Research Evaluation, 30(3), 2021, 323–335
doi: 10.1093/reseval/rvab002
Advance Access Publication Date: 11 April 2021
Article

OXFORD

The production of scientific and societal value in research evaluation: a review of societal impact assessment methods

Jorrit P. Smit [1,2,*] **and Laurens K. Hessels** [3,4]

[1]Department of Public Administration, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands, [2]Institute for Philosophy, Leiden University, Nonnensteeg 1-3, 2311 BE Leiden, The Netherlands, [3]Rathenau Instituut, Anna van Saksenlaan 51, 2593 HW, The Hague, The Netherlands and [4]CWTS, Leiden University, Kolffpad 1, 2333 BN Leiden, The Netherlands

*Corresponding author. Email: j.p.smit@essb.eur.nl

Abstract

Over the past two decades, several methods have been developed to evaluate the societal impact of research. Compared to the practical development of the field, the conceptual development is relatively weak. This review article contributes to the latter by elucidating the theoretical aspects of the dominant methods for evaluating societal impact of research, in particular, their presuppositions about the relationship between scientific and societal value of research. We analyse 10 approaches to the assessment of the societal impact of research from a constructivist perspective. The methods represent different understandings of knowledge exchange, which can be understood in terms of linear, cyclical, and co-production models. In addition, the evaluation methods use a variety of concepts for the societal value of research, which suggest different relationships with scientific value. While some methods rely on a clear and explicit distinction between the two types of value, other methods, in particular Evaluative Inquiry, ASIRPA, Contribution Mapping, Public Value Mapping, and SIAMPI, consider the mechanisms for producing societal value integral to the research process. We conclude that evaluation methods must balance between demarcating societal value as a separate performance indicator for practical purposes and doing justice to the (constructivist) science studies' findings about the integration of scientific and societal value of research. Our analytic comparison of assessment methods can assist research evaluators in the conscious and responsible selection of an approach that fits with the object under evaluation. As evaluation actively shapes knowledge production, it is important not to use oversimplified concepts of societal value.

Key words: research evaluation; impact assessment; societal impact; literature review

1. Introduction

Today, evaluation, audits, and accountability are widespread in society (Power 2000; Dahler-Larsen 2011). Evaluating productivity and impact has also become an integral part of scientific practice at all levels—from individuals and research groups to departments, faculties, and universities, and from grants and funding programmes to entire disciplinary fields (Wilsdon 2016). Over the past two decades, and in response to practical needs of science policymakers, research funding bodies, and university administrators, science studies scholars have developed a number of methods to evaluate the societal value of research. Several reviews have compared the functionality of these different methods for evaluating societal value (Bornmann 2013; Penfield et al. 2014; Miettinen, Tuunainen, and Esko 2015; Greenhalgh et al. 2016). Most literature about impact evaluation is user-oriented and driven by user needs. Lacking, however, is critical engagement with the methods and their policy context from a theoretical point of view (Donovan 2019; Muhonen, Benneworth, and Olmos-Peñuela 2020; Thomas et al. 2020; Williams 2020). This is urgent

because evaluation methods are not passive instruments but actively steer what counts as good, real, and relevant research (de Rijcke et al. 2016). As instances of valuation, research evaluations produce concepts of value that influence decisions about the research lines of individuals and research groups.

In this article, we analyse the theoretical aspects of 10 societal impact evaluation methods to understand how evaluation differentiates between the scientific and societal value of research. The performative nature of evaluation, and the outpacing of theory by fast developments in the practice of research evaluation, motivates us to scrutinize the conceptual presuppositions of these methods. This critical review aims to contribute to the required theoretical reflection on impact assessment, which can inform future evaluation practices.

We probe the conceptual reach and limits of these evaluation methods with a framework that is informed by constructivist studies of scientific practice. While Williams (2020) has recently proposed a sociological theory of power to describe research impact assessment, we understand societal value in a more constructivist way. This perspective will draw attention both to the interactions between academic researchers and societal actors and to the role allotted to *evaluators* in knowledge production. The focus on the practice of research abandons an upfront distinction between the scientific value and the societal value of research. Instead, a constructivist perspective suggests that a difference between the scientific and societal value of research has to be actively produced. We assume that evaluation practices have visibly contributed to such differentiation. An important purpose of evaluating societal value of research is to emphasize and illustrate the contribution that research activities can make to economic progress, societal well-being, or other public goods distinct from an, arguably more internal, epistemic contribution. This explains why methods for evaluating societal value typically position this type of value explicitly as a category separate from scientific value, which is subject to a longer tradition of evaluation (Wouters 1999).

The central research question for this article is*: how is the relationship between scientific and societal value of academic research understood and operationalized in different impact evaluation methods?* In Section 2, we discuss science studies literature about the relationship between societal and scientific value. In Section 3, we present our analytical framework and its three critical aspects: actors, exchange mechanisms, and the concept of societal value. Background and characteristics of the 10 evaluation methods are introduced in Section 4. Subsequently, in Section 5, we analyse and compare the methods with respect to the theoretical aspects of actors, mechanisms, and concepts. Ultimately, we will reflect on the possibilities for evaluation practices to balance practical (policy) demands for evaluation with the theoretical understanding of research practice.

2. Theoretical background

In this reflective study, we explore the theoretical assumptions underpinning a set of societal impact evaluation methods. Our perspective builds on a variety of historical, sociological, and philosophical studies that have described scientific research as a sociomaterial practice, embedded in networks of human and non-human actors. In this view, the production of scientific knowledge coevolves with the establishment of relations with non-academic actors. This implies that scientific and societal value of research are strongly related.

2.1 Societal and scientific value of research

The most common term to refer to evaluation methods addressing societal benefits or societal value of science is probably societal *impact* evaluation. In line with this, we will use the term societal impact evaluation to refer to our general object of study, but in our analysis of their theoretical presuppositions, we will focus in particular on the concept of societal *value*. While impact suggests a limited focus on (intended) changes in behaviour or practices, we regard the concept of societal value as a more open and inclusive concept, which can also refer to the appreciation of particular research outcomes that do not involve tangible effects, such as the cultural value of a better understanding of social phenomena. Informed by the approach in valuation studies that takes value itself as a social construct (Lamont 2012), we do not work with a precise circumscription of societal value but rather explore the way research evaluations construct the value of research activities in relation to a variety of actors in society.

Constructivist science studies have understood research primarily in terms of practices and networks, avoiding an a priori distinction between knowledge producers and knowledge users. Knowledge production takes place in translation networks, in which scientific claims start as 'fictions' that only develop the status of 'truth' if they receive sufficient interest from others (Latour 1987; Stengers 1997). Any new scientific claim can be stabilized by gathering a wide variety of allied 'actants'—including texts, devices, skills, institutions, and humans, from researchers and technicians to industrialists, politicians and activists (Callon 1997). Stabilizing a claim, or making it reliable, thus requires extension of the network. A strict boundary between the inside ('producers') and outside ('users') of research cannot be easily drawn in this view. Rather, it pleads for an ontologically open stance towards the actors that potentially contribute to the process of knowledge production.

The diversity of actors involved in knowledge production blurs the distinction between the activities of research and its distribution 'outside of the laboratory'. Current science policies aiming to improve the societal value of research typically concern themselves with the relations between scientific and societal actors. From a constructivist perspective, these relations appear integral to the network that sustains the primary research process. This implies that one cannot distinguish between active producers and passive recipients of knowledge and that the circulation of knowledge is an interactive process, integral to research as practice. This is of course in conflict with the value-free ideal of research and the linear model of innovation. Ideals of a pure and value-free science suggest that it is possible and necessary to separate the 'disinterested pursuit of truth' from economic, political, or moral interests and actors. The linear model adds that knowledge exchange runs only from fundamental research to society, eventually causing technological change and economic growth. Many analysts have shown that both perspectives are academic and rhetoric ideals rather than practical realities (Proctor 1991; Edgerton 2004; Godin 2006; Douglas 2014). For example, research categories like 'basic' and 'applied', or 'mode-1' and 'mode-2', do not really describe different methodologies but rather mirror political issues with respect to the practical organization of research (Godin 1998; Shinn 2002; Hessels and van Lente 2008; Kaldewey and Schauz 2018).

Constructivist science studies instead understand fundamental research itself as a value-laden practice that is the result of long chains of translations, flowing in both directions between science and society. In this perspective, the reliability or value of scientific claims partly depends on the heterogeneity and extension of the translation network. This view might induce the relativist reproach that this makes the credibility of knowledge dependent on the views and fads of industry, politicians and society (Sismondo 2017; Lynch 2017). The issue at stake, however, is not a choice between scientific *or* societal value. The constructivist premise is instead that scientific and societal values rely on similar actor networks. The process and results of evaluation could contribute to the understanding of which relations enable, prohibit, or blind the orientation and exchange of research to society.

2.2 Evaluation of research

Evaluations, assessments, and indicators have become widespread in the academic world in the last three decades as a part of a broader trend of accountability in public administration (Dahler-Larsen 2011). Compared to the evaluation of the economic and scientific value of research, evaluating the societal value of research is a recent phenomenon. Economists have modelled and measured the economic value of scientific research from the 1950s onwards, to account for observed growth rates of productivity (Godin and Doré 2004). The scientific value of science has been rendered evaluable by the field of scientometrics, which emerged in the 1960s, and focused on the circulation of knowledge within the scientific community (Wouters 1999). Scientometrics has strongly contributed to the production of a concept of scientific value linked to citation scores of articles, journals, and ultimately, individuals. Evaluations of the societal value of research have been developed in rudimentary forms since the 1980s in North-American and European public R&D funding programmes and have existed since 2000 as a systematic practice in various countries (Sand and Toulemonde 1993; Bozeman and Sarewitz 2011; de Jong, Smit, and van Drooge 2016; Smith et al. 2020). Today, the societal value of research is particularly prominent in science policies of a selection of high-income, European and North-American countries.

In response to the increasing demand from policymakers and society for (evidence of) valuable research, many reviews of the various assessment methods have appeared that aim to provide 'the basis for the development of robust and reliable methods of societal impact assessment' (Bornmann 2013). Reliable evidence of societal value matters to the research community to 'justify expenditure, showcase our work and inform future funding decisions' (Penfield et al. 2014). In this body of review literature, two contradictory types of problem formulation can be distinguished. Ironically, most review articles problematize either the lack of standardization or the lack of heterogeneity in evaluation methods. Several reviewers observe that there exists no shared accepted framework for evaluation of societal value and that routine capture of data is (therefore) also lacking (Penfield et al. 2014; Miettinen, Tuunainen, and Esko 2015). Some explain this by reference to the relative 'infancy' of the research field, compared to scientometrics (Bornmann 2013; De Silva and Vance 2017). Other reviewers stress instead that existing evaluation practices of societal value lack heterogeneity. Many of them disapprove of the overemphasis on economic impact (Bornmann 2013; Miettinen, Tuunainen, and Esko 2015). Some also criticize the dominance of STEM models of 'good' research and the focus on short-term proximal value (Molas-Gallart 2015; Greenhalgh et al. 2016; Reale et al. 2018). These reviewers argue that more sophisticated or additional approaches need to be developed that take into account the heterogeneity of societal value per discipline, especially for research in arts, humanities and social sciences (Budtz Pedersen, Grønvad, and Hvidtfeldt 2020). This is in line with studies that point out that the relations of research with societal actors differ fundamentally between and even within research fields (Hessels et al. 2011).

The authors of most of these reviews provide practical guidance regarding the most appropriate methods or novel approaches to capture societal value by either more diverse or more standardized techniques. For example, Penfield et al. (2014) conclude with a 'mixed method' approach to pull all data together; Molas-Gallart (2015) proposes a special methodology to gather evidence for value generated by arts and humanities research; Miettinen, Tuunainen, and Esko (2015) develop a framework for qualitative analysis in distinction to economic, quantitative, and constructivist approaches; De Silva and Vance (2017) suggest that an integration of altmetrics and bibliometrics could provide a 'near-complete picture of the impact of scientific research'. The authors of one review embrace precisely the lack of standardization and make it their goal to ensure that 'the most appropriate [method] is selected' for different situations (Greenhalgh et al. 2016).

Overall, the focus in these reviews is on the policy demands for, and practical requirements of evaluation of the societal value of research. To date, there is little theoretical analysis or conceptual comparison of the concepts of societal value in the evaluation practice (Muhonen, Benneworth, and Olmos-Peñuela 2020; Thomas et al. 2020). This is problematic, however, given the active, performative role of evaluation in knowledge production and the steering effects of metrics and indicators on scientific practice (de Rijcke et al. 2016; Wilsdon 2016). Evaluation methods embody an implicit or explicit theory of excellence, or 'good' research. These theoretical assumptions become most visible when researchers behave strategically in response, but it also influences the activities and priorities of researchers in more subtle ways. For example, the inclusion of a societal impact criterion in research assessment redefines what counts as valuable research (Oancea 2019).

The importance of conceptual presuppositions in evaluation methods can be illustrated by the fierceness and diversity of recent criticism of the impact criterion in the Research Excellence Framework (REF) in the UK (Smith et al. 2020). Scholars have criticized both the 'implicit optimism' and the overemphasis on the extraordinary of the societal impact agenda. Derrick et al. plead to study 'grimpact', extreme examples of negative impact, to question the dominance of non-controversial, economic, and indisputably good versions of impact (Derrick et al. 2018). In addition, Savigny (2019) showed how impact practices, and public engagement in particular, are infused with raced and gendered norms and expectations. Sivertsen and Meijer argue to look beyond the rare incidences where existing or new interactions have unexpected widespread implications. Instead they ask attention for 'normal' impact which follows from everyday active, productive, and responsible relations between academic and other organizations for the conduct of research (Sivertsen and Meijer 2020). It is these kinds of assumptions in evaluation methods about the concepts and mechanisms of the societal value of research that shape the researchers' responses about, and possibly the practical relations to, societal actors.

3. Analytical framework

To the best of our knowledge, no analytical framework exists for a systematic comparison of the conceptual foundations of evaluation methods dealing with societal value. To fill this gap, we have designed an analytical framework based on a constructivist approach to research and evaluation. The framework consists of four key aspects that enable the comparison of evaluation methods, in particular with respect to the relation between scientific and societal value.

First, we will compare the *types and roles of actors* that are considered *part of the production of knowledge* in different methods. According to constructivist science studies nobody, or even no thing, can be ruled out as a relevant actor in knowledge production. Due can be paid to this insight in evaluation by an ontologically open attitude to the question 'who is doing science, after all?' (Latour 1987). In the practice of the evaluation of publicly funded research, a special interest of course does exist for the role of the staff of the research institution, group, or project that is being assessed. The issue at stake is how evaluation methods balance their interest in the evaluated actors with connected networks of heterogeneous actors.

Second, we will compare the *interaction mechanisms* that different methods presume fundamental to the creation of societal value. Building on the literature about knowledge utilization, transfer, and exchange (Jacobson 2007; Best and Holmes 2010; Ward, House, and Hamer 2017), we distinguish three different understandings of the exchange mechanism. *Linear* models allocate a central place to research in relative isolation from society. Results are subsequently transferred to external parties, who function only as consumers or users. *Cyclical* models (also named relational or interaction models) describe the importance of recurrent, reciprocal and sometimes highly structured interactions between researchers and external agents for the agenda-setting, production, and dissemination of research. *Co-production* models (also named system, integration or dynamic models) point to a breakdown of the hierarchy between producers and users, and instead de- and prescribe participatory processes of research in which academic and non-academic actors are both actively involved.

Third, we will compare the *concepts of societal value* covered by different assessment methods. As we indicated above, we take the societal value of scientific research to be a social construct. Its meaning depends on its use in policy situations, scientific practices, and societal contexts. In actual evaluation available time and data can restrict the scope of the societal value concept. We will use an existing typology of three uses of the term 'societal impact' in evaluation that correspond to distinctive moments in the exchange process (Bornmann 2013; De Jong et al. 2014). Impact appears as *product* in terms of knowledge with potential value (also known as output); as *use* of this product by stakeholders (also known as outcome); and finally, as the *benefits* that follow from this use.

Fourth, we will use the three characteristics of actors, mechanisms and concepts to compare the *views on the relationships between scientific and societal value of scientific research*. Do assessment methods rely on an integrated concept of scientific and societal value or on a separation between these two? Our constructivist position implies that both scientific and societal value are generated in complex processes that can overlap to a large extent. Theoretically, one could argue for one integrated evaluation process of the value of scientific research, including all engaged actors and exchange mechanisms. Practically, most of the evaluation methods that we discuss here have been proposed in addition to a strongly institutionalized evaluation practice of 'scientific' value.

We have selected 10 assessment methods for analysis that all aim to capture the societal value of scientific research for evaluation purposes.[1] The selection is based on scientific database queries as well as comparison of the various reviews of evaluation methods (see Section 2.2). In our selection, we have attempted to achieve a reasonable coverage of the diversity in terms of the technical approaches (type of evaluation and data used). Our geographic and disciplinary focus is limited to regions and research fields with advanced science policies and evaluation practices. This makes it possible to critically explore the theoretical consequences of the link between policy, science studies, and scientific practice (see Table 1).

4. Review of 10 evaluation methods

Below, we introduce the reviewed evaluation methods. For each method, we highlight its context of origin, disciplinary focus and scope of data collection. We also introduce the actors, mechanisms, and concept of societal value each method includes, the three theoretical aspects on which we will compare the different methods in Section 5. We refer to some key publications and supporting literature for each method.

The *Payback Framework* was developed in the UK healthcare context for the specific aim of describing the wider social benefits that follow from providing evidence for policy. Subsequently, the framework has been used for, and further developed in, the evaluation of national funding programmes, also in the social and biomedical sciences. Various qualitative data sources are mobilized to analyse paybacks, such as surveys, interviews, and policy documents (Klautzer et al. 2011). The framework consists of a cyclical model of seven stages of research, from the inception of a research idea to the final societal outcomes, with several feedback loops to avoid the suggestion of linearity (Donovan and Hanney 2011). The *Payback Framework* emphasizes policymakers as main recipients of research, and it includes two interfaces for interaction between researchers, policymakers, and potential users of research, namely in the first stage of 'project specification' and the intermediate stage of 'dissemination'. The value of the research process is expressed in terms of 'paybacks', or 'impacts', which are classified in five dimensions that correspond with different steps in the model. Klautzer et al. (2011) generalized these impacts to make them applicable also beyond the health sector: knowledge (e.g. academic publications), impacts on future research (e.g. training new researchers), impacts on policy (e.g. at national level or within organizations), impacts on practice (e.g. cost savings in health), and broader social and economic impacts (e.g. commercial spin-offs or public debate).

Science and Technology (S&T) Human Capital has been developed to emphasize the 'socially embedded nature' of knowledge production and exchange. Based on Bourdieu's concept of social capital, this approach focuses on the growth of capacities and capabilities of individuals and, by addition, of groups and projects (Bozeman, Dietz, and Gaughan 2001). Human capital is operationalized by taking stock of individual career trajectories and the different types of knowledge acquired in that process (including tacit, craft, and know-how), as well as the productive social networks that sustain the creation of knowledge. The value of research is described in terms of human and social capital *increase*. This requires quite specific kinds of data: from activity diaries, resumes, and interviews

Table 1. Overview of the assessment methods reviewed in this article

Method	Evaluation type	Level of analysis	Qualitative data	Quantitative data	Original context	Key publication
Payback Framework	Ex post; summative	Programme	Documents, interviews, surveys	–	UK medical research	Buxton and Steve (1996)
Science and Technology Human Capital	Ex post; formative	Research group or programme	Interviews, surveys, diaries, resumes, contracts	Citation and patent patterns	US STEM research	Bozeman, Dietz, and Gaughan (2001)
Public Value Mapping	Ex ante and ex post; formative	Programme or organization	Case studies, documents, surveys, focus groups, expert opinions	Indicators	US science policy	Bozeman (2003)
Monetisation	Ex post; summative	Programme or system	–	Measures of investment and (health) gains	UK medical research	Health Economics Research Group (HERG), Office of Health Economics, RAND Europe (2008)
Flows of Knowledge	Ex post; summative	Programme	Case studies, documents, interviews, surveys, focus groups	Bibliometrics	UK research council funding	Meagher, Lyall, and Nutley (2008)
SIAMPI	Ex ante and ex post; formative	Project, programme, or organization	Case studies	Contextual response analysis and indicators of (im)material interactions	Research institutes (ICT, health, SSH, nano) for European Commission	Spaapen and van Drooge (2011)
Contribution Mapping	Ex post; summative and formative	Project or programme	Interviews with all actors	–	Global health sector	Kok and Schuit (2012)
Impact Narratives (REF)	Ex post; summative	Research group	Structured case studies, (user) expert opinions	Indicators for causal impact	UK assessment of university research (REF)	REF (2012)
ASIRPA	Ex post; summative	Programme or organization	Standardized case studies	Econometric, bibliometric and statistical methods	French public agricultural research institute	Joly et al. (2015)
Evaluative Inquiry	Ex post; formative	Research group or organization	Documents, interviews, workshop	Contextual scientometrics, contextual response analysis	Dutch assessment of university research (SEP)	de Rijcke et al. (2019)

The methods are ordered chronologically based on their key publications.

(to trace individual capacities), to contracts, citation, and patent patterns and questionnaires (to map the social network). The main actors are *mobile* scientists and engineers, who as such embody knowledge exchange between academia and other environments like start-ups, firms, or other universities. Ultimately, this 'holistic' approach has to speak to actors' own perception of the interconnected nature of all aspects of their work: the scientific, commercial, and social value of research depend on 'the conjoining of equipment, material resources (including funding), organizational and institutional arrangements for work, and the unique *S&T Human Capital* embodied in individuals'.

*Public Value Mapping (*PVM) was created to highlight the non-economic value of STEM research in US federal and state funding programmes *and* to align science policy with more diverse public values (Bozeman 2003). This approach builds on the previous method as well as on pragmatist theory. The mapping exercise starts with a case study consisting of the identification of relevant public values through document research and opinion polls. These values are subsequently hypothetically linked to research outputs so that these linkages can be empirically tested with indicators of social impacts (Bozeman and Sarewitz 2011). The latter are defined as the extent to which research contributes to broad social goals, a process in which 'knowledge value collectives' play an important role. These collectives include many parties, such as funding agents, end users, citizen groups, and commercial parties. The presupposition is that the production of research oriented at public values and its translation into uses takes place in contact with these broader collectives. In the *PVM* approach, value of knowledge, ultimately, consists in its use within the collective.

Monetization methods offer abstract evaluations of the societal value of research investments, in terms of economic returns. The evaluation method was first developed in the context of the UK health system and is mainly used in the highly institutionalized field of medical research. Societal value is defined as improvements to healthcare—in terms of cost-reduction or an increase in 'quality adjusted life years' (QALY) (Glover et al. 2014). For example, one can calculate the societal value of cardiovascular research in Canada by relating public investments in this type of research to estimates of the QALY for all the unique users of the different treatments (de Oliveria et al. 2013). The *Monetization* method only works with input and output indicators for funding and health gains and contains no explicit understanding of interaction mechanisms that produce societal value, apart from the linear chain from research to treatment to QALY increase. Similarly, no concrete actors are defined explicitly.

The *Flows of Knowledge* approach was developed for the evaluation of research council programmes that funded, for example psychology and mathematics research in the UK. It is inspired by a 'linkage-and-exchange' model that was used for Canadian health services research (Meagher, Lyall, and Nutley 2008). *Flows of Knowledge* is a multi-method approach including document research, surveys, case studies, semi-structured interviews, occasional bibliometrics and focus groups. It distinguishes between researchers, practitioners, policymakers, and private enterprises and especially highlights institutional and individual intermediaries (from funders and media to consultants and PhDs) in the process of societal value creation. Arrows of interaction generally flow both ways, emphasizing the point that long-term relationships of mutual respect, iterative dialog, and reciprocal benefits are an important proxy for non-academic impact (Meagher and Lyall 2013). Over the years, the concept of impact in the *Flows of Knowledge* method was diversified to five types (instrumental, conceptual, capacity, cultural, and connectivity) that could be realized by 27 different mechanisms (Meagher and Martin 2017).

SIAMPI (Social Impact Assessment through Productive Interactions) originated in a European funded project for which the production of societal value was studied at research institutes and departments in various European countries and disciplines (Spaapen and Van Drooge 2011). The premise of this approach is that one should not study the impact of research but the processes that function as proxy for impact. Productive interactions are exchanges (direct as well as mediated by material carriers or money) between researchers and stakeholders in which knowledge is produced and stakeholders make an effort to use this. The method prescribes to use field-specific quantitative indicators and qualitative data from case studies. Although this approach takes researchers as the main actors that produce knowledge, it considers a wide range of stakeholders to be a part of this process, including researchers in neighbouring fields, industry, public organizations, the government, and the general public. The *SIAMPI* approach does not make a clear distinction between productive interactions and societal impact 'because the transition from interaction to impact is often gradual' (Spaapen and Van Drooge 2011: 212).

Contribution Mapping was first used as a learning tool in the context of global health research and focuses on contributions and processes in order to avoid the overemphasis on impact and knowledge producers in other methods. This approach is inspired by actor-network theory and builds on the *Payback Framework*. To understand how research leads to action for health, *Contribution Mapping* focuses on the way users collect and combine knowledge (Kok and Schuit 2012). 'Process maps' are iteratively produced, from document analysis and interviews with researchers, potential key users, and other stakeholders, and used for improvement and accountability. Researchers and linked actors (practitioners, policymakers, participants, patient group representatives, opinion leaders) are considered equally involved in collective translation efforts that lead to knowledge utilization. Each of these actors can undertake 'alignment efforts' to enhance the likeliness that research contributes to action, for example by engaging linked actors in priority-setting or data interpretation. In *Contribution Mapping*, societal value is defined in terms of contributions to actor scenarios to stress that the role and meaning of research outcomes also depend on the context of users.

The *Impact Narratives* method has been developed as part of the Research Evaluation Framework (REF) for UK higher education institutes, which included impact as a criterion of assessment for the first time in 2014. This narrative case study approach is based on an elaboration of the *Payback Framework* with an additional impact rating scale focused on interactions with end users (Samuel and Derrick 2015). The *Impact Narratives* method of REF is applicable to all disciplines and based on expert review of the case studies. Research units produce exemplary narratives in which they causally relate high quality research to impact on societal stakeholders within detailed timeframes. This presupposes a rather linear process of exchange from clearly defined producers of knowledge to users outside of academia. Impact is defined 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'. A large diversity of potential beneficiaries is identified by the researchers themselves (Grant and Hinrichs 2015). Impact also includes 'reduction or

prevention of harm, risk, cost or other negative effects'. Ultimately, this broad definition is assessed according to two outcome-based criteria: 'significance' (meaning intensity of the effect) and 'reach' (meaning the spread amongst relevant constituencies).

ASIRPA ('Socio-Economic Analysis of the Impact of Public Agricultural Research') was developed as an ex post approach for assessing the socio-economic impact of public-sector research organizations, the French National Agricultural Research Institute in specific (Joly et al. 2015). Inspired by actor-network theories of innovation, this method focuses on the process of impact generation. The analysis consists of identifying 'the chain of translations', two-way processes that also transform problems, knowledge, and goals (Matt et al. 2017). The method consists of qualitative case studies, for 'thick description of specific situations', which are streamlined with the standardized outline of impact pathways so as to enable comparison of the diversity and amplitude of impact. These case studies include all the actors that happen to be involved in the different phases of an impact pathway according to a prescribed binary of academics, firms, extension agencies, public institutions, media, and farmers. Intermediaries, both people, organizations and artefacts, play an important role in the mechanisms of exchange. For example, technological objects contribute to the dispersion of knowledge from first users to massive utilization. *ASIRPA* uses as concept of impact 'direct and indirect effects of the various components of research on the economy, environment, health, etc. ... generated by lengthy and complex processes' (Matt et al. 2017). From 32 case studies, the *ASIRPA* authors extract four different impact ideal types based on the involvement of, and effects on, users.

Evaluative Inquiry is a recent approach to the evaluation of research groups and institutes that was developed in the Dutch university research assessment context. Building on *PVM*, *SIAMPI*, and *ASIRPA*, it consists of a mixed-methods approach, including contextualized scientometrics, productive interactions, and impact pathways, tailored to specific research units and evaluation purposes. The methods used may represent research 'numerically, verbally, and/or visually in ways that make visible the complexity of actual practice and its engagements' (de Rijcke et al. 2019). The authors of this approach identify networks of people, infrastructures, technologies, and resources as collectives and understand academic achievement as distributed over a host of academic and non-academic actors. Impact is an effect of translations within and between networks of actors that make up academic research and its environments. This method aims to do justice to the complexity and heterogeneity of research practices. Indicators are therefore not predefined but aligned with the work and mission(s) of the unit under assessment. Together, the actors under evaluation and the *Evaluative Inquiry* analysts determine the scope of actors and mechanisms, and the meaning of impact concepts. *Evaluative Inquiry* represents an explicit ambition to move away from a detached, clear delineation of academic value and to overcome the divide between the academic and the societal.

5. Analysis of evaluation methods

Table 2 summarizes our analysis of the 10 evaluation methods in terms of their main conceptual assumptions. We will elaborate on this in the following sections.

5.1 Roles of actors in research production, exchange, and evaluation

Unsurprisingly, all methods include researchers as primary actors involved in the process of knowledge production. But evaluation methods construe different objects of evaluation by dealing in diverse ways with other actors. Some methods focus only on the process of knowledge exchange or impact generation. This turns the research process effectively into a black box from which non-scientific actors are excluded. This is the case for example with *Impact Narratives*, *Flows of Knowledge*, and *S&T Human Capital*. The object of evaluation in these cases is impact in terms of the dissemination, use and effects of research results. The roles of non-academic actors tend, accordingly, to be limited to user or beneficiary. In other methods, like *Payback*, *PVM*, *SIAMPI*, *Contribution Mapping*, *ASIRPA*, and *Evaluative Inquiry*, knowledge exchange is situated as one part of a larger process of knowledge production. Such methods construct a more complex object of evaluation as they include the entire process of research in their understanding of impact, allowing more diverse roles for non-academic actors.

Many methods reproduce institutional distinctions in their perspective on relevant actors and their roles—for example particular investigators or research groups funded by a specific project or part of a public institution—and track 'forward' from knowledge production to eventual impacts. Relevant actors are in most cases then identified by the researchers themselves. Few methods work the other way around, tracking 'backwards' from societal change to knowledge production: *Monetisation* starts at health benefits and works backwards to public investments, while *Contribution Mapping* offers it as one of multiple strategies. *PVM* adopts backtracking as central methodological approach by testing hypotheses about connections between public values and previous research contributions. In this way, *PVM* analysts construct evaluation objects by identifying relevant political, societal, and research actors as part of one collective connected to certain values and a field of knowledge. Similarly, *SIAMPI* is based on 'productive interactions' to focus on the *process of interaction* from which researchers emerge not as detached first movers, but as one stakeholder amongst other actors.

Most methods include users as informants, for example in case studies about specific impacts (with the exclusion of *Monetisation* and *S&T Human Capital*). Only some methods also allow non-research actors to design the evaluation process itself, for example by identifying relevant actors and setting assessment criteria. Bozeman (2003), who designed *PVM*, considers 'knowledge users the proper evaluators' and the authors behind *Contribution Mapping* state that 'the roles and functions of those involved in the evaluation are not predetermined' but a topic of discussion at an early stage (Kok and Schuit 2012). Similarly, in the *Evaluative Inquiry* approach 'audiences are seen not only as (co)producers of knowledge and its impact, but also as (co)producers of the criteria by which such impact is to be evaluated' (de Rijcke et al. 2019).

A final difference that emerges from the comparison of actor roles in evaluation methods is the importance of the intermediary in relation to theories of knowledge production. Three methods explicitly include intermediaries or brokers as key actor in the production and/or exchange of knowledge: *PVM*, *Flows of Knowledge*, and *ASIRPA*. This centrality of intermediary actors in the impact process cannot be explained by a shared theoretical framework, which suggests that various methodological perspectives provide support for

Table 2. Comparison of the main theoretical assumptions of 10 evaluation methods

Method	Actor roles	Interaction mechanisms	Concept of societal value	Relationship societal–scientific value
Payback Framework	Policymakers and professionals as contractors, agenda-setters and users	*Cyclical*: 7 stages with interfaces and feedback	*Mixed*: Successively as products for, use by or benefits to research, policy, (health) practice and economy	Distinctive, successive categories
Science and Technology Human Capital	Scientists and engineers as producers and carriers of knowledge	*Linear*: People mobility	*Product*: Increase in human capital	Embodied
Public Value Mapping	Institutional, social and economic 'end users'; 'knowledge value collectives' as translators of research to new uses	*Cyclical*: Knowledge value collectives	*Mixed*: Tracked backwards from public benefits to societal use and research outcome	Integrated
Monetisation	Clinicians as users, patients as beneficiaries	*Linear*: Linear chain	*Benefit*: Improvements to healthcare	Implicitly connected
Flows of Knowledge	Practitioners and policymakers as specific users; organizations and individuals as intermediaries	*Cyclical*: Dynamic process of iterative dialogue and reciprocal benefits	*Benefit*: 5 types of impact (instrumental, conceptual, capacity, cultural and connectivity)	Distinctive categories
SIAMPI	Actors from science, industry, government and non-profits as stakeholders in knowledge use	*Cyclical*: Productive interactions	*Use*: (productive interactions)	Not clearly distinguishable
Contribution Mapping	Scientific and societal actors (including organizations, objects) engaged in priority-setting, proposal selection; producing, combining and using knowledge	*Co-production*: Alignment	*Use*: Contribution to actor scenarios	Integrated
Impact Narratives (REF)	Non-academic actors from society, economy, culture and public policy as (potential) beneficiaries	*Linear*: Linear exchange	*Benefit*: Effect, change or benefit beyond academia	Causally related
ASIRPA	Academic, economic, knowledge transfer and governmental actors as part of research production and, with media and farmers, as intermediaries and beneficiaries. Also objects as intermediaries	*Cyclical*: Translation networks and iterative learning processes	*Mixed*: Effects on economy, environment, health, etc.	Integrated
Evaluative Inquiry	Networks of people, technologies and resources connected to research units enable achievement of academic and societal value	*Co-production*: Translations within and between networks	Not pre-defined	Integrated

the importance of intermediaries, or knowledge brokers, in the impact process. When we look, on the other hand, at four methods that explicitly share a theoretical commitment, in this case to actor-network theory, we find that they try to include materials or non-human actors in their model of knowledge production and exchange: *ASIRPA* (technological objects); *SIAMPI* (material 'carriers' such as texts, exhibitions, models or films); *Contribution Mapping* (non-human actors in actor scenarios); and *Evaluative Inquiry* (research outcomes). It remains unclear to what extent this non-human agency results in more than a semantic twist to the impact narrative.

5.2 Mechanisms of knowledge exchange

In our sample, we have identified instances of all three mechanisms of knowledge exchange. Two methods work with a *linear model* of knowledge exchange: *Impact Narrative* and *Monetisation*. In their approach, knowledge users feature mainly as recipients of knowledge rather than as active co-producers. These methods have in common that they were designed for summative, rather than formative, purposes. Five methods fit with a *cyclical model*, emphasizing the feedback mechanisms between the production of knowledge and the application of knowledge. The authors of these methods refer to this with terms such as 'feedback loop' (*Payback*), 'non-linear pathway' (*ASIRPA*), and the 'churn model' (*PVM*). Two methods, at last, maintain a *co-production model* to explain both production and exchange of knowledge, which allocates more agency to users and intermediaries.

Examples of cyclical knowledge exchange models with several feedback loops are the *Payback Framework*, *Flows of Knowledge*,

ASIRPA, and *SIAMPI*. The *Payback Framework* consists of a detailed model including a number of specific feedback paths. The model includes arrows from outputs, adoption, and final outcomes back to topic identification, project specification, inputs to research and secondary outputs. *Flows of Knowledge* harbours an 'indirect, non-linear' understanding of research impact and considers bi-directional knowledge flows from researchers to policymakers and practitioners. The key assumption of *ASIRPA* is that impact of research develops 'over a non-linear pathway' in five main steps (inputs, outputs, intermediaries, impact 1, and impact 2). One-directional arrows between these steps suggest linear causality, but the approach explicitly emphasizes iterative learning processes. Also, the actors in *SIAMPI* mutually influence each other, so that societal value of research is the result of an iterative process between science, government, industry, and non-profit organizations (Molas-Gallart and Tang 2011). When productive interactions are deployed as method, the attention is typically put however more on the effects of productive interactions in society, than on scientific knowledge (De Jong et al. 2014; Muhonen, Benneworth, and Olmos-Peñuela 2020).

Two methods are based on a model of co-production. The mechanisms of knowledge exchange remain relatively unspecified. The *Contribution Mapping* advocates claim that agency in knowledge utilization is distributed between a number of actors and eventual change cannot be attributed to a single source. The assumption is that changes in action resulting from research are 'part of evolving, complex, and open systems in which change is continuous, non-linear, multi-directional and difficult to control' (Kok and Schuit 2012). However, it is not specified how these systems would evolve, for example in terms of feedback loops. *Evaluative Inquiry* emphasizes the distributed nature of knowledge production and the heterogeneity of the actors involved. The authors stress the active role of stakeholders as (co)producers of knowledge and impact as opposed to a passive role as 'audiences' (de Rijcke et al. 2019).

S&T Human Capital deviates from the pack because it does not consider knowledge users at all. As this method does not define societal value in terms of the usage of knowledge, but in terms of capacities, mobility, and careers of researchers, it puts forward a unique perspective on knowledge exchange in which knowledge travels in the people that embody it, rather than by transfer or interaction between different actors.

5.3 Concepts of societal value

Amongst the methods we have analysed, a variety of terms is used to describe the societal value of research: from impact and payback to public value, contributions, and social capital. We will compare these concepts with respect to the threefold characterization of societal value as product, use, and benefit.

S&T Human Capital is the only method that restricts its view on societal value to a product concept. This method gauges scientific research in its potential to contribute to non-academic environments in the sense of capabilities moving between spheres. Societal value as the use or uptake of research by stakeholders is difficult to identify in pure form in our sample of methods. The focus on proxies for societal impact in *SIAMPI* and *Contribution Mapping* comes closest to the use concept. Although their concepts of societal value do not imply actual use of knowledge products, they do regard societal value as contributions of research to the (potential for) actions of non-academic stakeholders. Lastly, we identified societal value as ultimate societal benefit or effect in three methods: *Impact Narratives*, *Flows of Knowledge*, and *Monetisation*. Each of these methods ultimately rests on indicators of change beyond academia, in policy, economy, or the environment. Note that there is also variety within this benefit concept of societal value. *Flows of Knowledge*, for example describes five types of benefits, ranging from the tangible to the intangible.

Three methods mix all three elements in their concept of societal value. In the *Payback Framework*, societal value successively emerges as product directly from the research process, as use after dissemination and adoption, and as benefit when the final outcome is reached. Interestingly, this is turned around in *PVM*: starting from a perceived public value (or benefit), the approach tries to establish a plausible link with a research result (or product) via intermediate users of that knowledge. Lastly, *ASIRPA* includes the three conceptual aspects of societal value in a matrix of four impact ideal types, ordered according to the levels of co-production and affectation of the users (Joly et al. 2015). In *Evaluative Inquiry*, it is left up to each evaluation context to choose a suitable definition of societal value, allowing in principle the product, use and benefit version of the concept.

5.4 The relation between societal and scientific values

This brings us to the conclusion of this analysis: the relation between the societal and scientific value of research. Some methods are not explicit about the relation between societal and scientific value. The two types of value may be implicitly assumed to be similar, or the precise relation between the two is not elaborated. *Monetisation* relies, for example on the highly institutionalized field of medicine, where there is a widely shared understanding of societal value (improving healthcare practices).

Other methods do make clear distinctions between scientific and societal value. The *Payback Framework* contains distinctions between five payback categories, two of which are situated in close proximity of the research process (knowledge and benefits for future research); the other three relate to societal effects (policy, health and economic benefits). The scholars behind the *Flows of Knowledge* approach implicitly distinguish clearly between the scientific and societal value of research. In the REF assessments, the two are not only clearly separated from each other but also causally related: *Impact Narratives* have to be based on societal impacts that can be related to research of 'high scientific quality'. The risk of this is that one creates blind spots for societal value based on 'mediocre', or normal, research. This is particularly relevant because it is not at all well established that scientific excellence is a proper predictor of societal value (Buxton 2011). For these three methods, it seems that their origin in a practical request from policy has isolated the production of research from its exchange and use.

Finally, there are several methods that incorporate the view that the networks that produce scientific and societal value coincide, at least partly. The *SIAMPI* authors do not distinguish in a generic way between scientific and societal value of research. In this approach, the precise relations between the dimensions of scientific value ('robustness') and societal value ('relevance') depend on the specific field of research (Spaapen and van Drooge 2011). For methods like *ASIRPA*, *Contribution Mapping*, *PVM*, and *Evaluative Inquiry*, the mechanisms that produce societal value—respectively, chains of translations, alignment efforts, knowledge value

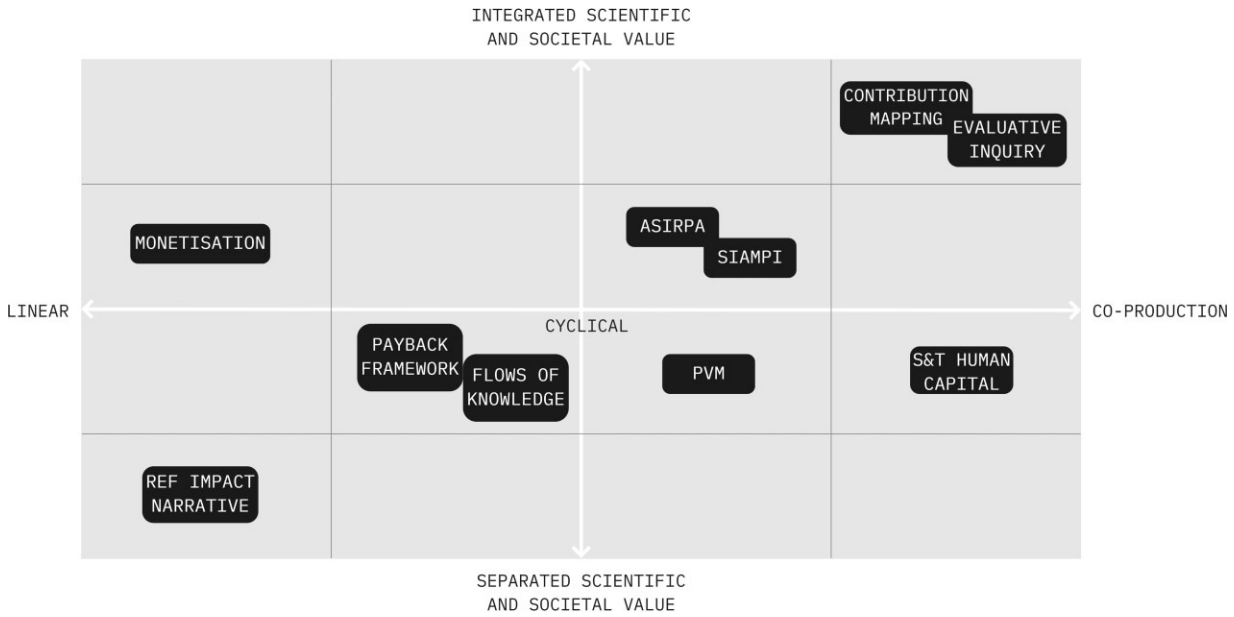

Figure 1. A classification of the evaluation methods with respect to the relation between scientific and societal values (vertical axis) and the knowledge exchange model (horizontal axis).

collectives, and socio-technical networks—are integral to the research process as such.

5.5 Comparative analysis

When comparing our analyses of the different methods, there seems to be a correlation between the level of aggregation at which an evaluation method approaches the research process, and their concept of societal value in terms of products, use, or benefit. The three methods that have a mixed concept (*Payback Framework*, *PVM*, *ASIRPA*) take entire research organizations, fields, or programmes into account. When methods evaluate the researchers and their groups in particular, the societal value concepts also remain closer to the research practice (i.e. product or potential use). With respect to the concept of societal value, two approaches stand out: *Impact Narratives* and *S&T Human Capital*. Both hold a product concept of societal value, take researchers as the primary actors, and use a linear model of knowledge exchange. All methods that take a broader set of actors and interactions into account have a use, benefit, or mixed concept of societal value.

Figure 1 illustrates our analysis by visualizing two of the four aspects of our analytical framework: the methods' knowledge exchange model and their understanding of the relationship between scientific and societal values. When comparing the knowledge exchange models with the perspective on scientific versus societal value, we only observe an association between a co-production model and integration of both values (see Figure 1). Apparently, the conviction that scientific and societal values are strongly related is not compatible with a traditional view on researchers as the primary actor of knowledge production.

The only method that does not use any qualitative data (*Monetisation*) corresponds with a linear model of knowledge exchange. This suggests that quantitative data can carry a bias towards a linear model, while qualitative data, like interviews, allow (but not prescribe) non-linear views on knowledge exchange. This does not imply, however, that quantitative data are of no use to evaluation methods that hold cyclical or co-production models of research. For one, we see that metrics and indicators can supplement qualitative data and moreover and that alternative 'contextual' types of metrics are employed, which do not presuppose a linear model of exchange.

When comparing the purpose of the methods with their conceptual principles, we see that most summative methods hold a linear model of knowledge exchange, and formative methods take either a cyclical or co-production model. But we find no relationship between the purpose of evaluation and the concept of societal value, or the relationship between scientific and societal values, probably because the summative/formative distinction is not equal to a product/process distinction (Scriven 1996).

6. Reflection and discussion

Over the past few decades, a rich set of tools has been developed to measure, compare and assess the societal value of research. The aim of this review article was to analyse how impact evaluation methods relate to, and operationalize, the distinction between scientific and societal value. Our analysis has shown that different methods construct different objects of evaluation and produce a variety of societal value concepts.

First, because of their theoretical starting points, the methods construct different objects of evaluation. Some focus the attention selectively on knowledge exchange, dissemination, or impact generation as activities separate from the research process, while others treat knowledge production, translation and transformation as one integrated process. Second, the methods also construct different stakeholders of scientific research. Some methods produce a strong contrast between academic and non-academic actors, by considering societal stakeholders exclusively as the users of the final outcomes of scientific research. Other methods use a more inclusive concept of stakeholders and evaluate a long chain of connected actors, including intermediaries. Finally, we found that the different methods enable the production and articulation of fundamentally different

types of value, as product, use, or benefit. Moreover, the societal value that they make visible and comparable relates to scientific value in various ways. While *Contribution Mapping* and *Evaluative Inquiry* articulate societal and scientific value in an integrated way, many methods specify societal value as separate from scientific value. These assessment methods reinforce a distinction between two systems of valuation, regardless of the fact that many authors ascribe to the theoretical principles and empirical findings of constructivist science studies.

The analytical and practical distinction between scientific and societal value in many methods is not surprising per se. The methods reviewed here were driven by a need to make societal value more explicit alongside the scientific value of research, which is traditionally more visible in the form of quantitative indicators and peer review assessments. Many of the methods have been developed at the request of funding or policy bodies, which suggests that a well-delineated object of evaluation existed beforehand. Only in some methods the starting point is the process of societal value creation rather than the actors of interest (publicly funded researchers or institutes in most cases). The balance that evaluation methods strike between societal value as separate performance indicator and as part of research practice will ultimately depend on the purpose of evaluation. Methods whose aim is comparison (in summative evaluations) employ generic approaches with clearly formulated actors and indicators, and a rigid concept of societal value. But for purposes of situational learning (in formative evaluations), methods take more tailored approaches. Hybrid approaches, which navigate between the extremities of standardization and specificity, offer interesting alternatives to the dichotomy of summative and formative evaluation (Lau 2016). We believe that an integrated concept of scientific and societal value, adaptable to the local situation of the department, group, or field under study, will encourage doing what we value most instead of doing what counts (Wouters 2017). Producing a strong distinction between scientific and societal value may stimulate researchers to concentrate on well-delineated activities that either yield peer recognition, like scientific publications, or that can be measured by indicators for societal value, like blogposts, patents, or policy reports. Moreover, a strict separation might limit the awareness of the heterogeneity of actors and institutions involved in knowledge production. This is especially urgent as intermediary and boundary actors, such as think tanks, consultancy firms and civil collectives, play increasingly important roles in science and innovation systems (Etzkowitz and Zhou 2017).

Our analysis therefore incites a reflection on the *position of evaluation* itself in the process of knowledge production. Evaluation methods do not simply *de*scribe but also *pre*scribe how societal value of research is produced (de Rijcke et al. 2016). The designers of some methods clearly show an awareness of the ways in which evaluation intervenes in research policy and practice. *Monetisation*, for example aims to advocate for future investments by demonstrating good value, while the *PVM* approach aims to align science policy with more diverse public values. With respect to research practices, methods like *SIAMPI*, *Flows of Knowledge*, and *Contribution Mapping* consider themselves as 'tools of enlightenment' that could support organizational learning about the conditions and obstacles to societal impact. The vantage point of *Evaluative Inquiry*, lastly, explicitly takes evaluation itself as knowledge production 'transforming evaluators and analysts into collaborators alongside evaluees' (de Rijcke et al. 2019). Although this co-productive approach of *Evaluative Inquiry* is unique, an active role of evaluation in knowledge production applies to all assessment methods discussed here.

For that reason, we hesitate to finish this critical review with a set of recommendations as to which evaluation method is most effective. Ultimately, we would advise policymakers and research managers to use evaluation tools that match both the research practice under evaluation and the theoretical convictions about knowledge production, exchange, and translation in that field. This suggestion implies that a discussion about the fundamentals of knowledge production in a particular field or institute has to take place between all relevant actors as part of the evaluation process. We hope that our analytic overview can help policymakers and research managers in selecting the method that fits best to their situation based on a consideration of policy goals, theoretical convictions, and practical constraints (available data, time, and money).

On a more fundamental note, we recommend to all our fellow science studies scholars to keep questioning the theoretical assumptions of the policymakers or research managers that ask for methods and tools. Our professional responsibility is to develop methods that are grounded in the theoretical developments of the field, also when they are at odds with direct practical needs. Evaluation methods that combine the different aspects of societal value and align with the way the various actors in knowledge production perceive value do most justice to the practice of research and impact. It is our conviction that these evaluation methods will contribute most to learning processes that improve the societal value of scientific research.

Note

1. Note that some methods focus entirely on the evaluation of societal value, while others include this variable next to other aspects under evaluation.

Acknowledgements

The authors gratefully acknowledge Caroline Wagner for many stimulating discussions and constructive feedback. They thank Stefan de Jong, Tjitske Holtrop, Leonie van Drooge, Sarah de Rijcke, Rodrigo Costas, and participants of the CWTS Seminar at Leiden University for helpful comments on earlier drafts. Lastly, special thanks to the anonymous reviewers for their excellent suggestions.

Funding

Part of this work was supported by the Netherlands Organisation for Scientific Research NWO [Grant number: 322-69-011]

Conflict of interest statement. None declared.

References

Best, A., and Holmes, B. (2010) 'Systems Thinking, Knowledge and Action: Towards Better Models and Methods', *Evidence & Policy: A Journal of Research, Debate and Practice*, 6: 145–59.

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

Bozeman, B. (2003) 'Public Value Mapping of Science Outcomes: Theory and Method', in Bozeman B., Sarewitz, D., Feinson, S., Faladori, G., Gaughan, M., Gupta, A, Sampta, B., and Zachary, G. (eds) *Knowledge Flows and Knowledge Collectives: Understanding the Role of Science and Technology*

Policies in Development, 2, pp. 3–48. Tempe: Consortium for Science, Policy & Outcomes.

Bozeman, B., Dietz, J. S., and Gaughan, M. (2001) 'Scientific and Technical Human Capital: An Alternative Model for Research Evaluation', *International Journal of Technology Management*, 22: 716–40.

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

Budtz Pedersen, D., Grønvad, J. F., and Hvidtfeldt, R. (2020) 'Methods for Mapping the Impact of Social Sciences and Humanities—A Literature Review', *Research Evaluation*, 29: 4–21.

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

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

Callon, M. (1997) 'Four Models for the Dynamics of Science', in Tauber A. I. (ed) *Science and the Quest for Reality*, pp. 249–92. London: Macmillan Press.

Dahler-Larsen, P. (2011) *The Evaluation Society*. Stanford: Stanford University Press.

De Jong, S. P. L., Barker, K., Cox, D., Sveinsdottir, T., and Van den Besselaar, P. (2014) 'Understanding Societal Impact through Productive Interactions: ICT Research as a Case', *Research Evaluation*, 23: 89–102.

De Jong, S. P. L., Smit, J. P., and van Drooge, L. (2016) 'Scientists' Response to Societal Impact Policies: A Policy Paradox', *Science & Public Policy*, 43: 102–14.

de Oliveira, C., Nguyen, H. V., Wijeysundera, H. C., Wong, W. W. L., Woo, G., Grootendorst, P., Liu, P. P., and Krahn, M. D. (2013) 'Estimating the Payoffs from Cardiovascular Disease Research in Canada: An Economic Analysis', *CMAJ Open*, 1/2: E83–90.

De Silva, P. U. K., and Vance, C. K. (2017) 'Assessing the Societal Impact of Scientific Research', in De Silva, P. U. K. and Vance C. K. (eds) *Scientific Scholarly Communication: The Changing Landscape*, pp. 117–32. Cham: Springer International Publishing.

De Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., and Hammarfelt, B. (2016) 'Evaluation Practices and Effects of Indicator Use—A Literature Review', *Research Evaluation*, 25: 161–9.

De Rijcke, S., Holtrop, T., Kaltenbrunner, W., Zuijderwijk, J., Beaulieu, A., Franssen, T., van Leeuwen, T., Mongeon, P., Tatum, C., and Valkenburg, G. (2019) 'Evaluative Inquiry: Engaging Research Evaluation Analytically and Strategically', *Fteval Journal for Research and Technology Policy Evaluation*, 48: 176–82.

Derrick, G. E. et al. (2018) 'Towards Characterising Negative Impact: Introducing Grimpact', in 23rd International Conference on Science and Technology Indicators (STI 2018), September 12–14, 2018. Leiden, The Netherlands: Centre for Science and Technology Studies (CWTS).

Donovan, C. (2019) 'For Ethical 'Impactology'', *Journal of Responsible Innovation*, 6: 78–83.

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

Douglas, H. (2014) 'Pure Science and the Problem of Progress', *Studies in History and Philosophy of Science Part A*, 46: 55–63.

Edgerton, D. (2004) ''The Linear Model' Did Not Exist: Reflections on the History and Historiography of Science and Research in Industry in the Twentieth Century', in Grandin K., Wormbs N., and Widmalm S. (eds) *The Science-Industry Nexus. History, Policy, Implications*, pp. 1–36. Sagamore Beach, MA: Science History Publications.

Etzkowitz, H., and Zhou, C. (2017) *The Triple Helix: University-Industry-Government Innovation and Entrepreneurship*. London: Routledge.

Glover, M., Buxton, M., Guthrie, S., Hanney, S., Pollitt, S., and Grant, J. (2014) 'Estimating the Returns to UK Publicly Funded Cancer-Related Research in Terms of the Net Value of Improved Health Outcomes', *BMC Medicine*, 12/1: 10.1186/1741-7015-12-99

Godin, B. (1998) 'Writing Performative History: The New New Atlantis?', *Social Studies of Science*, 28: 465–83.

Godin, B. (2006) 'The Linear Model of Innovation: The Historical Construction of an Analytical Framework', *Science, Technology, & Human Values*, 31: 639–67.

Godin, B., and Doré, C. (2004) *Measuring the Impacts of Science: Beyond the Economic Dimension*. Montreal: Canadian Science and Innovation Indicators Consortium.

Grant, J., and Hinrichs, S. (2015) *The Nature, Scale and Beneficiaries of Research Impact: An Initial Analysis of Research Excellence Framework (REF) 2014 Impact Case Studies*. HEFCE-Higher Education Funding Council for England.

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

Health Economics Research Group (HERG), Office of Health Economics, RAND Europe (2008) *Medical Research: What's It Worth? Estimating the Economic Benefits from Medical Research in the UK*. London: UK Evaluation Forum.

Hessels, L. K., and van Lente, H. (2008) 'Re-Thinking New Knowledge Production: A Literature Review and a Research Agenda', *Research Policy*, 37: 740–60.

Hessels, L. K., van Lente, H., Grin, J., and Smits, R. E. (2011) 'Changing Struggles for Relevance in Eight Fields of Natural Science', *Industry and Higher Education*, 25: 347–57.

Jacobson, N. (2007) 'Social Epistemology: Theory for the 'Fourth Wave' of Knowledge Transfer and Exchange Research', *Science Communication*, 29: 116–27.

Joly, P., Gaunand, A., Colinet, L., Larédo, P., Lemarié, S., and Matt, M. (2015) 'ASIRPA: A Comprehensive Theory-Based Approach to Assessing the Societal Impacts of a Research Organization', *Research Evaluation*, 24: 440–53.

Kaldewey, D., and Schauz, D. (2018) *Basic and Applied Research the Language of Science Policy in the Twentieth Century*. New York: Berghahn Books.

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.

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

Lamont, M. (2012) 'Toward a Comparative Sociology of Valuation and Evaluation', *Annual Review of Sociology*, 38: 201–21.

Latour, B. (1987) *Science in Action: How to Follow Engineers and Scientists through Society*. Cambridge: Harvard University Press.

Lau, A. M. S. (2016) 'Formative Good, Summative Bad?'—A Review of the Dichotomy in Assessment Literature', *Journal of Further and Higher Education*, 40: 509–25.

Lynch, M. (2017) 'STS, Symmetry and Post-Truth', *Social Studies of Science*, 47: 593–9.

Matt, M., Gaunand, A., Joly, P.-B., and Colinet, L. (2017) 'Opening the Black Box of Impact—Ideal-Type Impact Pathways in a Public Agricultural Research Organization', *Research Policy*, 46: 207–18.

Meagher, L., and Lyall, C. (2013) 'The Invisible Made Visible: Using Impact Evaluations to Illuminate and Inform the Role of Knowledge Intermediaries', *Evidence & Policy: A Journal of Research, Debate and Practice*, 9: 409–18.

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.

Meagher, L. R., and Martin, U. (2017) 'Slightly Dirty Maths: The Richly Textured Mechanisms of Impact', *Research Evaluation*, 26: 15–27.

Miettinen, R., Tuunainen, J., and Esko, T. (2015) 'Epistemological, Artefactual and Interactional–Institutional Foundations of Social Impact of Academic Research', *Minerva*, 53: 257–77.

Molas-Gallart, J. (2015) 'Research Evaluation and the Assessment of Public Value', *Arts and Humanities in Higher Education*, 14: 111–26.

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.

Muhonen, R., Benneworth, P., and Olmos-Peñuela, J. (2020) 'From Productive Interactions to Impact Pathways: Understanding the Key Dimensions in Developing SSH Research Societal Impact', *Research Evaluation*, 29: 34–47.

Oancea, A. (2019) 'Research Governance and the Future (s) of Research Assessment', *Palgrave Communications*, 5: 27.

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.

Power, M. (2000) 'The Audit Society—Second Thoughts', *International Journal of Auditing*, 4: 111–9.

Proctor, R. (1991) *Value-Free Science?: Purity and Power in Modern Knowledge*. Cambridge: Harvard University Press.

Reale, E., Avramov, D., Canhial, K., Donovan, C., 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. (2018) '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.

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

Sand, F., and Toulemonde, J. (1993) 'Changing Criteria in the Evaluation of European R&D Programmes', in: Süss W. and Becher G. (eds) *Politik Und Technologieentwicklung in Europa. Analysen Ökonomisch-Technischer Und Politischer Vermittlungen Im Prozess Der Europäischen Integration*, pp. 237–60. Berlin: Duncker & Humblot.

Savigny, H. (2020) 'The Violence of Impact: Unpacking Relations between Gender, Media and Politics', *Political Studies Review*, 18: 277–293.

Scriven, M. (1996) 'Types of Evaluation and Types of Evaluator', *Evaluation Practice*, 17: 151–61.

Shinn, T. (2002) 'The Triple Helix and New Production of Knowledge: Prepackaged Thinking on Science and Technology', *Social Studies of Science*, 32: 599–614.

Sismondo, S. (2017) 'Post-Truth?', *Social Studies of Science*, 47: 3–6.

Sivertsen, G., and Meijer, I. (2020) 'Normal versus Extraordinary Societal Impact: How to Understand, Evaluate, and Improve Research Activities in Their Relations to Society?', *Research Evaluation*, 29: 66–70.

Smith, K., Bandola-Gill, J., Meer, N., Stewart, E., and Watermeyer, R. (2020) *The Impact Agenda: Controversies, Consequences and Challenges*. Bristol: Policy Press.

Spaapen, J., and van Drooge, L. (2011) 'Introducing 'Productive Interactions' in Social Impact Assessment', *Research Evaluation*, 20: 211–8.

Stengers, I. (1997) *Power and Invention: Situating Science*. Minneapolis: University of Minnesota Press.

Thomas, D. A., Nedeva, M., Tirado, M. M., and Jacob, M.. (2020) 'Changing Research on Research Evaluation: A Critical Literature Review to Revisit the Agenda', *Research Evaluation*, 29: 275–88.

Ward, V., House, A., and Hamer, S. (2017) 'Developing a Framework for Transferring Knowledge into Action: A Thematic Analysis of the Literature', *Journal of Health Services Research & Policy,* 14: 156–164.

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

Wilsdon, J. (2016) *The Metric Tide: Independent Review of the Role of Metrics in Research Assessment and Management*. London: Sage Publications Ltd.

Wouters, P. F. (1999) *The Citation Culture*. Amsterdam: Universiteit van Amsterdam.

Wouters, P. F. (2017) 'Bridging the Evaluation Gap', *Engaging Science, Technology, and Society*, 3: 108–18.