Research Evaluation, 29(1), 2020, 4–21
doi: 10.1093/reseval/rvz033
Special Section

Methods for mapping the impact of social sciences and humanities—A literature review

David Budtz Pedersen*, Jonas Følsgaard Grønvad and Rolf Hvidtfeldt

Humanomics Research Centre, University of Aalborg, A.C. Meyers Vænge 15, Copenhagen DK-2450, Denmark

*Corresponding author. Email: davidp@hum.aau.dk.

Abstract

This article explores the current literature on 'research impact' in the social sciences and humanities (SSH). By providing a comprehensive review of available literature, drawing on national and international experiences, we take a systematic look at the impact agenda within SSH. The primary objective of this article is to examine key methodological components used to assess research impact comparing the advantages and disadvantages of each method. The study finds that research impact is a highly complex and contested concept in the SSH literature. Drawing on the strong methodological pluralism emerging in the literature, we conclude that there is considerable room for researchers, universities, and funding agencies to establish impact assessment tools directed towards specific missions while avoiding catch-all indicators and universal metrics.

Key words: research evaluation; impact assessment; social sciences and humanities; literature review.

Introduction

Across the international research and innovation community there is a growing interest in how to assess and communicate the diverse impacts of scholarly work. Being able to demonstrate the societal uptake and value of social sciences and humanities (SSH) research is increasingly seen as a crucial component in ensuring accountability and transparency (Penfield et al. 2014; Morton 2015; Greenhalgh et al. 2016; Ravenscroft et al. 2017). In recent years, the notion of 'research impact' has gained significant traction within the science system, and has been embedded in research policies, funding instruments, and evaluation regimes (e.g. Rip 2000; Holbrook and Frodeman 2011; Bornmann 2013; Buchanan 2013; Langfeldt and Scordato 2015; Derrick and Samuel 2017; Holbrook 2017; Reale et al. 2017). In this article, we provide an overview of the existing methods for broader impact assessments across SSH.

A key finding of the literature review is that different funding agencies, policy-makers, and research organizations operate with different models and methods for impact assessment. Impact simply does not mean the same thing across institutions, geographies, and research cultures. This conceptual diversity is reflected in the number of methods and frameworks which are used to track, demonstrate, assess, and incentivize the impact of research across the European SSH community and beyond. The diversity of the impact agenda in SSH reflects a broader trend within impact studies. The evolution of impact studies has shown that public research organizations do not just release their benefits to society following a linear model of growth and application. Instead, real-world effects of research occur at different stages in the research process, extending from knowledge dissemination and knowledge mobilization to long-term applications and dynamic effects.

Much progress has been made in measuring both the outcomes of research and the processes and activities through which these are achieved (Greenhalgh et al. 2016). However, as we demonstrate in this article, there exists a multitude of approaches to impact assessment reflecting the complex and multi-dimensional ways in which research is taken up by society. As Rafols (2017) noted at the Science, Technology, and Innovation Indicators Conference in 2017: 'The contributions of science to society are so varied, and mediated by so many different actors, that indicators used in impact assessment cannot be universal. Instead, they need to be developed for given contexts and used alongside qualitative assessment'. Assessing the impact of social science and humanities is indeed challenging. The ways in which research is taken up, used, and reused in real-world settings mean that linking research processes or outputs to wider changes is difficult, and timescales are hard to predict (Morton 2015). However, rather than being paralyzed by the lack

of 'golden standards' for impact assessment in SSH, the multiplicity of methods and instruments, leads to a much more context-sensitive and nuanced understanding of impact than is usually promoted by any individual assessment model.

Hence, the aim of this article is to provide a comprehensive overview of the literature and examine the various methodological components currently used in SSH impact assessment. We show that the literature is composed of a myriad of approaches and methods that allow for considerable flexibility, context-dependence, and value-based assessment models. More specifically, we show that methods for tracking and assessing research impact vary significantly at the project-, programme-, or system-level.[1] Different approaches for assessing impact inevitably make different assumptions about the nature of knowledge production, the purpose of research, the definition of research relevance, the role of values, and the mechanisms by which impact is achieved, measured, and evaluated (Penfield et al. 2014).

The article starts by presenting the scoping review of SSH impact assessment models and methods. We then proceed to a comparative analysis of the literature, examining strengths, and weaknesses of the different methods and frameworks. Finally, we discuss what the findings may imply for the future of research on research impact in the SSH.

Scoping review

The study can be defined as a scoping review, which is a research synthesis that aims to map the literature on a particular topic or areas of research identifying key concepts, gaps in the literature, and types and sources of evidence to inform practice, policy-making, and research. A scoping review tends to address broader topics where many different scientific traditions, study designs, and types of evidence may exist (Arksey and O'Malley 2005; Daudt, van Mossel and Scott 2013). This review strategy is especially suitable as there is currently no agreement on what is the best definition or the superior methodology for assessing impact in SSH. The literature on impact assessment is not to be regarded as a coherent scientific literature but rather as a partly policy-driven agenda. Due to this complexity, the review strategy does not address the quality of the included models or studies, but instead seeks to chart the methodological pluralism of different assessment models and their perceived strength and weaknesses in the literature.

The selection process

The review was carried out in 2016–2017 by the authors as part of the EU Horizon 2020 ACCOMPLISSH Project. The review was based on a systematic search of scientific databases and web searches (grey literature). The review identified 158 academic papers and 125 policy documents describing various impact models and frameworks and their applications. The data selection process can be divided into two sequential steps:

a. A systematic search of publications was performed using (1) scientific literature databases and (2) internet searches. The digital databases used in the review were SCOPUS and Web of Science. For the web searches the review combined keywords into different search strings using google.com as the search engine. For each search string, the first 50 hits were examined closer. The initial corpus of text was identified by applying the inclusion and exclusion criteria on title and then abstract or summary.
b. A manual search of reference and citation tracing were then conducted on the initial corpus of documents identified and screened as part of the systematic search. The reference and citation tracing used the same procedure and eligibility criteria on title and abstract.

Additional documents were continually found and added to the corpus through expert and peer consultations if they met all inclusion criteria and was not already included in the study.

Keyword and eligibility criteria

The selected keywords used in the systematic searches were based on different types of impact, for example, 'societal', 'economic', 'health', 'policy', 'cultural', 'public' in relation to the 'humanities' or 'social science'. The inclusion and exclusion criteria applied related to 'time period' (between 2005 and 2015), 'language' (English), 'population' (SSH research), 'publication type' (academic and policy documents), as well as 'study characteristics' (concepts, methods, and tools for understanding, assessing, or encouraging societal impact). The review has chosen not to include texts that only addresses quality assessment and impact within the academic community. As a consequence, much literature on peer review processes and bibliometric indicators are not accounted for.[2] A detailed description of the data selection process is summarized in the Figure 1. The keyword and precise eligibility criteria are provided in the Appendix.

Limitations

The chosen review design yields several limitations. The review has a disproportionate focus on UK experiences. The research assessment exercise and the research excellence framework (REF) were explicitly referred to in 29% of the included texts. This imbalance is a consequence of the review strategy to only include sources written in English. Furthermore, the review is strictly focused on SSH impact and does not include discussions of the impact of individual disciplines. This limitation was drawn to ensure a manageable corpus size and to focus on contributions that seek to capture the diversity of impacts across SSH. Finally, throughout the screening process it became clear that the literature is not situated in a single scientific tradition but draws on various research fields.

A substantial number of the included texts were located well beyond the conventional research evaluation literature. For instance, documenting the impact of research upon policy-making draws on the literature of science advice and evidence-informed policy-making (Court and Young 2006; British Academy 2008; Davies, Nutley and Walter 2008a; LSE Public Policy Group 2011; Nutley, Walter and Davies 2009). Likewise, the literature on the economic impact of research refers to studies of academic entrepreneurship, commercialization, and technology and knowledge transfer (Papagiannidis et al. 2009; Abreu and Grinevich 2014; Etzkowitz 2014). Contributions describing the cultural impact of SSH relate to a very broad discussion of the value of art and culture (Crossick and Kaszynska 2014, 2016). While this diversity made the inclusion and exclusion of texts technically challenging, it also represents a finding in its own right. There simply is no central SSH impact literature.

Despite these limitations, the review includes descriptions of all major frameworks and methods to assess and accelerate impact-oriented scholarship in the humanities and social sciences, including EU-funded projects and national funding and evaluation programmes. The approaches examined demonstrate that there is currently considerable academic and policy interest in understanding

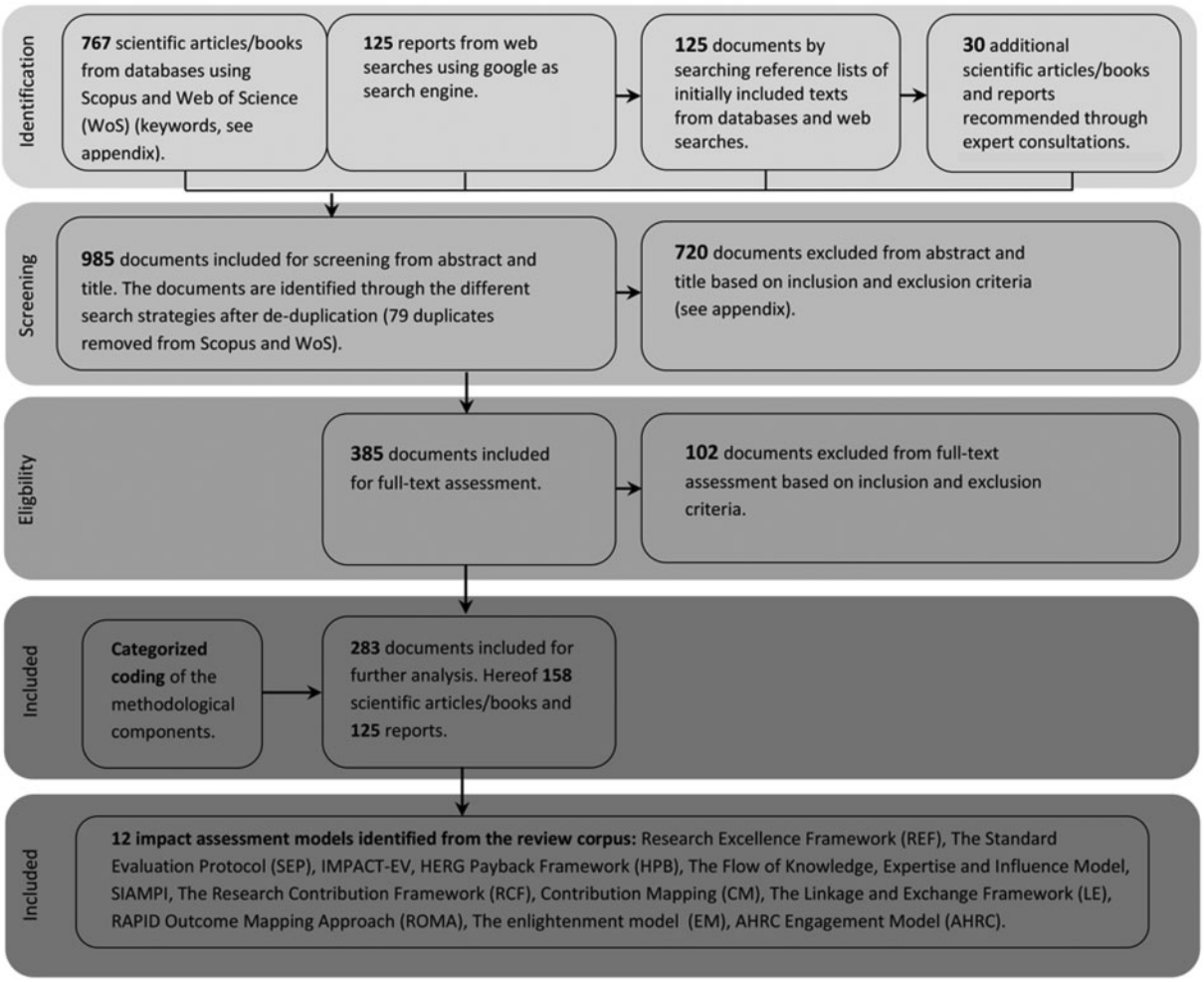

Figure 1. Flowchart of the review and data collection process.[3]

the various ways in which research can contribute to society and the multiple pathways through which impact is realized. Our study does not contain any ready-made recipes for creating or determining impact. Rather, the results highlight the eclectic and complex nature of SSH impact assessment. As such, the scoping review can help support institutions, faculty, and science policy-makers to get a more nuanced picture of the composite nature of SSH impact.

Approaches and instruments to assess the societal impact of SSH

In this section, we present approaches, instruments, and models of impact assessment identified in the literature before we move on to a discussion of their methodological components.

Research excellence framework

REF is a national evaluation system with the aim of assessing the impact and international quality of research carried out at UK universities. REF is conducted every 5 years in co-operation between the four national research policy authorities in England, Scotland, Wales, and Northern Ireland. The primary purpose of the REF is to inform the allocation of quality-related research funding. The Higher Education Funding Council for England, which is responsible for implementing the REF, defines research impact as 'an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia'. REF is based on extended peer review. Assessments are conducted by teams of academics and experts who are assigned to rank research from organizations other than their own. Assessment is carried out within 36 subject-based units of assessment (UOA), such as 'Clinical medicine', 'Law', 'Chemistry', and 'Philosophy'. In REF 2021 the research will be evaluated along three weighted dimensions: research output (60%), impact (25%), and research environment (15%) (REF 2014, 2015a, b, 2019).

The standard evaluation protocol

Standard evaluation protocol (SEP) is used to assess research at universities and other research institutions in the Netherlands. The purpose is to demonstrate and assess the quality and relevance of research and, if necessary, suggest improvements. The evaluation is conducted every sixth year. SEP is based on research units and is carried out by assessment committees to ensure a transparent and independent assessment process.

Panels of research institutions determine which research entities to evaluate. Initially, assessment committees carry out qualitative and quantitative assessments (score of 1–4) based on three overall assessment criteria: quality of research, relevance to society, and viability. Assessment criteria for research quality and social relevance include: (1) detectable products, (2) use of products, and (3) signs of recognition. Assessments of viability are based on SWOT analyses focusing on internal strengths and weaknesses, as well as external possibilities and barriers (KNAW 2010, 2013; VSNU, NWO and KNAW 2015).

The assessment model incorporates a large selection of indicators of both scientific quality and societal relevance identified and selected by the research groups themselves. From an SSH perspective, it emphasizes 'often forgotten' indicators related to teaching and training, which constitute important SSH contributions to both the academic community and broader society. PhD candidates and Master's students play a key role in knowledge dissemination and in applying scientific knowledge in their future work and in society at large (KNAW 2010, 2013).

IMPACT-EV

IMPACT-EV aims to develop a new permanent system for monitoring, selecting, and evaluating impact across SSH disciplines on the basis of quantitative indicators combined with qualitative analyses. As part of the project, IMPACT-EV has proposed a typology of SSH impact pathways. The main aim of the typology is to create a systematic method for analysing impact cases in SSH. Projects are assessed on the basis of the EU 2020 targets. The definition of impact distinguishes between scientific, political, and social impact of research and the impact on the European Research Area (Flecha et al. 2014). IMPACT-EV is funded by the European Commission, and specifically aims to evaluate the transnational impact of research projects sponsored by the Commission.

A clear strength of the model is that it tries to create a multidimensional description of social science and humanities research often absent in purely econometric impact assessments. The model also tries to develop a permanent system on the basis of quantitative indicators, which may be analysed and compared. This distinguishes the evaluation model from other assessment systems such as the REF and SEP that uses peer review to quantify and rank research. However, it is not straightforward to quantify the number of dissemination efforts, interactions and outcomes related to e.g. public policy-making, or use these data as a baseline for calculating and ranking impact. In many cases, only few direct relations and policy outputs exist, though they may be needed in order to influence policy. Thus, the quality of these engagements, interactions and outputs cannot necessarily be quantified.

HERG payback framework

The HERG payback framework is a problem-solving and policy-driven model with a linear representation of knowledge transfer and impact. The model is similar to REF in its focus on outcomes, benefits, and paybacks of research. The HERG payback framework initially focused on health sciences as one of the earliest models to incorporate both academic outputs and societal impact as criteria for assessment. It is one of the most used assessment models across scientific fields (Buxton and Hanney 1996; Buxton 2011).

The model has been modified to evaluate social science research through the Future Work Programme on how researchers inform policy-makers (Klautzer et al. 2011; Wooding et al. 2011) and has been adapted in an assessment of arts and humanities research at the University of Cambridge (Levitt, Celia and Diepeveen 2010). In this process, additional dimensions were added such as research impact on teaching, policy, and practice. Despite these adaptations, the model still faces several limitations in capturing SSH impact. SSH research involves complex, interactive variables, and factors which complicate the direct relations between inputs from research and the outputs generated. Even though the linearity of the model indicates a sequence of interdependent stages through which research progresses, it is important to acknowledge that research from SSH will most often not proceed in this simple, linear way.

The flow of knowledge, expertise, and influence model

The flow of knowledge, expertise, and influence models is an interactive assessment model with a focus on the *processes* leading to real-world research impact. These are represented as non-linear interactions and connections between researchers and users (Meagher, Lyall and Nutley 2008). The framework was developed as part of a research project funded by the UK Economic and Social Research Council, which aimed to investigate new approaches to assess policy and practice impacts (Meagher, Lyall and Nutley 2008).

The model distinguishes between instrumental and more conceptual types of impact. Instrumental impact refers to the use of research for solving specific problems or making decisions. Conceptual impact is a more wide-ranging notion of impact that comprises the complex and often indirect ways in which research influence and shape knowledge, understandings and attitudes of policy-makers and practitioners. The model acknowledges that SSH often improve the efficiency and effectiveness of specific policies or services through a more neutral stance or contribute a critical perspective on public policy in relation to specific agendas and policy trends (Weiss 1995). While such research impact may be less demonstrable—as they are often unapparent—they are no less important than more obvious, instrumental forms of knowledge use (Nutley, Walter and Davies 2007: 36).

Social impact assessment methods for research and funding instruments through the study of productive interactions

Study of productive interactions (SIAMPI) aims to uncover indicators for social impact through a conceptual framework based on the central notion of 'productive interactions' (De Jong et al. 2011; Molas-Gallart and Tang 2011; Spaapen and van Drooge 2011; De Jong et al. 2013). The model is built around a non-linear understanding of social impact as a result of dynamic interactions and joint efforts between several actors inside and outside the academic community. SIAMPI is funded through a research-led project consortium under the European Union Seventh Framework Programme (FP7/2007–2013), and further developed by the Rathenau Institute in the Netherlands and the pan-European League of Research Universities in 2017 (SIAMPI 2011; LERU 2017).

The SIAMPI model is particularly well suited for studying projects in SSH, where research is typically only one component in complex social and political processes where it is inappropriate to seek the direct effect of research. Studies of productive interactions found that SSH projects are quite often informal and decentralized and display unique dynamics and range in contacts with societal actors

(Molas-Gallart and Tang 2011; Olmos-Peñuela, Molas-Gallart and Castro-Martínez 2014b).

The RCF

Research contribution framework (RCF) is an evaluation approach initially used to assess public sector change programmes where several interrelated factors influence behaviour and change. In such cases, RCF has been implemented to help managers, researchers, and policy-makers track the contributions and outcomes of research programmes. RCF is a process-oriented approach that seeks to capture the uptake and use of research at separate stages of impact-creation. Based on a 'theory of change', the model looks at the audit trail that include factors such as the mission of a research programme, contextual factors, value assumptions, and risks involved in processes of change (Mayne 2001, 2008; Morton 2012, 2015).

RCF aims to take into account the many ways social science research is taken up, used, and reused in policy and practice, where it is difficult to link research processes and outputs to wider societal change (Morton 2015). The model is most suitable for accounting for research that has direct links to specific societal organizations, such as partnerships or formalized engagements. The model appears to be less suited for assessing research project characterized by indirect interactions with multiple and diverse target groups.

Contribution mapping

In contribution mapping (CM), research is represented as an ecosystem consisting of complex and unstable networks of people and technologies. Instead of looking at the final impacts of research, CM looks at the activities and alignment efforts of different actors during the research process. CM is a forward-tracking approach with extensive stakeholder involvement. It traces the routes to impact by interviewing researchers and other actors associated with the research project (Kok and Schuit 2012). This model seems appropriate for the engaged scholar and research that is co-produced alongside societal partners on the basis of commonly defined goals and solutions. The model is not suited for research that does not actively incorporate societal actors as part of the research process but instead seek to influence society from a distance through more indirect channels.

The linkage and exchange framework

Linkage and exchange framework (LE) emphasizes personal connections and interactions between researchers and societal actors and institutions. The framework seeks to promote research uptake in specific contexts as well as encourage research that is of use to stakeholders in society. LE has primarily been used by Adrian Cherney and colleagues to explore academic-industry collaborations in the social sciences (Cherney and McGee 2011; Cherney et al. 2013; Cherney 2015). The framework employs a scaling model inspired by empirical studies by Réjean Landry (Knott and Wildavsky 1980; Landry, Amara and Lamari 2001). The framework includes different forms of impact ranging from attempts to translate and communicate research to external partners and determining how research affects choices and decisions and finally give rise to societal applications (Cherney and McGee 2011).

RAPID outcome mapping approach

RAPID outcome mapping approach (ROMA) is developed by the Overseas Development Institute in order for planning international development work and measuring its results. The ROMA approach consists of theoretical work, case studies, and practical implementations. The aim is to identify interrelated factors which determine whether research-based evidence is likely to be adopted by policy-makers and practitioners. The framework is outcome-oriented and seeks to document the progressive changes in attitudes and behaviours of users, beneficiaries, and consumers. By involving external collaborators and stakeholders, it focuses on tracking specific policy change to determine the contributing factors and help identify and prioritize changes (Court and Young 2006; Young et al. 2014).

Conceptual framework

Apart from empirical assessment models, the literature review identified several conceptual frameworks that contain important reflections on impact assessment while not presenting data. The enlightenment model is a conceptual framework developed by Belfiore and Bennett (2010) to describe impact from the Arts. The framework is inspired by Carol Weiss' studies on research utilization in policy that criticizes simple and instrumental understandings of use and impact. Instead, Carol Weiss suggests that impact should be understood as a gradual sedimentation of insights, theories, concepts, and perspectives. She states that different kinds of diffuse and undirected outputs from research may gradually shift political thinking as well as perceptions of social problems and solutions (Weiss 1979, 1980, 1995). Likewise the UK Art and Humanities Research Council has developed a set of comprehensive guidelines for planning and demonstrating policy engagement. The framework suggests incorporating a range of qualitative and quantitative indicators adaptable to describe specific engagements. A basic assumption is that robust evidence of active and systematic engagement provides a good indicator of the impact of research (AHRC 2013). The framework acknowledges that impacts are often due to circumstances beyond the control of research institutions. Similar problems associated with a narrow focus on the realized value and end-impact are discussed by various contributions in the review (e.g. Brewer 2011; Benneworth 2014; Hammersley 2014; Upton, Vallance and Goddard 2014).

Findings

Interestingly, the models identified above extends from policy initiatives (AHRC, HERG) and funding models (REF, SEP) to academic initiatives and research projects occupied with creating a better understanding of the processes of knowledge translation and implementation (e.g. SIAMP, IMPACT-EV, RCF and CM). The dominance of academically driven initiatives should not be overemphasized. The influence of the UK REF is far more significant for UK universities than any other framework identified in the literature. Yet, the prevalence of academically driven initiatives is a testimony that scholars in SSH are increasingly interested in the impact agenda. In the next section, key methodological components of the different frameworks are described in detail.

Methods for SSH research impact assessment

In this section, we review the methods, procedures, and tools used for describing research impact within the different impact frameworks introduced in the previous section. This section provides an overview of methods within the individual frameworks, and gives insight into the evidence base of current impact assessment models

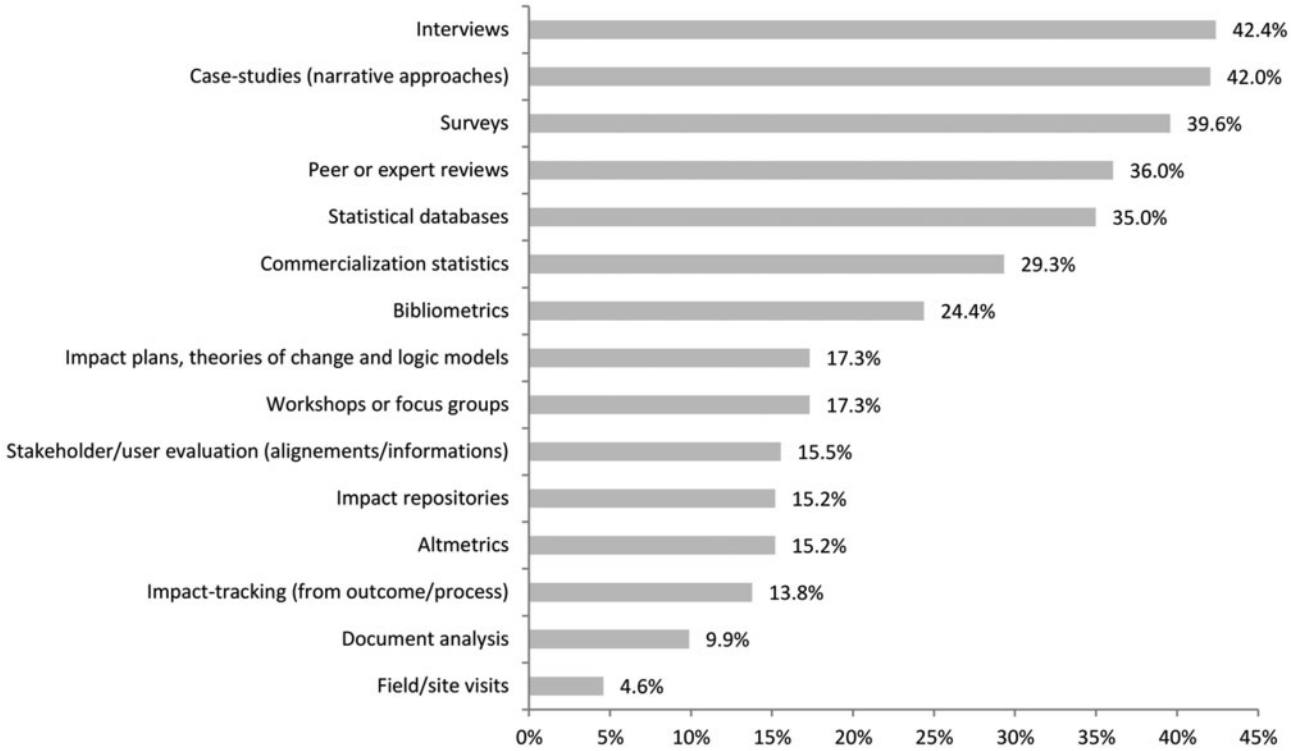

Figure 2. Share of documents mentioning the main methodological components in the literature corpus in percent (n = 283).

in SSH. The data show considerable methodological diversity. This finding indicates that different frameworks are focusing on different *aspects* of impact rather than on universal assessment. While the different methods share a common objective of providing evidence of impact there is considerable variability with regards to the preferred methods—whether it is expert judgement, stakeholder alignment, quantitative indicators, or combinations of these. Figure 2 displays the frequency of the main methodological components found in the literature. Figure 3 shows how different methods and components are distributed across the impact assessment models introduced above.

Interviews

Interviews are mentioned in 42% of the included contributions (120 texts) and is the most preferred method for studying the impact of SSH. Interviews and focus groups are part of several assessment models, and are used to assess first-hand perspectives with research collaboration or knowledge mobilization among partner institutions as an indicator of impact. Interviews typically involve non-academic partners and end-users but can also involve researchers responsible for impact-oriented activities (Cassity and Ang 2006; van Hemert, Nijkamp and Verbraak 2009; Blewden, Carroll and Witten 2010; British Academy 2010, 2014; Thelwall and Delgado 2015). Impact interviews are often highlighted as one of the most useful sources of information (Molas-Gallart and Tang 2011; Morton 2012).[4]

Interviews allow informants to reflect upon critical conditions for creating impact, and interviewers can react and customize questions based on informants' responses. Using a structured interview guide allows comparison between cases and projects, and may uncover motivations, enablers, or concerns related to the creation of impact. Qualitative interviews can be used throughout or even before a research project is initiated to collect information from stakeholders and manage expectations (Donovan and Hanney 2011; SIAMPI 2011; Kok and Schuit 2012; Morton 2012; Cherney 2015).

The literature highlights a number of weaknesses of the method. For instance, informants may not have perfect information about impact or pathways to impact. The timing of the interview, whether at the beginning, during, or after the research has been carried out, is crucial for the reliability and scope of the interview (Bell, Shaw and Boaz 2011). Furthermore, selecting informants may be difficult. There is a risk of overemphasizing impacts that did not originate from research itself. In other instances, it may be necessary to train interviewers to ensure the quality of data. Finally, transcribing, analysing, and comparing data can be time-consuming (Boaz, Fitzpatrick and Shaw 2008, 2009).

Case studies (narrative approaches)

Case studies are mentioned 119 times (42%) in the review corpus and are a central method in almost all assessment frameworks. The method is found to have several advantages. Case studies can deal with a high degree of complexity and provide descriptions of specific pathways that lead to use, uptake, and impact of research in real-world settings. Impact cases are centrally featured in REF and has inspired other national and international impact assessment exercises.[5] Case studies are popular in the SSH literature because they allow researchers and institutions to include a broad range of impact pathways, for example, policy, health, business, and cultural impact, which are not normally accounted for by, for example, econometric, bibliometric, or other data-driven approaches. In contrast, the case methodology gives research units a unique opportunity to produce a

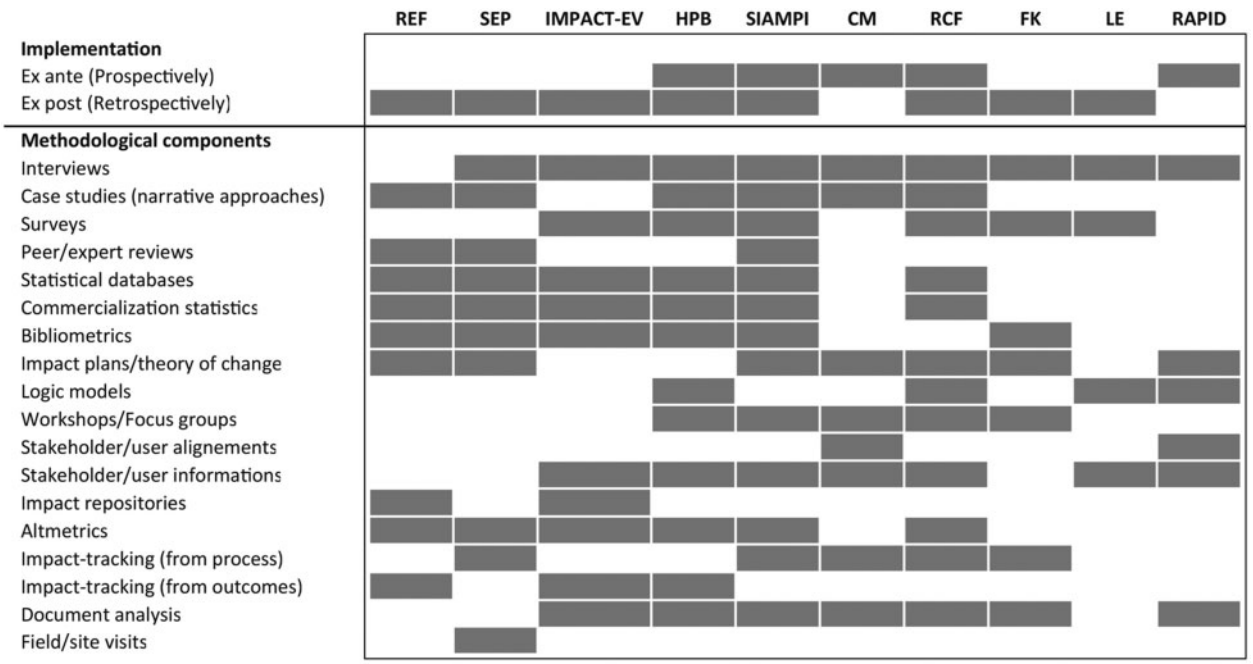

Figure 3. Distribution of the main methodological components across the identified impact assessment models.

coherent narrative, which describes who has benefited from or been influenced by the research. Appropriate evidence to support claims about impact vary significantly among impact assessment frameworks (Boaz, Fitzpatrick and Shaw 2008; Bornmann 2013).

Several criticisms of the method are found in the literature. Most importantly, the method has been criticized for its lack of objectivity as it can be difficult to compare and rank different case studies. Furthermore, the method gives priority to recent research (i.e. research produced within the last assessment cycle) not counting for the long-term impact of research contributions. Another criticism is that not all types of research are able to provide clear and unequivocal evidence for specific impacts. The narrative method puts emphasis on tangible impacts and outstanding examples rather than less visible (yet pervasive) impacts arising from the research process. Impact cases risk idealizing specific pathways and outcomes by not accounting for barriers, set-backs, feedback, and negative effects. Finally, the method is very labour-intensive for both researchers and assessors (Martin 2011).

Surveys

Surveys are mentioned in 40% of the texts in our sample. In the literature, several studies use surveys to explore the impact of SSH (Hughes et al. 2011; Cherney et al. 2013; Abreu and Grinevich 2014; Bastow, Dunleavy and Tinkler 2014; Upton, Vallance and Goddard 2014). These studies show that surveys are useful for collecting data on different variables, such as motivations, perceived barriers and enablers, and different types of engagements between researchers and society. Another advantage is that surveys allow comparative analysis of performance over time and throughout the research process.

The method also has limitations. Surveys primarily provide self-reported evidence of impact which have a bias towards mapping involvement and activities, seen through the lens of the assessment unit, rather than observed change and real-world effects. Another limitation is that surveys assume that research impact can be measured quantitatively and, to some degree, captured by means of standardized questionnaires. However, it is contestable whether dissemination efforts, relationships, and other interactions with, for example, policy-makers count for impact or rather indicate the *likelihood* for impact. In many cases, only few *direct* relationships can be established between research and 'policy change', while other, more *indirect* relations might indicate 'policy influence'. In the end, aggregating the mere number of engagements, interactions, and outputs may not lead to clear indicators of impact but should be supplemented by other methods. As a consequence, surveys often require other types of methods to validate self-reported evidence, for example, qualitative interviews, focus groups, or workshops. Examples of this can be found in the HERG payback framework, RCF, and SIAMPI. Finally, surveys are not very responsive to unforeseen impacts and context-specific factors (Boaz, Fitzpatrick and Shaw 2008, 2009).

Peer/expert review

Peer or expert review is mentioned in 36% of our review texts. Peer review is used as an umbrella term for expertise-based review practices, including the review of journal manuscripts, applications for funding, and hiring and promotion. Peer review is regarded as one of the most important methods for quality assessment across scientific domains (Wilsdon et al. 2015). Typically, quality is assessed by experts belonging to the relevant field, but peer review can also be supported by external indicators. Quality indicators include output indicators (e.g. publications or bibliometric indicators) or indicators of reputation (e.g. prizes, scholarly positions, and additional evidence of recognition) (KNAW 2005). Furthermore, in REF and SEP, the peer/expert review are used to assess both scientific excellence and the societal impact of research based on a case study approach in addition to assessments of scientific quality (REF 2015a; VSNU, NWO and KNAW 2015).

Peer review is a flexible and highly trusted method for research assessment that can be implemented at various stages in the research process. It is generally deemed useful in relation to understanding the broader societal impact of diverse research outputs. It is used for allocation of research funding or for mid-term and final evaluation (Holbrook and Frodeman 2011). Furthermore, it is possible to quantify data and evidence by attributing impact scores. As part of the peer review, feedback, learning, and criticism can be directed towards the scientific product or the strategy for achieving impact (European Science Foundation 2012; Wilsdon et al. 2015).

Because of the pervasive role peer review plays in academic life its use in impact assessment is contested (Derrick 2018). Peer review has been criticized for delivering an 'acceptance threshold' rather than a measure of impact. Peers may have a bias towards rewarding the work of highly recognized scholars in the field, leading to positive judgments and funding for researchers with a high a priori reputation (the so-called Matthew effect). The method is time-consuming and impractical because of the number of experts involved. Impact peer review requires that experts are well-informed and have an in-depth knowledge about the research context. This may be hard to provide (KNAW 2005, 2010; Wilsdon et al. 2015).

Statistical databases

Statistical databases are mentioned in 99 texts (35%). Statistical databases are used to describe research infrastructures, facilities, income, scientific degrees and prices, and may provide data and information from different scientific fields. It is possible to follow developments over time and between research units, and often it is possible to combine data across administrative and statistical records for comparative purposes.

The disadvantages of the method listed in the literature are that administration and user rights can hinder experiments on data and require repetitive renewal of agreements between researchers about joint utilization of established datasets. This can prevent access to data for both scientists and other interested parties. Registries may be inadequate and require ongoing maintenance, documentation, and validation. It may be difficult to document and derive impact from specific projects from administrative and statistical databases alone, especially informal engagement efforts (Martin and Tang 2006; Bekkers and Bodas Freitas 2008; Abreu and Grinevich 2013; Olmos-Peñuela, Molas-Gallart and Castro-Martínez 2014b). On the other hand, several contributions perceive statistical databases to be useful for national evaluation systems where it is important to take into account the relationship between different disciplines, the academic environment, and the academic and societal outputs and contributions (Reale et al. 2017; REF 2014; VSNU, NWO and KNAW 2015).

Commercialization data

In the literature, we found references in 83 texts (29%) to different forms of commercialization data. Commercialization data belongs to a quantitative assessment regime and is frequently used in combination with bibliometrics (university-industry co-authorships, etc.). Commercialization statistics are often narrowly associated with measuring different types of economic effects, primarily in relation to technology transfer and the business sector. Standard indicators used are patents, licences, joint research and development, contracted research, industry funding and spin-outs. Commercialization data can also be used to study various other types of entrepreneurial activities (Martinelli, Meyer and Von Tunzelmann 2008; Abreu and Grinevich 2013; D'Este et al. 2013; Perkmann et al. 2013). Like administrative databases, commercialization data makes it possible to identify formal and contractual relationships between researchers and society.

However, it is often hard to compare different types of commercial impacts—especially across disciplines and contexts. Furthermore, the method and its underlying transaction model may be useful for describing economic and technological impact in the natural and technical sciences, but is generally found to be insufficient for understanding impact in SSH (Bullen, Robb and Kenway 2004; Hughes et al. 2011; Abreu and Grinevich 2013). In SSH, impact relations are established through diverse (non-economic) channels such as policy reports, stakeholder meetings, public lectures, or books written for general audiences. Such are not accounted for in terms of economics (Hughes et al. 2011; Abreu and Grinevich 2014; Hazelkorn 2014; Olmos-Peñuela, Benneworth and Castro-Martinez 2014a). For example, debates concerning innovation in SSH are often oriented towards 'creative', 'social', and 'civic' innovation rather than product innovation (Aalestad et al. 2006; Phipps and Shapson 2009; Etzkowitz 2014; Gulbrandsen and Aanstad 2015; Phipps and Morton 2013).

Considerable criticism of a narrow focus on commercial impact can thus be found in both the academic and policy literature. Several contributions emphasize that commercial exploitation of research always takes place within a broader socio-technical system, and is mediated by values, norms, and practices that are not themselves commercial. This argument, which can be found in several policy reports, highlights the entanglement of economic and non-economic factors in creating impact. Following this line of reasoning, research on, for example, sustainable infrastructures, public service, migration, values, or inter-cultural understanding, may have substantial impact on the economy—although this cannot be accounted for directly through standard economic indicators (Griffin 2006; European Commission 2009; European Science Foundation 2009, 2013; ALLEA 2013; LERU 2013; Science Europe 2013).

Bibliometrics

Bibliometric methods are mentioned in 69 (24%) of the reviewed texts. Bibliometric methods are primarily focused on assessing the academic impact of research (measured in terms of publications, citations, co-authorships, etc.) rather than the societal impact, which is the focus of our review. However, bibliometric data also includes information about research outputs such as datasets, reagents, software, researcher interactions, and funding. Bibliometrics represent a widely accepted statistical method to asses and rank scientific performances, mainly on the basis of individual authors as the unit of analysis (Wouters and Costas 2012). Citations analysis such as H-index allows for studies on whether research is being pursued at the highest level and cited by other researchers in the field (LSE Public Policy Group 2011; Wilsdon et al. 2015). Bibliometric analyses may contribute to the objectivity and transparency of the research evaluation process and provide an overview of publication patterns and scientific networks that cannot be seen from the perspective of the individual researcher. This may provide useful information on how specific research fields are interconnected, growing, or declining over time (KNAW 2005).

However, we found a number of substantial criticisms of bibliometric assessment in the SSH literature. Bibliometrics as a field has primarily been developed as a tool to assess life and natural sciences,

which have historically recorded publication and citations patterns. Coverage of SSH in the main bibliometric databases have traditionally been poor and unable to provide a coherent picture of publications (including monographs) and citations. For instance, several contributions mention the importance of non-journal publications in SSH, as well as the prevalence of non-English language publications, which cannot at this stage be accounted for by using standard bibliometric indicators (KNAW 2005; Nederhof 2006). Adding to the criticism, the notion of 'quality' is perceived by SSH scholars to be field-specific. Comparison between publication or citation patterns may be analytically problematic.

Responses to this line of criticism is generally divided in two: one group of contributions is focused on creating better metrics for SSH and building better publication and citation records, whereas another group criticizes the use of metrics and the ambitions to develop indicators across research fields (Rafols et al. 2012; Zuccala 2012; Wouters et al. 2015). Despite these limitations, bibliometrics are generally considered to be a useful method in the assessment of the quality of research, when used responsibly and cautiously (Wilsdon et al. 2015). However, in the context of SSH, bibliometric indicators cover only a small part of the written communication between researchers, and display no evidence of the broader engagement with wider society (KNAW 2005; Bornmann 2014a). The literature further emphasizes that an exclusive use of bibliometric indicators has performative consequences for the evaluation system as a whole. Performing well on bibliometric measures is not sufficient for making research useful and relevant in a broader societal context (Donovan and Butler 2007; Bornmann 2014b; Hammarfelt 2014; Thelwall and Delgado 2015). Overreliance on bibliometrics can lead to a distorted view of scientific outputs (publications) as more or less an end in itself (Martin 2011; Dahler-Larsen 2012, 2014; Wilsdon et al. 2015). In the different impact assessment frameworks included in the present review, bibliometrics are exclusively used for *academic* impact assessment which do not capture research activities oriented towards broader audiences outside academia (HERG, IMPACT-EV, REF, and SEP) whereas other methods (e.g. altmetrics) are used to deliver evidence of the broader societal uptake of publications and content.

Impact plans theory of change and logic models

Impact plans, templates, and logic models are mentioned in 17% of the reviewed texts. Impact plans may be developed *retrospectively* to describe context, activities, and outputs of research that have influenced society. In REF, for example, impact plans and templates are used to account for the context and evidence underlying the individual case study and provide additional information about the wider range of activities associated with the unit of assessment. This method allows assessment panels to take into account particular circumstances that may have influenced submissions.

Impact plans can also be used *prospectively* to explain how a project is expected to bring about desired results. The literature also includes frameworks (ROMA, RCF) which use impact plans in the preparation phase based on a 'theory of change'. This approach can be useful for determining expected contributions of projects. This method makes it possible to examine how different factors and assumptions affect research uptake by revisiting and reformulating assumptions and indicators throughout the lifecycle of a project. A theory of change ideally functions as a road map, which guides those engaged in the project. This can help ensure that researchers and relevant collaborators keep track of project outcomes throughout the research process. A theory of change requires that the project team is able to provide ex-ante descriptions of its activities, key partners, and societal effects. However, for highly innovative and explorative projects it may be difficult to determine expected changes at the planning phase (Mayne 2008). Instead, researchers may strive to maintain an open mind and acknowledge the iterative nature of the research, where the impact of research cannot necessarily be established at the outset. Typically, theories of change are not used to describe specific causal links between specific research activities and societal changes. Instead, the 'theory' is altered and refined during the research process providing a learning opportunity for the involved partners (Meagher, Lyall and Nutley 2008; Spaapen and van Drooge 2011; Kok and Schuit 2012; Morton 2012; Young et al. 2014). A 'theory of change' requires an explicit understanding of the mission of research, the context and interest of external actors, and the specific activities linking research to society. Only if these factors are included in a coherent framework is it possible to account for the impact of SSH.

Several impact assessment models (e.g. HERG) combine a logic model (to depict input-activities) with impact plans, templates, and case studies. The purpose of this combined methodology is to capture the complex processes and interactions through which knowledge is produced, taken up, and used in society. Logic models can be useful tools for tracking specific outcomes resulting from research (Morton 2012; Young et al. 2014) as well as they can be used to break down different types of outcome associated with the research (Donovan and Hanney 2011). In both cases, sensitivity towards indirect impacts and links between research and society is required.

Workshops and focus groups

Workshops and focus groups are mentioned in 49 texts (17%) in the review corpus. These methods can be described as an organized form of discussion that involves researchers, partners, and other stakeholders. The method is applied in different phases of an assessment and research process. Workshops may be used early on to bring researchers and relevant stakeholders together in order to co-produce ideas and set up shared objectives (Eerd et al. 2011). As a method to accelerate and assess research impact, workshops can encourage the mobilization of research by brining attention to specific projects or programmes among external audiences. Workshops can also serve as a platform to co-develop dissemination strategies, and explore potential sources of impact evidence (Nutley 2003). Workshops and focus groups are relatively cost-effective ways to increase the influence of a research project. Workshops and focus groups can also be applied after a research project has ended in order to capture the immediate effects of research in relation to specific partner institutions or target groups. On the downside, the literature identifies risks that participants in focus groups may not describe all positive or negative effects of a research project. Participants are primed to present positive results and may have selective memory when reconsidering potential impacts and risks (Boaz, Fitzpatrick and Shaw 2008, 2009).

Stakeholder/user evaluation

Stakeholder evaluation is mentioned in 44 texts (16%). In some cases, the literature discusses how stakeholders are part of the research project as partners and collaborators participating in the co-production of knowledge (Antonacopoulou 2010; Armstrong and

Alsop 2010; Martin 2010; Orr and Bennett 2012; Campbell and Vanderhoven 2016). In such cases, stakeholders play an integral role in developing plans for how the research is organized, implemented, and assessed. Including stakeholders throughout the research process may also lead to the co-production of evaluative indicators. In many instances, researchers and stakeholders have little or no influence on impact indicators. Taking a stakeholder approach to impact assessment may foster wider participation and engagement in the research process.

Stakeholders may also serve as important informants for learning about how research is taken up and used. Methods such as surveys, interviews, workshops, and focus groups may serve to generate insight into partner organizations and help improve awareness, understanding, and communication amongst researchers and stakeholders as (Eerd et al. 2011; O'Brien, Marzano and White 2013). Involvement of stakeholders can happen at the design phase in order to orient research towards specific needs and goals (Kok and Schuit 2012; Cherney 2015). This may help break down barriers and cultural differences between researches and potential partners and beneficiaries. Stakeholder involvement may also be implemented during the research process to gather information of preliminary benefits (Molas-Gallart and Tang 2011; Young et al. 2014; Morton 2015).

According to several contributions, there is a risk that partners may become too involved in the research, and hence compromise the integrity and academic freedom of the research units. In many situations, it is not always possible to identify users or beneficiaries of research, and user-involvement can be a costly and time-consuming method during all phases (Boaz, Fitzpatrick and Shaw 2009). Yet, gathering information from users is a common method in large parts of the reviewed literature, and is in FK, RCF, and HERG. Stakeholder involvement and alignment in the evaluation or design of indicators is less common (RAPID and CM). In RAPID, stakeholder involvement is used to develop tools and guidelines for shared understanding of objectives and requirements (Court and Young 2006; Young et al. 2014). In CM, potential end-users and other informants are involved throughout the research process to explore possible contributions. This may include training staff, adapting organizational practices, establishing relation with policy-makers, recruiting participants, and so on. At later stages, researchers and related actors may further investigate how they can disseminate the produced knowledge and stimulate utilization and knowledge absorption directly through, for example, personal interactions, presentations, or publications (Kok and Schuit 2012).

Impact repositories

In the review, we found 43 mentions (15%) of impact repositories and databases, especially in relation to the policy literature (LERU 2013; OECD 2013; Flecha et al. 2014; HERA 2014; European Commission 2015). Databases and repositories are part of the emerging Open Science agenda in which academic and policy actors are focusing on opening up research for collaboration and knowledge sharing. As part of REF, an impact database has been created in which all submitted impact case studies are listed (REF 2012). AHRC has developed a portfolio with over 100 impact case studies in an ongoing effort to create an evidence base for research impact. The European Commission has promoted initiatives such as OpenAIRE and CORDIS to support impact. OpenAIRE is a data collection for measuring impact of Horizon2020 projects, whereas CORDIS functions as the primary public repository for dissemination of information on all EU-funded research (European Commission 2015).

Databases make it possible to locate individuals (peers within universities and collaborators outside academia) linked to specific research projects. They are also used to share research results and data to facilitate a broader dissemination across projects and fields. Repositories facilitate explorative studies of impact data, and help provide comprehensive and nuanced representations of research impact across fields (King's College London and Digital Science 2015). The disadvantages of repositories mainly pertain to the fact they require researchers to invest time in documenting and describing their impacts and pathways. There are also ethical considerations in relation to sharing and utilizing data, for instance related to securing sensitive information.

The generation of different forms of impact repositories is considered an important component in many of the impact assessment frameworks. In the literature, we find a strong focus on the broader communication and dissemination of research through the establishment of repositories. Often knowledge translation, knowledge exchange, and knowledge mobilization involves explicit research brokerage, which primarily benefit the immediate partners (Knight and Lightowler 2010; Lightowler and Knight 2013; Sebba 2013). In contrast, impact repositories may stimulate wider interest in research and make data about research accessible to a wider audience of practitioners.

Altmetrics

In the literature, we locate 43 contributions (15%) which mention alternative metrics or methods (altmetrics) for assessing scientific impact—most often based on scientometrics (Bornmann 2014b; Hammarfelt 2014; Waltman and Costas 2014; Zahedi, Costas and Wouters 2014; Bornmann, Haunschild and Marx 2015; Bornmann and Marx 2014). While much attention has been devoted to text-based outputs such as peer-reviewed publications, researchers in SSH produce a wide variety of outputs which are not normally taken into account (Hazelkorn 2014; Wilsdon et al. 2015). Altmetrics aims at covering broader outputs by tracking communication and sharing of research through social media and digital platforms such as Twitter, Facebook, blogs, and digital sharing services, for example, Mendeley, Cite U Like, Altmetric.com, or Impact Story. Recent attempts to improve altmetrics include the integration of non-traditional outputs such as policy reports, white papers, or manuals. Different forms of data can be collected from these sources, ranging from citations, views, downloads, clicks, tweets, shares, likes, bookmarks, and comments beyond traditional bibliometrics (Bornmann 2014a; Hammarfelt 2014; Holmberg and Thelwall 2014, Zahedi, Costas and Wouters 2014). The main advantage of altmetrics is the ability to harvest big data making it possible to quantify wider dissemination efforts and effects on scholarly conversations in the digital sphere. Altmetrics is used to track broader societal outputs, including media presence, attention, and other indicators (Mohammadi and Thelwall 2014). Altmetrics methods work well as supplements to case studies and narrative approaches.

Altmetric has several limitations. Data are not easily compared between different disciplines, or even across different research themes and topics (the problem of normalization). In addition, altmetric data are often biased towards specific users and it is rarely possible to access precise user-statistics or samples for different types of social media. Furthermore, it is unclear what a mention or

download indicates. In contrast to citation analysis, references and citations in policy reports or social media are less standardized. Research used in such contexts is not always cited and not everything cited is actually used (Bornmann and Daniel 2008; Neylon, Willmers and King 2014). Therefore, scholars recommend a reflexive and responsible approach to altmetrics, acknowledging that it cannot be used as a universal method for assessing societal impact at this stage (Hicks et al. 2015; Wilsdon et al. 2015).

Impact-tracking and activity registration

Process-tracking is found in 39 texts (14%). It is a technique for tracing impact pathways, either *forward* from initial research towards research use, outcomes, and impacts, or *backwards* from identified outcomes to specific research initiatives and outputs (Meagher, Lyall and Nutley 2008). A backward-tracking approach is used by HERG to account for specific outcomes of research (Buxton and Hanney 1996; Wooding et al. 2011). The method has several advantages. It traces return on investment from a variety of empirical sources. This makes it possible to uncover 'how' as well as 'why' specific outcomes or impacts have succeeded. The methods allow research projects to be measured against goals set by research institutions or funding agencies themselves. On the other hand, key drawbacks are that backward-tracking tends to rely on the quality of and access to relevant documentation. Further difficulties relate to the 'attribution' of outcomes to specific research results—especially when accounting for indirect pathways to impact (Boaz, Fitzpatrick and Shaw 2009).

SIAMPI and RCF make use of forward-tracking approaches. These frameworks seek to identify links and productive interactions that may lead to socially relevant applications. SIAMPI finds that 'it is only by analysing the processes that induce social impact that we have a chance of recognizing potential research impacts and the contributions made by research that might otherwise not be evident' (Spaapen and van Drooge 2011: 213). RCF tracks relations and pathways based on a theory of change linking different factors that may influence use, uptake, and outcomes (Mayne 2008; Morton 2012). Forward-tracking techniques can help establish a connection between (1) research objectives, framing, and design, (2) the research processes and outputs, and (3) outcome and impact. It is also able to identify barriers and enablers of achieving research impact.

Process-tracking can be based on qualitative (e.g. interviews, impact logs) as well as quantitative approaches (e.g. social network analysis, geo-referencing, contextual response analysis etc.). However, it is difficult to systematically describe ways to achieve impact during, or even at the end of a project. For example, impact assessment studies using the SIAMPI model found that information on productive interactions were hampered by the lack of systematic data. Researchers often claimed importance for specific forms of interactions, while no independent data on these interactions are available for assessment (SIAMPI 2011).

Review and analysis of documents

Document analyses mentioned in 28 texts (10%) cover the review and interpretation of existing documents such as books, policy reports, white papers, grey literature, etc. Review of documents can be used both qualitatively and quantitatively in combination with computational text analysis (e.g. text mining, topic models, semantic text analysis, etc.) or traditional coding strategies (e.g. categorized coding, thematic syntheses, etc.). The method provides an understanding of content and context of specific outputs. It depends heavily on the quality of existing outputs and the ability to find and collect them systematically and says little about the non-written outputs from research. Furthermore, no single method can be easily adapted and used across the board, but many different methodological strategies may be required, which may complicate study design. Consequently, it often requires extensive expertise and time to adapt chosen methods.

Field/site visits

Field visits is a method mentioned in 13 (5%) texts from the corpus. As part of the national evaluation system in the Netherlands, the approach is used to assess research at universities and other research institutions. The purpose is to demonstrate the quality and societal relevance of research (VSNU, NWO and KNAW 2015). Field visits, can give evaluators a chance to observe research on the ground and talk to principal investigators and other staff members about their experience, plans, and strategies. Field visits can generate valuable insights, especially when triangulated with data from other sources. However, the reliability of qualitative data can vary, and the method consequently faces the same challenges as other qualitative measures, such as workshops and interviews.

Summary

In this section, we have introduced and discussed a number of the most prominent methods for assessing the impact of SSH. Findings from the scoping review suggest that the field of SSH impact assessment is characterized by pluralism, and that different models, theories, and frameworks make use of different methods and assumptions. Which specific method is applied to document impact is inevitably value-laden. Different organizations tend to focus on different levels of impact, and adapt methods suitable for tracking that. The literature review also suggests that the contributions of SSH to society are mediated by numerous factors, and that it is unlikely or even impossible to use universal or catch-all indicators. Instead, methods need to be developed for given contexts and used in combination with other types of assessment. Our review also identified that methods for documenting research impact appear to be established both in policy and academic contributions. Impact itself is indeed emerging as a key concept of attention across universities, funding agencies, as well as in peer-reviewed articles and grey literature.

Main distinctions

In this section, we discuss how the models and methodological techniques presented in the previous sections point in different directions, and how the main axes of agreement and disagreement found in the literature may be used to identify specific challenges and questions for research on research impact in the future.

From linear to multi-dynamic and cyclical assessment

First of all, the discussed frameworks all present multi-dynamic or cyclical approaches to impact assessment. Apart from The HERG payback framework, no model in the review defines impact as a 'linear' process—and even HERG acknowledges that processes of knowledge transferral are anything but simple. Rather, it is described how research is 'embedded in networks' (SIAMPI 2011); is 'dynamic and complex'; and involves 'two-way processes between

research, policy, and practice' (Court and Young 2006, Kok and Schuit 2012, Morton 2012). Research processes often resemble knowledge ecosystems rather than linear relationships (British Academy 2008). A key finding of the literature review therefore is that societal impact does not occur in a social or cultural vacuum, but is realized in a network of interacting actors, interests, values, and institutions.

A move from single methods to mix-methods approaches

Due to the complexity and diversity of research impacts, different assessment models incorporate diverse and flexible methods to describe impact and its pathways. Quantitative methods (such as citation analysis and commercialization data) and qualitative methods (such as case studies, interviews and field visits) need to be used reflexively and preferably in combination. As we have shown in previous sections, different methods each have their advantages and disadvantages. For example, bibliometric indicators may be useful for assessing academic impact. But bibliometrics are unfit to account for the broader communication and engagement of research in society. Hence, an important finding from our study is that there is no one-size-fits-all model for measuring and assessing impact. The methodological strategies chosen need to be tailored towards specific goals, for example, the mission of specific research projects and funding programmes.

Incorporation of ex-ante, in media res and ex-post drivers

The methods presented in this article are not only measuring impact *after* research have been carried out (ex-post), but are occupied with tracing and accelerating outputs and activities during the research process (in media res) and even at the planning stage before research is initiated (ex-ante). Impact plans, and theory of change, can be used ex-ante to clarify the mission, identify potential partners, and develop strategies for broader dissemination and user-involvements. Similarly, workshops, focus groups, and interviews are useful to create stakeholder alignment for specific types of projects by directing research towards societal needs at the planning and implementation phase.

In addition to ex-ante assessment, several frameworks pay attention to the different benefits obtained throughout the research process. In several contributions, SSH activities and contributions are intimately linked and created in collaboration and co-creation with actors outside the research community. Impact assessments therefore need to take into account the contributions of external partners and stakeholders which are not directly linked to the research project but constitute the wider social web of relations necessary for realizing the impact of research. Linking research with societal institutions may lead to the development of proxy indicators of connectivity to show how research is actually being co-produced, implemented, and exploited for specific purposes. Yet, this also highlights the 'attribution problem' found in the literature, that is, that specific impacts occur as the result of multiple stakeholders working together and providing contributions rather than only (or primarily) academic institutions driving the impact uptake chain.

Case-based vs. metric-based approaches

Due to technological advancements, a number of digital methods and indicators are currently being developed to create alternative indicators of research impact. These methods should in principle be less labour-intensive and require less manual assessment. Metric-based methods may contribute to more sophisticated and automated techniques for research evaluation and provide quantitative (big) data for comparing UOA. However, scientometrics and altmetrics have been challenged for a number of reasons. Using scientometric indicators as a silver bullet for measuring societal impact may give an inaccurate understanding of research performance which is analytically narrow and unfair to some types of research. Indicators may be arbitrary, and impact assessments risk being led by data and evidence rather than what is significant and interesting in real-world cases. This is the pitfall of letting only what is countable count.

Nevertheless, the literature review identifies several contributions that develop principles and guidelines for the responsible use of metrics (Hicks et al. 2015; Wilsdon et al. 2015). In the Leiden Manifesto, Hicks et al. (2015) describe 10 principles to guide research evaluation. They argue that metrics should always protect locally relevant research and be able to account for variation in specific research fields against the mission of research. Assessments need to be open and transparent, and quantitative evaluations should always support qualitative and expert assessment. Finally, narrow indicators may lead to unintended consequences in terms of adaptive (or performative) behaviour (Dahler-Larsen 2012, 2014). A narrow focus on short-term impacts may, for instance, create perverse incentives to direct research towards easily attainable results while not supporting the long-term indirect uptake of knowledge.

Case-based methodologies may be more suitable (either as stand-alone or in combination with metrics-based approaches) to account for broader and more nuanced forms of impact, especially cultural and policy impact where there are no coherent quantifiable data available. In contrast to metrics-based approaches, impact case studies do not only account for fixed artefacts (such as texts, citations, patents, etc.) but can include descriptions of dynamic impacts unfolding from a multiplicity of sources. Translations, artworks, public debates, curricula, reports, performances, exhibitions, etc. can be included as evidence to support the impact case. Likewise, combinations of impact repositories, process-tracking, metrics, and case studies may give a much more coherent understanding of research impact than is currently available. Being able to account for the number of activities, interactions, and relations any unit of assessment has to the external environment (e.g. number of meetings, interviews, media appearance, workshops, etc.) may create valuable evidence used in impact case studies. Combinations of qualitative and quantitative methods may include descriptions of relationships between research and society, which are less codified but highly relevant to understand real-world changes (Pilegaard, Moroz and Neergaard 2010; Hughes et al. 2011; Spaapen and van Drooge 2011; Abreu and Grinevich 2014; Olmos-Peñuela, Molas-Gallart and Castro-Martínez 2014b). However, a comprehensive case-based assessment system risks becoming a very expensive and burdensome (Martin 2011).

Positive vs. negative impact

A final important distinction relates to the balancing of positive and negative impacts. Several contributions highlight that impact of research can be contested and that research results can be co-opted by external stakeholders and used for purposes beyond the control of academic scholars. Brewer (2011) speaks of 'disguised' impact which occurs when the results of research are rejected—not because

the research is useless or irrelevant but because it is in conflict with societal values. For instance, research on vulnerable groups may be highly useful and relevant and at the same time rejected because of conflicting value systems within the dominant political discourse.

Hammersley (2014) further describes situations in which the desire to create impact does not lead to the expected positive changes. For example, behavioural psychology or economic theory can be used to justify very different social policies. This observation shows that impact assessments will differ according to norms and interests. Bozeman and Sarewitz (2011) argue that a basic social theory for science is needed to inform science policy on whether research serves specific public values (Bozeman and Sarewitz 2011). They argue that value assumptions are helpful when framing impact and evaluation frameworks, and that explicit attention to values will help inform a more tangible and contextualized approach to mapping the societal impact of research. At the macro-level, research policies should move beyond a mere 'marked failure model' and instead include social goals and public values to determine 'social outcomes and roles of research as part of the web of institutions, networks, and groups giving rise to social impacts' (Bozeman 2003: 12). A similar focus on values is introduced by Nussbaum (2010) who highlights the need to reflect upon the cognitive and deliberative capabilities that research is bringing about. She argues that a narrow concern with national economic growth will neglect important initiatives in research and teaching, such as the promotion of critical thinking, sympathy towards marginalized groups, and competence in dealing with complex global problems (Nussbaum 2010). To avoid a narrow perspective of research that produce easy-to-measure impact, assessments require a 'public value theory' which determine desirable and undesirable effects of research (Hammersley 2014).

This last point is pivotal to SSH disciplines. Key SSH fields such as science and technology studies, gender studies, area studies, and policy studies often notice how science and innovation do not necessarily lead to social progress. Research impact may also cause harm—sometimes purposefully, as with nuclear weapons, sometimes accidentally, as with asbestos, thalidomide, or opioids (Derrick et al. 2018). Often, there is uncertainty and disagreement regarding what is desirable, which generally calls for a responsible approach to impact assessment (Budtz Pedersen, Grønvad and Hvidtfeldt 2017). One way of bridging the value gap when designing future impact assessments is to include both scholars and external partners in the creation of indicators. External partners may have significant contributors throughout the entire research process including in the evaluation design (Ross et al. 2003).

Conclusion

In this review article, we have presented various ways in which the impact of SSH is conceptualized and modelled. We have demonstrated a widespread interest across the academic and policy literature in the assessment and communication of SSH impacts. The SSH are—like other scholarly fields—committed to demonstrating the vital role of research in society, in terms of the production and dissemination of knowledge, and through sustained engagement with non-academic stakeholders.

Our findings confirm prior research on research impact, which states that the mechanisms through which research lead to socially valuable knowledge are deeply complex. Different impact frameworks and definitions focus on different *aspects* of impact ranging from academic impact to policy impact, social impact, educational impact, cultural impact, and economic impact. Our findings also demonstrate that there is considerably room for social values, missions, and objectives to enter into the assessment cycle. Because of the complexity and open-ended nature of impact assessment in SSH, it is to a large extent up to universities, funding agencies, and national academies to define how and with which methods individual organizations and ecosystems assess the impact of the humanities and social sciences.

Beyond researchers and research organizations, stakeholders, partners, and collaborators play an important role in realizing the impact of SSH. Models such as knowledge mobilization, knowledge brokering, and co-creation share a commitment to create stronger alignment among academic and non-academic institutions. Yet the literature review finds that some types of research may require more steps and long-term capacity-building before the impact is realized, or may not lead to benefits or products at all (e.g. Oancea and Furlong 2007; Spaapen and van Drooge 2011; Thorpe et al. 2011; Kok and Schuit 2012, Ellwood, Thorpe and Coleman 2013; Morton 2015; Budtz Pedersen 2016). An increased awareness of the entire knowledge translation chain, and the capacities needed to absorb research among user communities, may in the future create a better appreciation of the various types of research impact and the resources needed to realize them.

The available literature is rich and varied but also contains certain deficits. Notably, most contributions to the current literature are *conceptual* (and often speculative). While the conceptual interest in research impact pervades the academic and policy literature, it is surprising that no more empirical studies (especially among academic contributions) are found. Demonstrating the viability of impact assessment models in practice rather than in theory is an important next step when developing SSH impact studies. Future research in this field should provide a more substantial understanding of the nature and breadth of impacts, and the literature would benefit from contributions addressing and experimenting with new data and in-depth studies of engagement, networks, products, and processes leading to research uptake. Future research in the field should also address the question of responsibility and attribution. We have noted that knowledge exchange and knowledge brokering is not a task only for researchers but include a shared responsibility across academic and non-academic partners. For this co-responsibility to be acknowledged, assessing SSH impacts should be based on (1) contributions of research to societal impact instead of attributing societal impact to specific research projects, (2) it should be based on a clear understanding of values and scales (interests, temporality, locality, etc.), and finally (3) it should be based on an acknowledgement that successful impact and innovation takes place in a complex interplay between actors of many kinds.

Notes

1. When considering the impact that is generated as a result of research, a number of authors and government recommendations have advised that a clear definition of impact is required (Duryea, Hochman and Parfitt 2007; Klautzer et al. 2011; Russell Group 2012). For a recent discussion, see Alla et al. (2017).
2. The relationship between academic impact and societal impact is often described as a hierarchy, in which academic impact (excellence) is concerned defined by the quality and validity of

research, whereas societal impact refers to the broader engagement, uptake, and relevance of research to society. In this basic distinction only research that is found to be scientifically sound by peers should be judged by its wider relevance to society.

3. From the text corpus (n = 283), 10 unique impact assessment models were identified in addition to two more conceptual and strategic frameworks. The empirical assessment models were broken down into their main methodological components. A categorized coding were performed in order to structure the methodological discussion of each individual methodological component. The coding was conducted manually by reading the documents.
4. As Norman Denzin notes: "This may be because many qualitative researchers don't have data and findings, tables and charts, statistics and numbers. We have stories, narratives, excerpts from interviews" (Denzin 2009: 151).
5. The UK Research Excellence Framework (REF) introduced impact cases as a methodology in 2014 in the form of short (two pages) descriptions that document how a concrete outcome is related to a specific research product such as a publication. This method has later spread to other countries, including Norway where it has been used in at least four large-scale evaluations of research both among research institutes and larger fields including universities and colleges.

Acknowledgements

We appreciate the valuable input and commentary that the scoping review has received in various stages. In particular, we would like to thank Sharon Smit, Peter Meister-Broekema, Peter Dobers, Jonas Stier, Frederik Stjernfelt, Gunnar Sivertsen, Jack Spaapen, and feedback from the journal's anonymous reviewers. We are grateful for financial support received from the European Union's Horizon 2020 programme (grant no. 693477), Velux Foundation (grant no. 342044).

References

Aalestad, P., Acham, K., Caughie, J., Moeran, B., and Holm, J. (2006) *The Humanities as a Source of Creativity and Innovation*. Brussels: Humanities in the European Research Area (HERA).

Abreu, M., and Grinevich, V. (2013) 'The Nature of Academic Entrepreneurship in the UK: Widening the Focus on Entrepreneurial Activities', *Research Policy*, 42: 408–22.

Abreu, M., and Grinevich, V. (2014) 'Academic Entrepreneurship in the Creative Arts', *Environment and Planning C: Government and Policy*, 32: 451–70.

AHRC (2013) *Guidance on Planning and Demonstrating Effective Policy Engagements*. UK: AHRC.

Alla, K., Hall, W. D., Whiteford, H. A., Head, B. W. and Meurk, C. S. (2017) 'How Do We Define the Policy Impact of Public Health Research? A Systematic Review', *Health Research Policy and Systems*, 15: 84.

ALLEA (2013) *Embedding the Social Sciences and Humanities in Horizon 2020. All European Academies, ALLEA*. Berlin: ALLEA.

Antonacopoulou, E. P. (2010) 'Beyond Co-production: Practice-Relevant Scholarship as a Foundation for Delivering Impact through Powerful Ideas', *Public Money & Management*, 30: 219–26.

Arksey, H., and O'Malley, L. (2005) 'Scoping Studies: Towards a Methodological Framework', *International Journal of Social Research Methodology*, 8: 19–32.

Armstrong, F., and Alsop, A. (2010) 'Debate: Co-production Can Contribute to Research Impact in the Social Sciences', *Public Money & Management*, 30: 208–10.

Bastow, S., Dunleavy, P., and Tinkler, J. (2014) *The Impact of the Social Sciences*. London, UK: Sage.

Bekkers, R., and Bodas Freitas, M. I. (2008) 'Analysing Knowledge Transfer Channels between Universities and Industry: To What Degree Do Sectors Also Matter?', *Research Policy*, 37: 1837–53.

Belfiore, E., and Bennett, O. (2010) 'Beyond the "Toolkit Approach": Arts Impact Evaluation Research and the Realities of Cultural Policy-Making', *Journal for Cultural Research*, 14: 37–41.

Bell, S., Shaw, B., and Boaz, A. (2011) 'Real-World Approaches to Assessing the Impact of Environmental Research on Policy', *Research Evaluation* 20: 227–37.

Benneworth, P. (2014) 'Tracing How Arts and Humanities Research Translates, Circulates and Consolidates in Society. How Have Scholars Been Reacting to Diverse Impact and Public Value Agendas?', *Arts and Humanities in Higher Education*, 14: 45–60.

Blewden, M., Carroll, P., and Witten, D. K. (2010) 'The Use of Social Science Research to Inform Policy Development: Case Studies from Recent Immigration Policy', *Kōtuitui: New Zealand Journal of Social Sciences Online*, 5: 13–25.

Boaz, A., Fitzpatrick, S., and Shaw, B. (2008) *Assessing the Impact of Research on Policy: A Review of the Literature for a Project on Bridging Research and Policy through Outcome Evaluation*. London: Kings College London & Policy Studies Institute.

Boaz, A., Fitzpatrick, S., and Shaw, B. (2009) 'Assessing the Impact of Research on Policy: A Literature Review', *Science and Public Policy*, 36: 255–70.

Bornmann, L. (2013) 'What is Societal Impact of Research and How Can It Be Assessed? A Literature Survey', *Journal of the American Society of Information Science and Technology*, 64: 217–33.

Bornmann, L. (2014a) 'Do Altmetrics Point to the Broader Impact of Research? An Overview of Benefits and Disadvantages of Altmetrics', *Journal of Informetrics*, 8: 895–903.

Bornmann, L. (2014b) 'Validity of Altmetrics Data for Measuring Societal Impact: A Study Using Data from Altmetric and F1000Prime', *Journal of Informetrics*, 8: 935–50.

Bornmann, L., and Daniel, H. D. (2008) 'What Do Citation Counts Measure? A Review of Studies on Citing Behavior', *Journal of Documentation*, 64: 45–80.

Bornmann, L., Haunschild, R., and Marx, W. (2015) 'Policy Documents as Sources for Measuring Societal Impact: How is Climate Change Research Perceived in Policy Documents?', *Scientometrics*, 109: 1477–95.

Bornmann, L., and Marx, W. (2014) 'How Should the Societal Impact of Research Be Generated and Measured? A Proposal for a Simple and Practical Approach to Allow Interdisciplinary Comparisons', *Scientometrics*, 98: 211–19.

Bozeman, B. (2003) 'Public Value Mapping of Science Outcomes: Theory and Method', Knowledge Flows and Knowledge Collectives: Understanding the Role of Science and Technology Policies in Development. Working Paper. <http://citeseerx.ist.psu.edu/> accessed 13 Oct 2018.

Bozeman, B., and Sarewitz, D. (2011) 'Public Value Mapping and Science Policy Evaluation', *Minerva*, 49: 1–23.

Brewer, J. D. (2011) 'The Impact of Impact', *Research Evaluation*, 20: 255–6.

British Academy (2008) *Punching Our Weight: The Humanities and Social Sciences in Public Policy Making*. London: British Academy.

British Academy (2010) *Past, Present and Future: The Public Value of the Humanities and Social Sciences*. London: British Academy.

British Academy (2014) *Prospering Wisely: How the Humanities and Social Sciences Enrich Our Lives*. London: British Academy.

Buchanan, A. (2013) 'Impact and knowledge Mobilisation: What I Have Learnt as Chair of the Economic and Social Research Council Evaluation Committee', *Contemporary Social Science*, 8: 176–90.

Budtz Pedersen, D. (2016) 'Integrating Social Sciences and Humanities in Interdisciplinary Research', *Palgrave Communications*, 2: 1–7.

Budtz Pedersen, D., Grønvad, J., and Hvidtfeldt, R. (2017) *Responsible Impact Manifesto* <http://react.aau.dk> accessed 26 Nov 2018.

Bullen, E., Robb, S., and Kenway, J. (2004) 'Creative Destruction": Knowledge Economy Policy and the Future of the Arts and Humanities in the Academy', *Journal of Education Policy*, 19: 3–22.

Buxton, M. (2011) 'The Payback of "Payback": Challenges in Assessing Research Impact', *Research Evaluation*, 20: 259–60.

Buxton, M., and Hanney, S. (1996) 'How Can Payback from Health Services Research Be Assessed?', *Journal of Health Service Research and Policy*, 1: 35–43.

Campbell, H., and Vanderhoven, D. (2016) *Knowledge That Matters: Realising the Potential of Co-production*. UK: N8 Research Partnership and ESRC.

Cassity, E., and Ang, I. (2006) 'Humanities-Industry Partnerships and the "Knowledge Society": the Australian Experience', *Minerva*, 44: 47–63.

Cherney, A. (2015) 'Academic-Industry Collaborations and Knowledge co-Production in the Social Sciences', *Journal of Sociology*, 51: 1003–16.

Cherney, A., Head, B., Boreham, P., Povey, J., and Ferguson, M. (2013) 'Research Utilization in the Social Sciences: A Comparison of Five Academic Disciplines in Australia', *Science Communication*, 35: 780–809.

Cherney, A., and McGee, T. R. (2011) 'Utilization of Social Science Research: Results of a Pilot Study among Australian Sociologists and Criminologists', *Journal of Sociology*, 47: 144–62.

Court, J., and Young, J. (2006) 'Bridging Research and Policy in International Development: An Analytical and Practical Framework', *Development in Practice*, 16: 85–90.

Crossick, G., and Kaszynska, P. (2014) 'Under Construction: Towards a Framework for Cultural Value', *Cultural Trends*, 23: 120–31.

Crossick, G., and Kaszynska, P. (2016) *Understanding the Value of Arts & Culture: The AHRC Cultural Value Project. Arts and Humanities Research Council*. UK: AHRC.

Dahler-Larsen, P. (2012) *The Evaluation Society*. Stanford CA: Stanford Business Books.

Dahler-Larsen, P. (2014) 'Constitutive Effects of Performance Indicators', *Public Management Review*, 16: 969–86.

Daudt, H. M. L., van Mossel, C., and Scott, S. J. (2013) 'Enhancing the Scoping Study Methodology: A Large, Inter-Professional Team's Experience with Arksey and O'malley's Framework', *BMC Medical Research Methodology*, 13: 48.

Davies, H., Nutley, S., and Walter, I. (2008) 'Why "Knowledge Transfer" is Misconceived for Applied Social Research', *Journal of Health Services Research and Policy*, 13: 188–90.

D'Este, P., Tang, P., Mahdi, S., Neely, A., and Sánchez-Barrioluengo, M. (2013) 'The Pursuit of Academic Excellence and Business Engagement: Is It Irreconcilable?', *Scientometrics*, 95: 481–502.

De Jong, S., Barker, K., Cox, D., Sveinsdottir, T., and Van den Besselaar, P. (2013) *Understanding Societal Impact through Studying Productive Interactions*. Den Haag: Rathenau Institute.

De Jong, S. P. L., van Arensbergen, P., Daemen, F., van der Meulen, B., and van den Besselaar, P. (2011) 'Evaluation of Research in Context: An Approach and Two Cases', *Research Evaluation*, 20: 61–72.

Denzin, N. K. (2009) 'The Elephant in the Living Room: Or Extending the Conversation about the Politics of Evidence', *Qualitative Research*, 9: 139–60.

Derrick, G. (2018) *The Evaluators' Eye: Impact Assessment and Academic Peer Review*. London/New York: Springer.

Derrick, G. E., Faria, R., Benneworth, P., Budtz Pedersen, D., and Sivertsen, G. (2018) 'Towards characterising negative impact: Introducing Grimpact', *23rd International Conference on Science and Technology Indicators (STI 2018)*, September 12–14, 2018. Leiden, The Netherlands: Centre for Science and Technology Studies (CWTS).

Derrick, G., and Samuel, G. (2017) 'The Future of Societal Impact Assessment Using Peer Review: Pre-Evaluation Training, Consensus Building and Inter-Reviewer Reliability', *Palgrave Communications*, 3: 17040.

Donovan, C., and Butler, L. (2007) 'Testing Novel Quantitative Indicators of Research "Quality", Esteem and "User Engagement": an Economics Pilot Study', *Research Evaluation*, 16: 231–42.

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

Duryea, M., Hochman, M., and Parfitt, A. (2007) 'Measuring the Impact of Research', *Research Global*, 1: 8–9.

Eerd, D. V., Cole, D., Keown, K., Irvin, E., Kramer, D., Gibson, J. B., Kohn, M., Mahood, Q., Slak, T., Amick, B., Phipps, D., Garcia, J., and Morassaei, S. (2011) *Report on Knowledge Transfer and Exchange Practices: A Systematic Review of the Quality and Types of Instruments Used to Assess KTE Implementation and Impact*. Toronto, ON: Institute for Work & Health.

Ellwood, P., Thorpe, R., and Coleman, C. (2013) 'A Model for Knowledge Mobilisation and Implications for the Education of Social Researchers', *Contemporary Social Science*, 8: 191–206.

Etzkowitz, H. (2014) 'Making a Humanities Town: Knowledge-Infused Clusters, Civic Entrepreneurship and Civil Society in Local Innovation Systems', *Triple Helix*, 2: 1–22.

European Commission (2009) *Emerging Trends in Socio-Economic Sciences and Humanities in Europe. The METRIS Report (Advisory Report)*. Bruxelles: European Commission.

European Commission (2015) *Impact of Open Science methods and practices on the economics of research and science. Case Studies from Life, Mathematical and Social Sciences*. Brussels: European Commission.

European Science Foundation (2009) *Vital Questions. The Contributions of European Social Science. SCSS Science Position Paper.* France: European Science Foundation, ESF.

European Science Foundation (2012) *The Challenges of Impact Assessment*. France: European Science Foundation, ESF.

European Science Foundation (2013) *Cultural Literacy in Europe Today. Science Policy Briefing. European Science Foundation, ESF and COST*, 48. France: European Science Foundation, ESF.

Flecha, R., Soler, M., Oliver, E., Puigvert, L., Sordé, T., Schubert, A., Donovan, C. (2014) *Impact Evaluation of FP6 (Last Call) and FP7 SSH Research Projects. Report 3*. Amsterdam: IMPACT-EV.

Greenhalgh, T., Raftery, J., Hanney, S., and Glover, M. (2016) 'Research Impact: A Narrative Review', *BMC Medicine*, 14: 78.

Griffin, G. (2006) 'Balancing Agendas Social. Social Sciences and Humanities in Europe', *Arts and Humanities in Higher Education*, 5: 291–304.

Gulbrandsen, M., and Aanstad, S. (2015) 'Is Innovation a Useful Concept for Arts and Humanities Research?', *Arts and Humanities in Higher Education*, 14: 9–24.

Hammarfelt, B. (2014) 'Using Altmetrics for Assessing Research Impact in the Humanities', *Scientometrics*, 101: 1419–30.

Hammersley, M. (2014) 'The Perils of "Impact" for Academic Social Science', *Contemporary Social Science*, 9: 345–55.

Hazelkorn, E. (2014) 'Making an Impact: New Directions for Arts and Humanities Research', *Arts and Humanities in Higher Education*, 14: 25–44.

HERA. (2014) *HERA Joint Research Programme: Outcomes and Achievements. "Cultural Dynamics: Inheritance and Identity" and "Humanities as a Source of Creativity and Innovation"*. Brussels: European Commission. Humanities in the European Research Area, HERA.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., and Rafols, I. (2015) 'The Leiden Manifesto for Research Metrics', *Nature*, 520: 429–31.

Holbrook, J. B. (2017) 'The Future of the Impact Agenda Depends on the Revaluation of Academic Freedom', *Palgrave Communications*, 3: 1–8.

Holbrook, J. B., and Frodeman, R. (2011) 'Peer Review and the Ex Ante Assessment of Societal Impacts', *Research Evaluation*, 20: 239–46.

Holmberg, K., and Thelwall, M. (2014) 'Disciplinary Differences in Twitter Scholarly Communication', *Scientometrics*, 101: 1027–42.

Hughes, A., Kitson, M., Probert, J., Bullock, A., and Milner, I. (2011) *Hidden Connections: Knowledge Exchange between the Arts and Humanities and the Private, Public and Third Sectors*. UK: AHRC.

King's 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*. London: Higher Education Funding Council of England (HEFCE).

Klautzer, L., Hanney, S., Nason, E., Rubin, J., Grant, J., and Wooding, S. (2011) 'Assessing Policy and Practice Impacts of Social Science Research: The Application of the Payback Framework to Assess the Future of Work Programme', *Research Evaluation*, 20: 201–9.

KNAW (2005) *Judging Research on Its Merits: An Advisory Report by the Council for the Humanities and the Social Sciences Council*. Amsterdam: Royal Netherlands Academy of Arts and Sciences.

KNAW (2010) *Quality Indicators for Research in the Humanities (Advisory Report)*. Amsterdam: Royal Netherlands Academy of Arts and Sciences.

KNAW (2013) *Towards a Framework for the Quality Assessment of Social Science Research (Advisory Report)*. Amsterdam: Royal Netherlands Academy of Arts and Sciences.

Knight, C., and Lightowler, C. (2010) 'Reflections of "Knowledge Exchange Professionals" in the Social Sciences: Emerging Opportunities and Challenges for University-Based Knowledge Brokers', *Evidence & Policy*, 6: 543–56.

Knott, J., and Wildavsky, A. (1980) 'If Dissemination is the Solution, What is the Problem?', *Knowledge: Creation, Diffusion, Utilization*, 1: 537–78.

Kok, M. O. M. O., and Schuit, A. J. A. J. (2012) 'Contribution Mapping: A Method for Mapping the Contribution of Research to Enhance Its Impact', *Health Research Policy and Systems*, 10: 21.

Landry, R., Amara, N., and Lamari, M. (2001) 'Climbing the Ladder of Research Utilization: Evidence from Social Science Research', *Science Communication*, 22: 396–422.

Langfeldt, L., and Scordato, L. (2015) 'Assessing the Broader Impacts of Research: A Review of Methods and Practices', NIFU Working Paper 2015-8. Norwegian Institute for Innovation and Research Studies (NIFU).

LERU (2013) *The Future of the Social Sciences and Humanities in Europe: Collected LERU Papers on the SSH Research Agenda*. EU: League of European Research Universities, LERU.

LERU (2017) *Productive Interactions: Societal Impact of Academic Research in the Knowledge Society*. EU: League of European Research Universities, LERU.

Levitt, R., Celia, C., and Diepeveen, S. (2010) 'Assessing the Impact of Arts and Humanities Research at the University of Cambridge', Technical Report. Cambridge: RAND Corporation.

Lightowler, C., and Knight, C. (2013) 'Sustaining Knowledge Exchange and Research Impact in the Social Sciences and Humanities: Investing in Knowledge Broker Roles in UK Universities', *Evidence & Policy*, 9: 317–34.

LSE Public Policy Group (2011) *Maximizing the Impacts of Your Research: A Handbook for Social Scientists*. London: LSE Public Policy Group.

Martin, S. (2010) 'Co-Production of Social Research: Strategies for Engaged Scholarship', *Public Money & Management*, 30: 211–18.

Martin, B. R. (2011) 'The Research Excellence Framework and the "Impact Agenda": Are we Creating a Frankenstein Monster?', *Research Evaluation*, 20: 247–54.

Martin, B. R., and Tang, P. (2006) *The Benefits from Publicly Funded Research. SPRU Electronic Working Paper Series Paper.* UK: SPRU – Science and Technology Policy Research, University of Sussex.

Martinelli, A., Meyer, M., and Von Tunzelmann, N. (2008) 'Becoming an Entrepreneurial University? A Case Study of Knowledge Exchange Relationships and Faculty Attitudes in a Medium-Sized, Research-Oriented University', *The Journal of Technology Transfer*, 33: 259–83.

Mayne, J. (2008) 'Contribution Analysis: An Approach to Exploring Cause and Effect', ILAC Brief 16 *Evaluation*, 1–4.

Mayne, J. (2001) 'Addressing Attribution through Contribution Analysis: Using Performance Measures Sensibly', *The Canadian Journal of Program Evaluation*, 16: 1–24.

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: 163–73.

Mohammadi, E., and Thelwall, M. (2014) 'Mendeley Readership Altmetrics for the Social Sciences and Humanities: Research Evaluation and Knowledge Flows', *International Review of Research in Open and Distance Learning*, 14: 90–103.

Molas-Gallart, J., and Tang, P. (2011) 'Tracing "Productive Interactions" to Identify Social Impacts: An Example from the Social Sciences', *Research Evaluation*, 20: 219–26.

Morton, S. (2012) *Exploring and Assessing Social Research Impact: A Case Study of a Research Partnership's Impacts*. UK: The University of Edinburgh.

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

Nederhof, A. J. (2006) 'Bibliometric Monitoring of Research Performance in the Social Sciences and the Humanities: A Review', *Scientometrics*, 66: 81–100.

Neylon, C., Willmers, M., and King, T. (2014) *Rethinking Impact: Applying Altmetrics to Southern African Research*. Ottawa, Canada: International Development Research Centre.

Nussbaum, M. C. (2010) *Not for Profit: Why Democracy Needs the Humanities*. Princeton, NJ: Princeton University Press.

Nutley, S. (2003) *Increasing Research Impact: Early Reflections from the ESRC Evidence Network*. UK: ESRC Centre for Evidence Based Policy and Practice.

Nutley, S., Walter, I., and Davies, H. (2007) *Using Evidence: How Research Can Inform Public Services*. Bristol, Policy Press.

Nutley, S., Walter, I., and Davies, H. T. (2009) 'Promoting Evidence-based Practice: Models and Mechanisms from Cross-sector Review', *Research on Social Work Practice*, 19: 552–9.

O'Brien, L., Marzano, M., and White, R. M. (2013) 'Participatory Interdisciplinarity": Towards the Integration of Disciplinary Diversity with Stakeholder Engagement for New Models of Knowledge Production', *Science and Public Policy*, 40: 51–61.

Oancea, A., and Furlong, J. (2007) 'Expressions of Excellence and the Assessment of Applied and Practice-Based Research', *Research Papers in Education*, 22: 119–37.

OECD (2013) *New Data for Understanding the Human Condition: International Perspectives*. Paris: OECD.

Olmos-Peñuela, J., Benneworth, P., and Castro-Martinez, E. (2014a) 'Are Sciences Essential and Humanities Elective? Disentangling Competing Claims for Humanities' Research Public Value', *Arts and Humanities in Higher Education*, 14: 61–78.

Olmos-Peñuela, J., Molas-Gallart, J., and Castro-Martínez, E. (2014b) 'Informal Collaborations between Social Sciences and Humanities Researchers and Non-Academic Partners', *Science and Public Policy*, 41: 493–506.

Orr, K., and Bennett, M. (2012) 'Public Administration Scholarship and the Politics of Coproducing Academic–practitioner Research', *Public Administration Review*, 72: 487–95.

Papagiannidis, S., Li, F., Etzkowitz, H., and Clouser, M (2009) 'Entrepreneurial Networks: A Triple Helix Approach for Brokering Human and Social Capital', *Journal of International Entrepreneurship*, 7: 215–35.

Penfield, T., Baker, M. J., Scoble, R., and Wykes, M. C. (2014) 'Assessment, Evaluations, and Definitions of Research Impact: A Review', *Research Evaluation*, 23: 21–32.

Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Broström, A., D'Este, P., Fini, R., Geuna, A., Grimaldi, R., Hughes, A., Kitson, M., Krabel, S., Llerena, P., Lissoni, F., Salter, A., and Sobrero, M. (2013) 'Academic Engagement and Commercialisation: A Review of the Literature on University-Industry Relations', *Research Policy*, 42: 423–42.

Phipps, D. J., and Shapson, S. (2009) 'Knowledge Mobilisation Builds Local Research Collaborations for Social Innovation', *Evidence and Policy*, 5: 211–27.

Phipps, D., and Morton, S. (2013) 'Qualities of Knowledge Brokers: Reflections from Practice', *Evidence & Policy*, 9: 255–65.

Pilegaard, M., Moroz, P. W., and Neergaard, H. (2010) 'An Auto-Ethnographic Perspective on Academic Entrepreneurship: Implications for Research in the Social Sciences and Humanities', *Academy of Management Perspectives*, 24: 46–61.

Rafols, I. (2017) *There's No Silver Bullet for Measuring Societal Impact*. Research Europe. <http://www.researchresearch.com/news/article/?articleId=1370687> accessed 12 Oct 2017.

Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P. and Stirling, A. (2012) 'How Journal Rankings Can Suppress Interdisciplinary Research: A Comparison between Innovation Studies and Business & Management', *Research Policy*, 41: 1262–82.

Ravenscroft, J., Liakata, M., Clare, A. and Duma, D (2017) 'Measuring Scientific Impact beyond Academia: An Assessment of Existing Impact Metrics and Proposed Improvements', *PLoS One*, 12: e0173152.

Reale, E., Avramov, D., Canhial, D., Donovan, D., Flecha, R., Holm, P., Larkin, C., Lepori, B., Mosoni-Fried, J., Oliver, E., Primeri, E., Puigvert, L., Scharnhorst, A., Schubert, A., Soler, M., Soòs, S., Sordé, T., Travis, C., and Van Horik, R. (2017) 'A Review of Literature on Evaluating the Scientific, Social and Political Impact of Social Sciences and Humanities Research', *Research Evaluation*, 27: 298–308.

REF (2012) *Panel Criteria and Working Methods*. UK: REF2014.

REF (2014) *Excellence Framework 2014: The Results*. UK: REF2014.

REF (2015a) *Research Excellence Framework 2014: Overview Report by Main Panel D and Sub-Panels 27 to 36*. UK: REF2014.

REF (2015b) *Research Excellence Framework 2014: Overview Report by Main Panel C and Sub-Panels 16 to 26*. UK: REF2014.

REF (2019) *REF 2021: Guidance on Submissions*. HEFCE. <https://www.ref.ac.uk> accessed 1 Dec 2019.

Rip, A. (2000) 'Higher Forms of Nonsense', *European Review*, 8: 467–85.

Ross, S., Lavis, J., Rodriguez, C., Woodside, J, and Denis, J. L. (2003) 'Partnership Experiences: Involving Decision-Makers in the Research Process', *Journal of Health Services & Policy*, 8(Suppl 2): 26–34.

Russell Group (2012) *The Social Impact of Research Conducted in Russell Group Universities*. Russell Group Papers - Issue 3, London.

Science Europe (2013) *Embedding Social Sciences and Humanities in the Horizon 2020 Societal Challenges. Science Europe Position Statement*. Brussels: Science Europe.

Sebba, J. (2013) 'Mediators in Social Science', *Evidence & Policy*, 9: 391–409.

SIAMPI (2011) *SIAMPI Final Report: Social Impact Assessment Methods for Research and Funding Instruments through the Study of Productive Interactions between Science and Society*. Amsterdam: SIAMPI.

Spaapen, J., and van Drooge, L. (2011) 'Introducing "Productive Interactions" in Social Impact Assessment', *Research Evaluation*, 20: 211–18.

Thelwall, M., and Delgado, M. M. (2015) 'Arts and Humanities Research Evaluation: No Metrics Please, Just Data', *Journal of Documentation*, 71: 817–33.

Thorpe, R., Eden, C., Bessant, J., and Ellwood, P. (2011) 'Rigour, Relevance and Reward: Introducing the Knowledge Translation Value-Chain', *British Journal of Management*, 22: 420–31.

Upton, S., Vallance, P., and Goddard, J. (2014) 'From Outcomes to Process: Evidence for a New Approach to Research Impact Assessment', *Research Evaluation*, 23: 352–65.

Van Hemert, P., Nijkamp, P., and Verbraak, J. (2009) 'Evaluating Social Science and Humanities Knowledge Production: An Exploratory Analysis of Dynamics in Science Systems', *Innovation: The European Journal of Social Science Research*, 22: 443–64.

VSNU, NWO, and KNAW (2015) *Protocol for Research Assessments in the Netherlands*. Amsterdam: Association of Universities in the Netherlands.

Waltman, L., and Costas, R. (2014) 'F1000 Recommendations as a Potential New Data Source for Research Evaluation: A Comparison with Citations', *Journal of the Association for Information Science and Technology*, 65: 433–45.

Weiss, C. (1979) 'The Many Meanings of Research Utilization', *Administration Review*, 39: 426–31.

Weiss, C. (1980) 'Knowledge Creep and Decision Accretion', *Knowledge: Creation, Diffusion, Utilisation*, 1: 381–404.

Weiss, C. (1995) *The Haphazard Connection: Social Science and Public Policy*. Ithaca, NY: Cornell University Press.

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., and Johnson, B. (2015) *Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. UK: Higher Education Funding Council for England, HEFCE.

Wooding, S., Nason, E., Klautzer, L., Rubin, J., Hanney, S., and Grant, J (2011) *Policy and Practice Impacts of Research Funded by the Economic and Social Research Council. A Case Study of the Future of Work Programme, Approach and Analysis*. EU: RAND Europe.

Wouters, P., and Costas, R. (2012) *Users, Narcissism and Control—Tracking the Impact of Scholarly Publications in the 21st Century*. Utrecht: SURF foundation.

Wouters, P., Thelwall, M., Kousha, K., Waltman, L., de Rijcke, S., Rushforth, A., and Franssen, T. (2015) *The Metric Tide Literature Review: Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management*. UK: Higher Education Funding Council for England, HEFCE.

Young, J., Shaxson, L., Jones, H., Hearn, S., Datta, A., and Cassidy, C. (2014) *ROMA—A Guide to Policy Engagement and Influence*. London: Overseas Development Institute (ODI).

Zahedi, Z., Costas, R., and Wouters, P. (2014) 'How Well Developed Are Altmetrics? A Cross-Disciplinary Analysis of the Presence of "Alternative Metrics" in Scientific Publications', *Scientometrics*, 1491–513.

Zuccala, A. (2012) 'Quality and Influence in Literary Work Evaluating the "Educated Imagination', *Research Evaluation*, 21: 229–41.

Appendix

Table A.1. Inclusion and exclusion criteria used in the review

	Inclusion criteria	Exclusion criteria
Publication types	Published peer-reviewed articles, published peer-reviewed books, published reports	Working documents, conference papers, newspaper articles, websites, or blogs
Linguistics	Written in English	Written in other languages
Time period	Published between 2005 and 2016	Published before 2005 Published after 2016
Authors	Reports published by a national or supranational research organization in the European Research Area	Reports published by a national or supranational research organization outside the European Research Area
Population	Include social science and/or humanities research as a collective research field	Only include individual research disciplines in the social science and/or humanities Only include research from scientific fields other than social science and/or humanities
Study characteristics	Developing and/or discussing conceptual models, methods, or tools for understanding, measuring, or driving impact	Not developing and/or discussing conceptual models, methods, or tools for understanding, measuring, or driving impact
	Addressing impact of research towards the wider society and actors outside the academic community (e.g. the limitations of bibliometric indicators, expert review for broader impact assessment)	Addressing impact of research towards the academic community only (e.g. bibliometric indicators of academic performances, peer review concerning quality assessment of scientific work)

Table A.2. Keywords used for searches in scientific databases and web searches

impact AND (cultural OR social OR societal OR health OR public OR policy OR political OR broader OR wider OR economic) AND (social science OR humanities)
knowlegde AND (transfer OR mobilization OR exchange OR circulation) AND (social science OR humanities)
impact AND (evaluation OR assessment OR valorization) AND (social science OR humanities)
co-production AND (social science OR humanities)
co-creation AND (social science OR humanities)
academic AND entrepreneurship AND (social science OR humanities)