations (REF2021 2018). In 2011, a joint statement from HEFCE, RCUK (now UKRI) and Universities UK (UUK) (2011) summarized their commitment to the agenda; making impact an essential component in all areas of the research lifecycle. The UK is one of the few countries that has undertaken a national assessment of impact within higher education (Wilkinson 2017) but similar assessments have been piloted in other countries, for example Australia (Morgan Jones et al. 2013).

The next REF is in 2021, and there is an ongoing commitment to continue to assess institutions including on the basis of impact, despite some changes to other aspects of the process (REF Steering Group 2019). The increasing influence of research impact has led some to report that researchers now consider impact at earlier stages of their projects, seemingly due to a major shift in awareness around impact (Marcella, Lockerbie and Bloice 2015). It has also been suggested that the agenda's influence is broadening the potential for credit for activities, which may have been previously overlooked in assessments, such as public engagement (Watermeyer 2012; Williams 2020).

Impact information was also built into grant application processes for the UK Research Councils with 'Pathways to Impact' encouraging researchers to think about what can be done to ensure their research makes a difference at the point of applying for funding (UKRI 2018b), though there have recently been some changes to this approach. The Research Councils concurrently use an online system called 'Researchfish', an international system used by over 100 funders in the UK, North America, and Europe, to collect information from their grant holders regarding the outcomes from their work (UKRI 2018b).

Beyond these large-scale policy approaches, there is a lack of literature as to how other service providers (Kassab, Mutz and Daniel 2020), like CEDA, are collecting information about the impact they have on users and wider society but there is evidence some are producing impact stories for use on their websites or in reports (NCEO 2017; NEODAAS 2018). There is also evidence of non-public impact reporting for environmental organizations; such as a national capability commissioning project NERC has recently undertaken with its data centres (including CEDA) via a survey, although this is not yet published publicly (NERC 2016). As the impact agenda continues, with funders requiring this information more frequently from service providers, solutions that work over a range of sectors are likely to become more important.

2.3.Challenges in evidencing impact

Beyond the difficulties in defining and reporting on impact, evidencing impact is also not without its problems. Penfield et al. (2014) summarizes five challenges associated with impact, most relating to the timeframe in which impact occurs; time-lag (impact takes time), developmental nature of impact (impact does not just happen in 1 day), attribution (not easy to attribute to one single source), knowledge creep (research understanding increases over time), and finally, gathering evidence (difficult to collect information retrospectively).

McCowan (2018) similarly reports on some well-known difficulties in assessing impact such as; difficulties of evaluating impact, the possible threat to blue skies research, and short-termism, alongside less covered perils like impact relationships not being linear. However, McCowan (2018: 292) states the principle danger is 'when it starts to dominate, emphasis on impact can undermine the practice of enquiry that is at the heart of the university'. For CEDA users, this is likely to be a particular concern, as not all research in which they are involved is likely to be particularly impactful

(as defined by REF) for example. Blue skies theoretical research may have considerable benefits longer term but these may not be immediately relatable to public issues (Watermeyer 2011), or driving the research in the first instance.

There are then logistical and personal difficulties for researchers involved in recording and evidencing impact. Marcella, Lockerbie and Bloice (2015) interviewed established researchers who had work submitted for REF 2014. Despite being established, they still had difficulties with translating their work into REF templates; stating communicating impact using the REF template was difficult and time-consuming. The lack of agreed definitions for impact was also identified as a difficulty for researchers producing case studies; with some stating advice was ill informed and lacking, a problem also identified in other studies (Samuel and Derrick 2015). These difficulties sit against a backdrop of broader changes to academic work, with increasing pressures on researchers, and an erosion of trust acknowledged (Macfarlane 2011). There are also criticisms that the pressure to demonstrate impact has been so pronounced that policy interventions have 'outpaced theory' (Williams 2020).

Literature then suggests that academics are ill-prepared or trained in areas around documenting impact (Kellard and Sliwa 2016), and challenges with demonstrating impact, such as it not being easily measured or tangible, increase these difficulties (Marcella et al. 2018). Career stage can also lead to different views or understanding around impact (Wilkinson 2017). Marcella et al. (2018) found early career researchers tended to not be able to identify impact from their own personal research and did not sense any obligation to do so either. Few felt comfortable with giving conclusive evidence for cause and effect of their research impact. Marcella, Lockerbie and Bloice (2015) hypothesized early career researchers would learn about impact from experiences of established researchers; however, early career researchers have reported minimal mentoring opportunities and lack of institutional support (Marcella et al. 2018).

In summary, many of the issues highlighted around defining, recording, and evidencing impact are relevant to CEDA and its users. Whilst there is the possibility that the impact culture will increasingly be built into the research lifecycle in future (Marcella, Lockerbie and Bloice 2015; Solans-Domènech et al. 2019) the research community is yet to be at that point. Therefore, this research aimed to further explore this context. The research focussed on one overarching research question, informed by two sub-questions:

1. How can services such as CEDA gather evidence of impact?
   a. How willing are users to provide impact information to services such as CEDA?
   b. What processes are most useful in gathering impact information?

3. Methods

Tashakkori and Teddlie (1998) argue using a mixed methods approach, utilizing both qualitative and quantitative data, can give broader and more detailed perspectives within a research project. Both qualitative and quantitative designs were employed within this research; via a series of six focus groups and an online survey, which sought to gather CEDA users' views, preferences, and suggestions. Young et al. (2013) state early engagement with key stakeholders can lead to increased support, implementation and uptake of an intervention by a community. Therefore being equipped with a greater understanding of CEDA users' needs and perspectives, as well as involving them in the design of impact processes from the beginning, intended to promote a culture change for the provision of impact evidence, which also equipped users with a better understanding of why impact knowledge is necessary for CEDA.

Barbour (2014) states that one of the most efficient way of collecting views of individuals is via focus groups; and six focus groups were facilitated for this research, allowing detailed exploration of participants' views, opinions, and their own experiences of reporting impact and leading to rich discussions between participants (Vaughn, Shay Schumm and Sinagub 1996). Focus groups are often used within preliminary stages of a research project, drawing on the data to develop themes for inclusion in questionnaires (McLeod et al. 2000; Wackerbarth, Streams and Smith 2002; Leiss, Lyden and Klein 2011; Galliott and Graham 2016). This was an approach utilized in this research whereby the focus group conversations then informed the design of an online questionnaire survey which was sent out to CEDA users in July 2018.

The focus groups were held in June and July 2018 at existing events and conferences which CEDA users were known to attend, this included the JASMIN conference (three focus groups), the National Centre for Atmospheric Science (NCAS) staff meeting (two focus groups), and the Atmospheric Science Conference (one focus group). The events were expected to include a range of CEDA users such as JASMIN users, NCAS staff members, and potentially non-specialist members of the public (who can also access CEDA services) at the Atmospheric Science Conference. Focus group participants were recruited via a self-sign-up sheet at the events, with details also sent in advance to attendees of the JASMIN and NCAS events, and included on all event websites. Focus groups ranged in size from six participants in the largest group to four in the smallest group and in total, 26 people participated, including three people who were CEDA staff members now or in the past. Discussions varied in length from 20 to 40 min. This was constrained by each events timetable, with the same list of topics, and questions used for each discussion (see Supplementary information). In all focus groups, there was a general welcome and ice breaker, followed by questions covering topics such as participants dependency on CEDA and how they currently report impact.

The focus group discussions then informed the design of the online survey. This mainly related to questions 10, 11, and 12 on the survey, which asked users about the potential benefits (if any) of CEDA collecting and sharing users impact stories, how that should be done, and ways in which the process could be incentivized. The focus group data were able to inform the options for these questions. Survey participants were sampled by selecting all JASMIN users (less than 2,000), as well as CEDA Archive users that had used their accounts within the past three years. Once duplicates and bounced emails were removed, the survey was sent to 17,118 users. An initial email invitation was sent to users between 19 and 21st July, with a reminder email sent between 15 and 16th August 2018. Questions in the survey covered topics, which were similar to those in the focus groups (see Supplementary information), as well as some new themes from the focus group data. This included how often to collect impact information, and current ways that respondents were reporting impact.

All six focus group discussions were digitally recorded, with multiple devices, and transcribed verbatim by one researcher (P.T.) to increase familiarization with the data (Sutton and Austin 2015). Thematic analysis was applied to transcripts, identifying nodes for

common themes using Nvivo 12. Transcripts were analysed by one researcher (P.T.) and followed five steps outlined in Castleberry and Nolen (2018). These include compiling (creating transcripts), disassembling (creating meaningful codes), reassembling (putting codes into themes), interpreting (make analytical conclusions from data), and concluding (relating interpretations to research questions asked). Themes were discussed with the second researcher (C.W.) and revised when relevant. Survey data were explored by using both univariate and bivariate analysis, via Microsoft Excel and SPSS. Statistical analysis was carried out using Pearson Chi-Square, with the Likelihood Ratio and Fishers Exact Test used when appropriate to cell size. An alpha level of 0.05 was used for all statistical tests. Full ethical approval was granted by UWE Bristol.

4. Results

4.1 Focus group data

Over the six focus groups, seven key themes were identified including dependency on CEDA services, processes for reporting impact, collecting impact information, incentives, benefits for researchers, reasons for providing information, and the sharing of impact stories. 'Dependency on CEDA services' often came to light early in the focus groups, when an ice breaker opened with their use of the service, and was a theme identified over all six focus groups. Comments at groups 1–3 (held at the JASMIN conference) tended to focus heavily on the use of JASMIN with only some mentions of the CEDA Archive, whereas groups 4–6 tended to talk more about their reliance on the Archive and this was likely related to the different users attending each conference:

> I couldn't do my job without it [JASMIN]... We have so much data coming from so many places. There's not a local resource that can handle it. (Ryan, Focus Group 5)

A number of participants suggested CEDA services were essential for their work, whilst other focus group participants pointed out if CEDA did not exist they would find some other way to do their work; but agreed the services made it easier:

> If we didn't have [CEDA], then we could just chop and change, make up our own minds, and things wouldn't be standard. We could survive [without CEDA], but it wouldn't be as good. (Alan, Focus Group 5)

The focus groups also took the opportunity to discuss 'Processes for reporting impact' based on users' current activities. Responses included formal or informal reporting. However, some participants were confused about whether they report impact to their institutes at all. A range of formal reporting processes was mentioned by participants, such as reports, email exchanges, and forms/surveys, as well as Researchfish for those who were supported by Research Councils, but there was not one consistent formal process that all participants reported their impact to. Some mentioned informal processes like choosing to update webpages or telling a contact impact had happened, without a request for information, but there was a sense amongst many participants that this was also something they were responsible for and informal reporting (n = 5) was far less frequently referred to in the transcripts than formal (n = 26):

> For us as a research group, we have our own webpage ... Partly for our own profiles, we're not required to do it. (Michelle, Focus Group 6)

> I think you have to take on your own responsibility for it as well ... [Institute contact] can highlight them for us, but if we've got something that we think would make a good impact case study then it's our responsibility to report that. (Ruth, Focus Group 6)

However, when probed on 'collecting impact information', and why that might happen a range of views were expressed in five of the six groups. Firstly, participants thought impact stories could help raise awareness of resources available to users, reduce duplication of effort, and 'stop people reinventing the wheel' (Zoe, Focus Group 6):

> Often people don't know whether a particular facility is useful to them or not. But if you can see a story of how somebody else with a similar problem has used a facility, then you can decide if it's appropriate for you to use it too. (Bryan, Focus Group 2)

> It's nice to hear how other people are doing things. We can either say I don't want to do it like that I'll do it my own way still, or piggyback off it. (Anonymous, Focus Group 4)

A further reason for collecting impact information was to provide evidence of the service's usefulness to science; whether that is to funding bodies such as UKRI, relevant government departments (BEIS), or the UK taxpayer, but the complimentary nature of services like CEDA, to research in and of itself was seen to make this more complex:

> Do we still need this spanner? No-one uses the 9/16th anymore, let's just get rid of it... [Laughter] (Alan)
> [Laughing] That's not a metric. (Ryan)
> No but, I think there is some kind of a disconnect between the impact which occurs from one research group, and you guys [at CEDA] who facilitate that. (Mat, Focus Group 5)

CEDA was seen as an important facilitator for science, yet, there was confusion about how best to evidence this and concern that such a service could disappear if it is not adequately recognized. How to evidence impact of CEDA was therefore discussed over all focus groups and suggestions largely revolved around collecting via existing processes; such as university impact officers, Researchfish, posters/talks from JASMIN conference, and project contacts, in order to reduce the repetition of impact mechanisms:

> Make sure you're not asking people for information they're already giving to someone else ... there's nothing worse than being repeatedly asked for the same information but in a slightly different form. (Louise, Focus Group 4)

The most common suggestion was to talk to contacts sitting within universities, such as impact officers, and align how CEDA collect impact information with how universities are already capturing it. Other suggestions included a user survey, with one participant even suggesting it could be compulsory to complete for continued access to the services.

> It's a fantastic resource and we're given it for free. We should be required ... [to] send information, so you [CEDA] get a better idea of impact. (Andrew, Focus Group 1)

However, many participants indicated they do not like surveys and have 'survey fatigue', and that there should be some kind of warning mechanism before a service was ever removed.

There were, therefore, mixed views on how to ask for information. There was an overwhelming consensus from participants about the necessity to keep length of text and effort required minimal.

Many agreed they would like to 'give it in bullets and then someone else translate that into impact speak' (Victoria, Focus Group 3). The idea of building upon basic initial information was popular. Participants agreed they would rather build upon information when they know it is relevant and useful to CEDA; rather than wasting time on something that would not be used. The use of templates or examples was also suggested, or a CEDA-based staff member to help with narratives. Realistic deadlines with reminders were also seen to be important. How often to collect impact information was discussed, with many agreeing annually was about right. However, there were difficulties with knowing when to stop asking about a project's impact, as discussed in Group 5, due to the time-lag as to when impact develops.

> You'd probably want to follow up every year, for 10 years, after that [project ends] (Mat)
> Yeah, it won't come in the first six months. (Ryan)
> Exactly. (Mat)
> It's normally like two years later. (Ryan)
> And if what you want to say is 'government policy was changed' ... You're looking at a decadal timeline between data going onto CEDA, and the policy change occurring. (Mat, Focus Group 5)

There were then a number of recognized challenges within the groups. Difficulties for individuals often revolved around issues associated with measuring and evidencing impact, alongside a feeling of not knowing how to explain or describe impact.

> I think it's a challenge for all of us scientists to formulate the real world impacts ... I was being interviewed ... they asked me to say 'and so what does that enable' and the words just weren't there ... We need to get better at putting that message across because that's what the real public want to know. (Victoria, Focus Group 3)

Participants indicated they would like to overcome this difficulty, but a lack of knowledge and/or confidence was holding them back. They were also sometimes unaware that information was not getting back to CEDA. One participant gave the example of a paper in *Nature* and the role of CEDA being promoted, but this is challenging when the service may not be aware of such a papers existence, and this was recognized to be a common problem over a number of similar services.

'Incentives' for providing impact information were also discussed, with rewards being more popular than compulsory options, despite a number of participants saying that removing access to resources would be 'the only way I would do it [provide information]' (Ryan, Group 2). Incentives suggested included sharing the work of researchers (e.g. on a website), offering digital object identifiers (DOIs), more resources and financial rewards, or even sweets and chocolates, in some humorous interactions.

However, it was also recognized that there were wider 'benefits for researchers' including that thinking about impact could be personally useful to them for developing their work and ideas, or to reflect on the impact their work was having:

> It helps you to focus on the problem you're trying to achieve. Sometimes you can drift a little bit and it just brings you back to the real issues. It can actually help you to think of new ideas. (Anonymous, Focus Group 1)

> It helps us keep a record of what we're doing and it means external people ... can see the different things we do. (Michelle, Focus Group 6)

In terms of the 'reasons for providing information', the most commonly expressed reason was due to the external pressure to provide it, be it by a project, a funder, or an institutional contact asking them for information. Conversely, there were also many reasons why it was not being provided, the most common being that they did not know what they were being asked for or 'what impact is'. Users were dissuaded if they thought information was not going to be used, had no benefit to them, or they were frustrated with the collection process.

There also appeared to be variations based on peoples' career stage, with those earlier in their careers less clear what impact requirements there might be:

> I don't think I do [report impact]. Not impact. I report what I'm doing but... (Marios)
> Fortunately, one of the beauties of being a PhD student is that you don't have to do that, your supervisor does. (Bryan)
> Facilitator: That's another issue – that it doesn't get filtered down to PhD's and postdocs?
> Yes, I think that's very much a problem. Most tools that gather impact information are pointed at people that have got money versus people that are doing the work. They are not the same people. (Bryan, Focus Group 2)

Despite the issues outlined, when the focus groups finally discussed ways to 'share impact stories', the conversations were often animated and constructive. Firstly, participants came up with several ideas around sharing impact stories including written monthly highlights or staff presentations at meetings and conferences. Other participants suggested CEDA could 'sell' its services more and focus on benefits to users, discussing how it 'will make people's lives so much easier' (Freya, Focus Group 6):

> I'd like to hear [about] impact stories, but I know other people are using CEDA more effectively than I am, and I'd like to know how we can do other things. (Ryan, Focus Group 5)

However, it was also highlighted that it was important impact was the key focus, and the primary reasons for gathering knowledge of impact, rather than straying too far into a marketing agenda.

4.2 Questionnaire data

Following the focus groups, questionnaires were sent to ~17,000 users with a response of 520 users over the 4 weeks the survey was open (3% response rate). The initial section of the survey explored the demographics of the respondent in order to compare to the wider users of the CEDA database. In terms of interest areas, the most popular area of interest was climate change (27%, n = 140), followed by earth science (11%, n = 59) mirroring the most popular areas of interest in the database. Most respondents were working in universities (66%, n = 367), followed by government (13%, n = 74) which is again similar to the main database. Thirty-three per cent (n = 173) of respondents were postgraduate students, 28% (n = 150) were postdoctoral researchers, 13% (n = 70) were professors, and 22% (n = 117) defined themselves as 'other'. As career stage information is not regularly updated on the CEDA database, it could not be compared to wider users.

The majority of respondents 84% (n = 436) used the CEDA Archive and 28% (n = 148) used JASMIN, with 64 respondents using both services. Most respondents were using the services on a weekly or monthly basis. The majority of users (CEDA Archive 61%, n = 268, JASMIN 68%, n = 100) indicated that they would be

willing to provide impact information to CEDA. Users that disagreed (CEDA Archive 6%, n = 26, JASMIN 7%, n = 10) tended to use the services least frequently.

In terms of current reporting processes, half of users (n = 260) said that their institution currently requires them to submit information on impact. Of those who were not submitting impact information, common reasons included not being senior enough, not being asked, the type of institution based in, lack of time, lack of importance, or because they were no longer affiliated to an institution. Although there was some variation in responses by career stage, this was not statistically significant (see Figure 1).

The most common impact processes highlighted in the questionnaire data were reporting to a funder (74%, n = 187) and project leads (72%, n = 173), whereas the least common was talking to the departmental impact officer (27%, n = 54) or other processes not described (21%, n = 25) (see Figure 2). It is interesting to note a quarter of respondents were unsure whether they submit REF case studies (27%, n = 55) or talk to their impact officers (26%, n = 52). The majority of participants keep track of their impact information via a personal record (69%, n = 150), and it should not be assumed they were necessarily engaging in formal reporting mechanisms, like REF impact studies or reporting on pathways to impact. In a separate question, 8% (n = 42), of those surveyed, had participated in an REF impact case study, and 4% (n = 21) in an RCUK/UKRI impact case study.

Despite there being relatively little experience with existing impact reporting mechanisms, there was some understanding as to why such processes are and might be put in place. Participants agreed that the collecting and sharing of CEDA impact stories could 'highlight the impact CEDA services are having on researchers' (82%, n = 420), 'provide evidence for funding bodies' (82%, n = 418), as well as 'impact on wider society' (79%, n = 405) and 'highlight the work of groups of researchers' (74%, n = 380) (see Figure 3). Relatively few respondents saw 'no foreseeable benefits' (11%, n = 50) in gathering the information.

In regards to how these data should be collected, there was very little agreement that it should be compulsory. Only 22% (n = 112) of respondents thought this should be the case, and the most popular options were to use channels that impact information is already submitted to (67%, n = 336) or to submit information at the end of a research project (64%, n = 332). Though 33% (n = 167) respondents were neutral regarding the creation of a CEDA impact portal or web form, 56% (n = 279) agreed or strongly agreed that this could be an option for collection. There were no obvious patterns as to whether information should be collected yearly, biannually or over longer periods, with the most preferred choice simply being at the end of a project.

Building on the focus group data, the questionnaire took the opportunity to explore the potential incentives and disincentives in any impact capturing process. Perhaps unsurprisingly the most popular incentives were those which offered an obvious benefit to researchers including the 'use of new datasets' (74%, n = 372), 'provision of a DOI' (67%, n = 334), and 'advertising their research to other users' 64% (n = 312) (see Figure 4). The potential to inform funders that impact information had not been correctly provided by a user (18%, n = 89) or to remove access to resources (18%, n = 88) had far lower levels of agreement when compared to other options. Open responses to this question suggested small gifts such as vouchers, sharing news on impactful research, making it as easy as possible, building it into existing mechanisms in the services, providing support, and being clearer about why CEDA needs the impact information, could all add to incentivize the process of impact reporting.

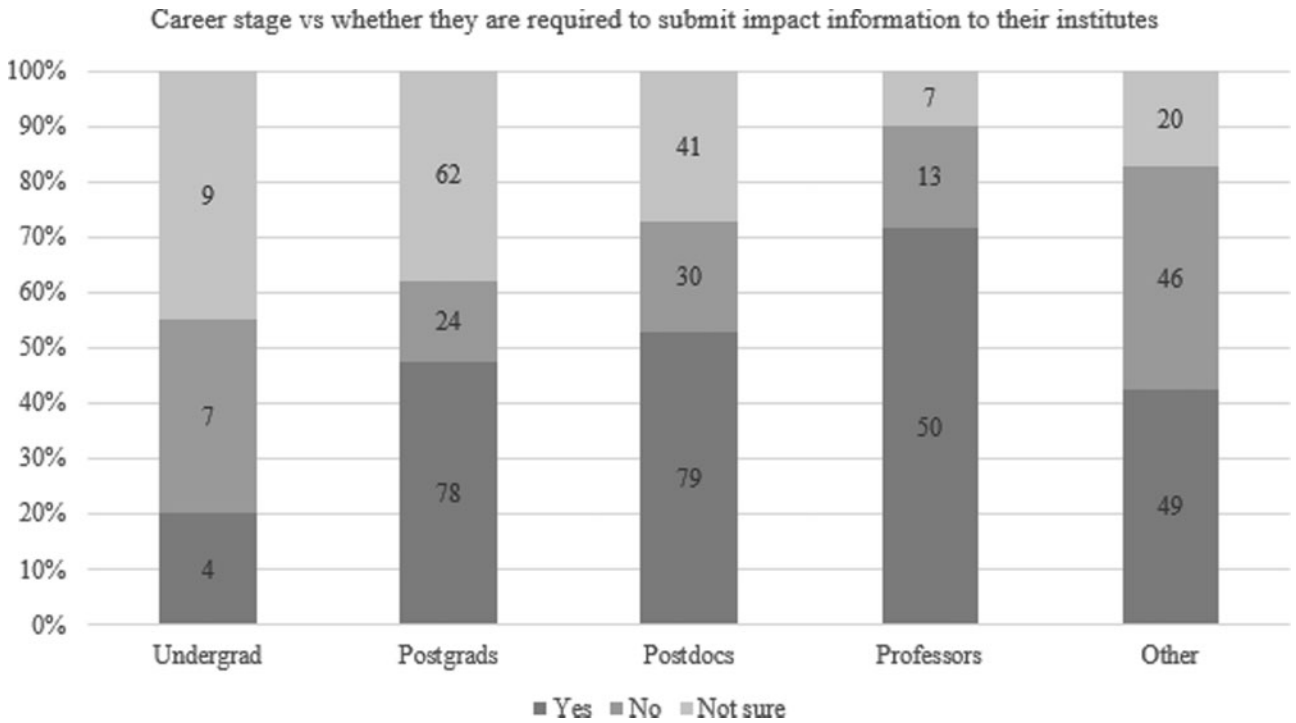

Figure 1. Respondents career stage and their answers to the question 'Does your institution require you to submit information about the impact of your research?'. Note—one non-response.

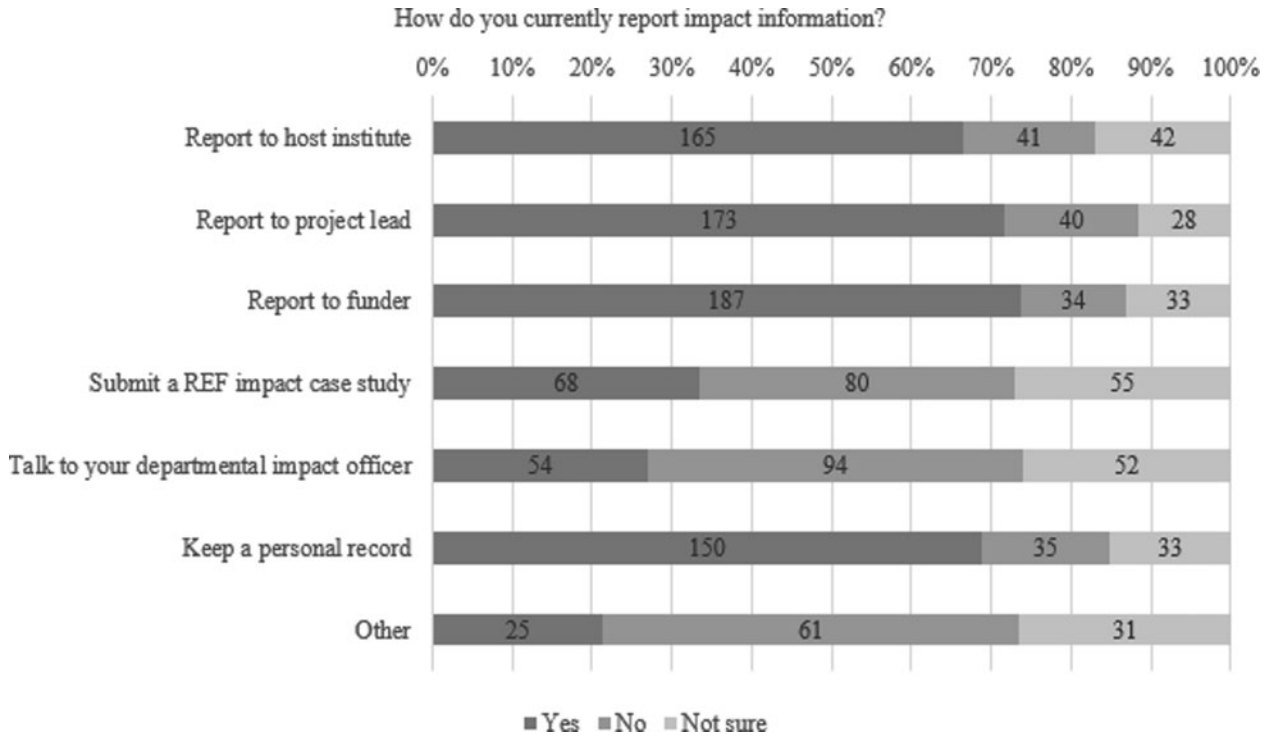

Figure 2. How respondents currently report impact information. This was a filtered question so reduced response rates were expected—highest participant response was 254 (report to funder) and lowest 117 (other).

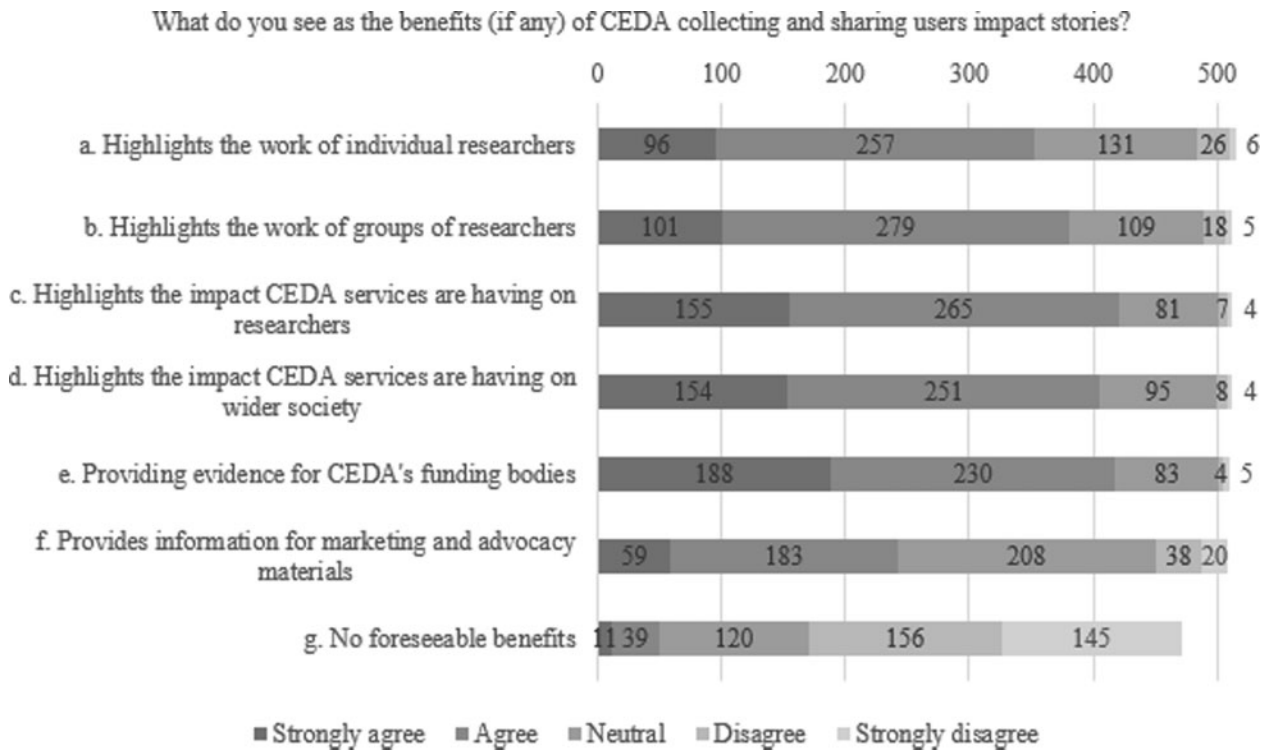

Figure 3. Respondents agreement to the different benefits that could come from CEDA collecting impact information.

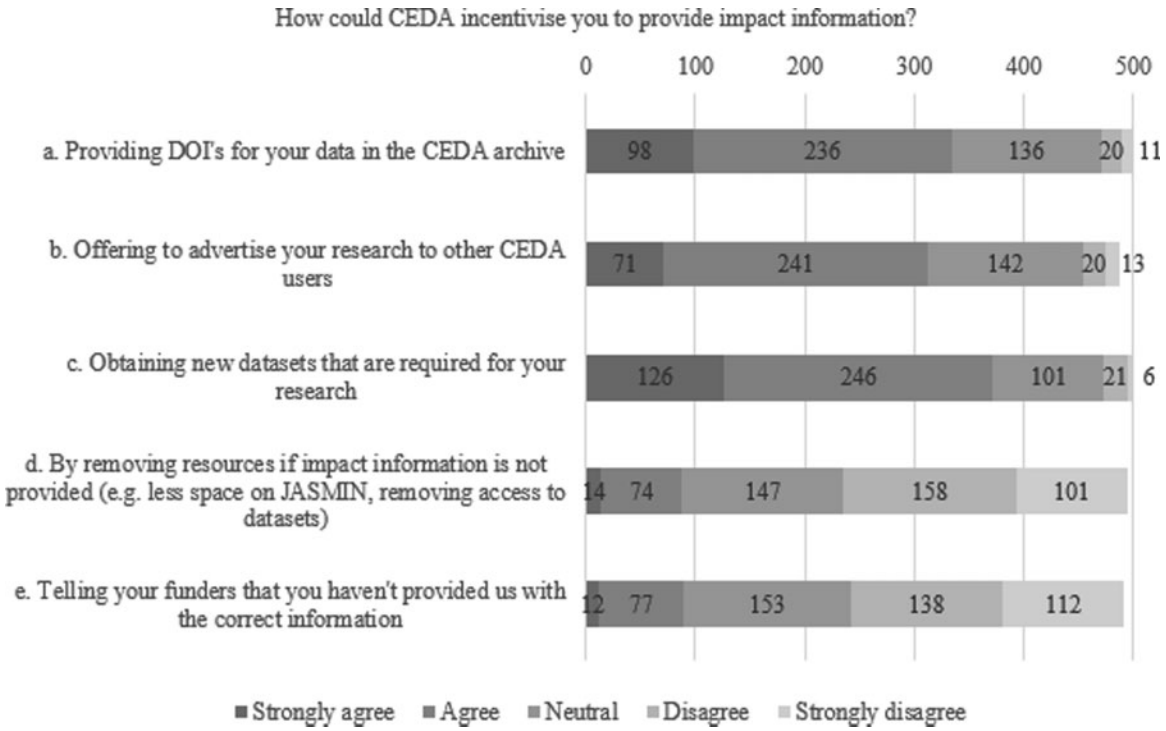

Figure 4. Incentives.

Although some researchers discussed how they may or may not benefit from evidencing the impact of CEDA, this mainly focussed on minimizing the 'costs' of administration that would be required in relation to benefits for them. There were rarely comments in the focus group or questionnaire data around the potential for 'game playing' or researchers making more cynical use of such processes, though one person commented:

> Unfortunately, I think that reporting impact to a data provider would cause a circle to develop; the data provided would be led by the researchers which have the greatest impact. This is of course flawed by the fact that impact varies depending on the media, other current events, investors' interests etc. . . I do not believe a measure of impact should in any way impact the priority of one researcher over another, for example for on websites or marketing for funding. (Questionnaire Respondent)

In addition to incentives, respondents also reflected a number of other points. They discussed a lack of time, being over-burdened, or feeling the pressure of 'red-tape', struggling to understand definitions of impact, as well as the potential to exaggerate or neglect blue skies research. However, often these comments were contextualized by positive comments about the service itself and a desire to continue to utilize it. Thus, there is the dilemma once again of how to balance research users' needs and time, with that of the desire to reflect impact:

> I am sure you are doing great work, but please don't generate even more work load for researchers by forcing them to invent implausible 'impact stories'. The real impact of most research won't be known for decades most of the time, and that's fine. Einstein didn't sit down and ponder over differential geometry for years because he wanted to make satellite navigation possible. (Questionnaire respondent)

5. Discussion

There are several limitations in this research, the most pertinent being potential bias due to one of the researchers being a CEDA staff member and that the survey response rate could be further increased. Every effort to reduce bias was implemented by designing a clear methodology and being aware and open to potential conflicts of interest. In addition, the self-selecting nature of the methods used means respondents may have chosen to participate for reasons that are systematically related to the characteristics of the study (Lavrakas 2008), most obviously that they are already open and favourable towards the impact agenda. Several areas in the research could be improved; including more thorough piloting of the focus group questions to improve clarity and encourage methodological rigour (Lancaster, Dodd and Williamson 2002). Although carrying out focus groups during relevant meetings helped to identify potential participants (Breen 2006), competition with parallel sessions during the host events was challenging and wider promotion could have helped increase participation.

Despite these issues, the results suggest that the majority of users (61%, n = 312), who participated in the survey, were willing to provide impact information to CEDA. However, the research also found there was confusion around what 'impact' is and many found it challenging to evidence; reflecting the current literature (Marcella, Lockerbie and Bloice 2015; Samuel and Derrick 2015; Wilkinson 2017; Marcella et al. 2018; Solans-Domènech et al. 2019). In the qualitative data, this was slightly more prominent for early career researchers who lacked awareness and/or the requirement to report the impact of their work (Marcella et al. 2018). This presents issues for services such as CEDA, who must then compete with the other demands on researchers' time, but it also suggests there are opportunities to engage with users earlier in their careers, to think about

impact prior to the demands of funding reporting requirements or sustained engagement with REF.

In terms of the processes that are most useful in gathering impact information, the research suggests three key suggestions for collecting impact information amongst service providers such as CEDA. Firstly, collect information via existing processes. Many participants indicated frustrations at repeatedly giving impact information to various sources. Previous research has suggested intermediaries or 'impact officers' are most effective for collecting impact information when close relationships are formed (Manville et al. 2015; Wilkinson 2017), however, less than 30% of survey respondents liaised with their impact officers or central systems in their universities. The data showed that no single process for collecting impact information was used across the user community. Either this means, information has to be collected from a variety of existing sources or, services such as CEDA are required to create their own systems, which potentially could include drawing down information from multiple sources. Secondly, users indicated the frequency of collection was important, with the most popular suggestion being at the end of a project. However, for services such as CEDA, which are not tied to dedicated project funding, this is difficult to manage. There are also wider concerns relating to the period for collecting information, and the appropriate 'impact window' (Kings College London and Digital Science 2015; Manville et al. 2015; Morton 2015) which accounts for the time-lag of impact development (Penfield et al. 2014; Solans-Domènech et al. 2019). Reporting at the end of a project may only be possible if a system is put in place, which would specifically prompt users to report. The final key suggestion was to collect preliminary impact information first, and then follow-up for more details only if the impact is relevant to the service. It was also deemed important to provide support and templates to the researchers when writing the more detailed impact story, mirroring recent research (Wilkinson 2017).

The suggestions emerging in the research have a common theme; to avoid time wasting and repetition, approaches which can be considered by similar service providers seeking to measure impact. As the impact agenda continues, with governments and funding bodies expecting impact evidence (Morgan 2014; Reed 2016), this type of research will become increasingly important in discovering best practice for service providers. Researchers are already under immense pressure (Macfarlane 2011; Kassab, Mutz and Daniel 2020;) to provide impact information to multiple sources; institutes, REF, project reports, amongst many others. Manville et al. (2015) found that evidencing impact imposes burdens on researchers. Service providers, like CEDA, should be mindful of these pressures and aim to minimize the additional requirements on users as much as possible. Considering how impact information systems can be rationalized requires careful consideration. It may not be a case of one size fits all. For smaller services and providers, playing an integral role in the impact landscape, evidence of impact may need to be commensurate to the scale of an organization and the interactions users have with it.

As the results show no-one process is endorsed by all users but there was considerable evidence that users were keen to learn and benefit from impact reporting but are often distracted by reporting mechanisms which focus more on accountability than understanding (Penfield et al. 2014). It is therefore important that the need to provide impact information is integral to requests made of researchers, but also that those gathering information keep in mind potential benefits in learning and sharing as one of their commitments. As repetition may be unavoidable, it is essential to explain the need and purpose in gathering such information, but also to elaborate on the wider benefits that may exist for researchers.

Supplementary data

Supplementary data are available at *Research Evaluation Journal* online.

Acknowledgements

The authors would like to thank all focus group participants and survey respondents who gave up their time to contribute in this research, as well as CEDA for access to information and permission to carry out the research.

Funding

This work was supported by the Centre for Environmental Data Analysis (CEDA) who funded Poppy Townsend's MSc Science Communication, which this research formed a part of.

References

Barbour, R. (2014) 'Analysing Focus Groups', in: Flick U. (ed.) *The SAGE Handbook of Qualitative Data Analysis*, pp. 313–26. London: SAGE Publications Ltd.

Breen, R. L. (2006) 'A Practical Guide to Focus-Group Research', *Journal of Geography in Higher Education*, 30: 463–75.

Castleberry, A., and Nolen A. (2018) 'Thematic Analysis of Qualitative Research Data: Is It as Easy as It Sounds?', *Currents in Pharmacy Teaching and Learning*, 10: 807–15.

CEDA (2018a) *Data and Information Services for Environmental Science* <www.ceda.ac.uk> accessed 8 March 2018.

CEDA (2018b) *CEDA Archive* <www.archive.ceda.ac.uk> accessed 8 March 2018.

ESRC (2017) *Research Funding Guide October 2017* <http://www.esrc.ac.uk/funding/guidance-for-applicants/research-funding-guide/> accessed 9 March 2018.

Galliott, N., and Graham L. J. (2016) 'Focusing on What Counts: Using Exploratory Focus Groups to Enhance the Development of an Electronic Survey in a Mixed-Methods Research Design', *Australian Educational Researcher*, 43: 567–85.

HEFCE (2016) *REF Impact* <http://www.hefce.ac.uk/rsrch/REFimpact/> accessed 3 March 2018.

HEFCE, RCUK and UUK (2011) *HEFCE, RCUK and UUK Joint Statement on Impact* <http://webarchive.nationalarchives.gov.uk/20120104102800/http://rcuk.ac.uk/kei/maximising/Pages/Impactstatement.aspx> accessed 7 Nov 2018.

JASMIN (2018) *Petascale Storage and Cloud Computing for Big Data Challenges in Environmental Science* <www.jasmin.ac.uk> accessed 8 March 2018.

Kassab, O., Mutz R., and Daniel H. D. (2020) 'Introducing and Testing an Advanced Quantitative Methodological Approach for the Evaluation of Research Centers: A Case Study on Sustainability Science', *Research Evaluation*, 29: 135–49.

Kellard, N., and Sliwa M. (2016) 'Business and Management Impact Assessment in REF2014: Analysis and Reflection', *British Journal of Management*, 27: 693–711.

Kings College London and Digital Science (2015) *The Nature, Scale and Beneficiaries of Research Impact: An Initial Analysis of Research Excellence Framework (REF) 2014 Impact Case Studies*. Bristol: HEFCE.

Lancaster, G. A., Dodd S., and Williamson P. R. (2002) 'Design and Analysis of Pilot Studies: Recommendations for Good Practice', *Journal of Evaluation in Clinical Practice*, 10: 307–12.

Lavrakas, P. J. (2008) *Encyclopedia of Survey Research Methods*. Thousand Oaks, CA: Sage Publications, Inc.

Lawrence, B. N., and Townsend P. (2016) *JASMIN Science Case*. Centre for Environmental Data Analysis (CEDA) <http://cedadocs.ceda.ac.uk/1350/> accessed 8 March 2018.

Leiss, J. K., Lyden J. T., and Klein C. (2011) 'Using Formative Research to Design an Epidemiologic Survey: The North Carolina Study of Home Care and Hospice Nurses', *Epidemiology and Health*, 3: 1–7.

Macfarlane, B. (2011) 'The Morphing of Academic Practice: Unbundling and the Rise of the Para- Academic', *Higher Education Quarterly*, 65: 59–73.

Manville, C. Morgan Jones, M., Henham, M.L., Castle-Clarke, S., Frearson, M., Gunershekar, S., and Grant, J. .(2015) *Preparing Impact Submissions for REF 2014: An Evaluation*. Cambridge: RAND Corporation.

Marcella, R., Lockerbie H., and Bloice L. (2015) 'Beyond REF 2014: The Impact of Impact Assessment on the Future of Information Research', *Journal of Information Science*, 42: 369–85.

Marcella, R. et al. (2018) 'The Effects of the Research Excellence Framework Research Impact Agenda on Early- and Mid-Career Researchers in Library and Information Science', *Journal of Information Science*, 44: 608–18.

McCowan, T. (2018) 'Five Perils of the Impact Agenda in Higher Education', *London Review of Education*, 16: 279–95.

McLeod, P. J. et al. (2000) 'Using Focus Groups to Design a Valid Questionnaire', *Academic Medicine*, 75: 671.

Morgan, B. (2014) 'Research Impact: Income for Outcome', *Nature*, 511: S72–5.

Morgan Jones, M. et al. (2013) *Assessing Research Impact: An International Review of the Excellence in Innovation for Australia Trial*. Santa Monica, CA: RAND Corporation.

Morton, S. (2015) 'Progressing Research Impact Assessment: A "Contributions" Approach', *Research Evaluation*, 24: 405–19.

NCEO (2017) *The National Centre for Earth Observation Impact Brochure* <https://www.nceo.ac.uk/about-us/nceo-impact-brochure-2017/> accessed 23 October 2018.

NEODAAS (2018). *Our Impact*. https://www.neodaas.ac.uk/supportedscience/> accessed 23 Oct 2018.

NERC (2016) *NERC Environmental Data Centres—Future User Needs Survey* <https://nerc.ukri.org/about/whatwedo/engage/engagement/datacentres-survey/> accessed 23 Oct 2018.

NERC (2018) *NERC Research Grants and Fellowships Handbook* <http://www.nerc.ac.uk/funding/application/howtoapply/forms/grantshandbook/> accessed 9 March 2018.

Penfield, T. et al. (2014) 'Assessment, Evaluations, and Definitions of Research Impact: A Review', *Research Evaluation*, 23: 21–32.

Reed, M. S. (2016) *The Research Impact Handbook*, 1st edn. Newcastle: Fast Track Impact.

REF2021 (2018) *What Is the REF?* <https://www.ref.ac.uk/about/what-is-the-ref/> accessed 7 Nov 2018.

REF Steering Group (2019) *Key Decisions* <https://www.ref.ac.uk/about/key-decisions/> accessed 12 June 2019.

Samuel, G., and Derrick G. (2015) 'Societal Impact Evaluation: Exploring Evaluator Perceptions of the Characterization of Impact under the REF2014', *Research Evaluation*, 24: 229–41.

Solans-Domènech, M. et al. (2019) 'Development and Validation of a Questionnaire to Measure Research Impact', *Research Evaluation*, online first, 28: 1–10.

STFC (2018) *Research Grants Handbook—5.8 Supporting Information* <https://www.stfc.ac.uk/research-grants-handbook/5-applying-for-a-grant/5-8-supporting-information/#5.8.3> accessed 9 March 2018.

Sutton, J., and Austin Z. (2015) 'Qualitative Research: Data Collection, Analysis, and Management', *Canadian Journal of Hospital Pharmacy*, 68: 226–31.

Tashakkori, A., and Teddlie C. (1998). *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Thousand Oaks, CA: Sage Publications.

UKRI (2018a) *Excellence with Impact* <https://www.ukri.org/innovation/excellence-with-impact/> accessed 23 Oct 2018.

UKRI (2018b) *Pathways to Impact* <https://www.ukri.org/innovation/excellence-with-impact/pathways-to-impact/> accessed 23 Oct 2018.

Vaughn, S., Shay Schumm J., and Sinagub J. (1996) *Focus Group Interviews in Education and Psychology*. Newbury Park, CA: Sage.

Wackerbarth, S. B., Streams M. E., and Smith M. K. (2002) 'Capturing the Insights of Family Caregivers: Survey Item Generation with a Coupled Interview/Focus Group Process', *Qualitative Health Research*, 12: 1141–54.

Watermeyer, R. (2011) 'Challenges for University Engagement in the UK: Towards a Public Academe?', *Higher Education Quarterly*, 65: 386–410.

Watermeyer, R. (2012) 'From Engagement to Impact? Articulating the Public Value of Academic Research', *Tertiary Education and Management*, 18: 115–30. p

Wilkinson, C. (2017) 'Evidencing Impact: A Case Study of UK Academic Perspectives on Evidencing Research Impact', *Studies in Higher Education*, 44: 72–85.

Williams, K. (2020) 'Playing the Fields: Theorizing Research Impact and Its Assessment', *Research Evaluation*, 29: 191–202.

Young, J. C. et al. (2013) 'Does Stakeholder Involvement Really Benefit Biodiversity Conservation?', *Biological Conservation*, 158: 359–70.