What is social impact of research and how can it be assessed? A literature survey

Article *in* Journal of the American Society for Information Science and Technology · February 2013

DOI: 10.1002/asi.22803

Lutz Bornmann
Max Planck Society

What Is Societal Impact of Research and How Can It Be Assessed? A Literature Survey

Lutz Bornmann
Division for Science and Innovation Studies, Administrative Headquarters of the Max Planck Society, Hofgartenstr. 8, 80539 Munich, Germany. E-mail: bornmann@gv.mpg.de

Since the 1990s, the scope of research evaluations becomes broader as the societal products (outputs), societal use (societal references), and societal benefits (changes in society) of research come into scope. Society can reap the benefits of successful research studies only if the results are converted into marketable and consumable products (e.g., medicaments, diagnostic tools, machines, and devices) or services. A series of different names have been introduced which refer to the societal impact of research: third stream activities, societal benefits, societal quality, usefulness, public values, knowledge transfer, and societal relevance. What most of these names are concerned with is the assessment of social, cultural, environmental, and economic returns (impact and effects) from results (research output) or products (research outcome) of publicly funded research. This review intends to present existing research on and practices employed in the assessment of societal impact in the form of a literature survey. The objective is for this review to serve as a basis for the development of robust and reliable methods of societal impact measurement.

Introduction

Nowadays, most Organisation for Economic Co-operation and Development (OECD) countries spend approximately 2 to 3% of gross domestic product (GDP) on research and development (R&D) each year (Martin, 2007). Until the 1970s, there was never any doubt in the minds of policymakers that public investment in R&D also would have a positive impact on areas such as communication; the way we work; our housing, clothes, and food; our methods of transportation; and even the length and quality of life itself (Burke, Bergman, & Asimov, 1985). Many countries worked on the principle that "science is the genie that will keep the country competitive, but the genie needs to be fed" (Stephan, 2012, p. 10). In the United States, Bush (1945) argued that any investment in science is inherently good for the society. However, from the late 1980s onward, the empty public coffers increasingly compelled science to account for its accomplishments in the form of internal assessment (otherwise known as *peer review*) and indicators to measure scientific output and scientific impact (the buzzwords being *audit society* and *new public management*). The only aspect of interest when measuring impact was the impact of research on academia and scientific knowledge. The assumption was that a society could derive the most benefit from science conducted at the highest level.

Since the 1990s, there has been a visible trend away from automatic trust in the validity of this assumption; the expectation is that evidence shall be provided to demonstrate the value of science for society (Martin, 2011). What are the results of public investment in research from which society actually derives a benefit (European Commission, 2010)? The scope of research evaluations becomes broader (Hanney, Packwood, & Buxton, 2000; van der Meulen & Rip, 2000) as the societal products (outputs), societal use (societal references), and societal benefits (changes in society) of research come into scope (Mostert, Ellenbroek, Meijer, van Ark, & Klasen, 2010). "What one expects today is measures of the impact of science on human lives and health, on organizational capacities of firms, institutional and group behaviour, on the environment, etc." (Godin & Doré, 2005, p. 5). Society can reap the benefits of successful research studies only if the results are converted into marketable and consumable products (e.g., medicaments, diagnostic tools, machines, and devices) or services (Lamm, 2006). For example, under the Science and Technology for America's Reinvestment: Measuring the Effect of Research on Innovation, Competitiveness and Science program

Received April 30, 2012; revised August 10, 2012; accepted August 10, 2012

 • Published online 17 December 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.22803

(STAR METRICS; https://www.starmetrics.nih.gov/), the effect of federal research grants and contracts on outcomes such as employment and economic activity is traced (Frodeman & Briggle, 2012; Macilwain, 2010).

Primarily, the benefits from basic research have been under scrutiny since the 1990s (Salter & Martin, 2001; C.H.L. Smith, 1997) because perhaps "the importance of fundamental research for . . . society is not fully understood" (Petit, 2004, p. 192). "Research that is highly cited or published in top journals may be good for the academic discipline but not for society" (Nightingale & Scott, 2007, p. 547). R. Smith (2001), for example, suggests that "scientists would think of the original work on apoptosis (programmed cell death) as high quality, but 30 years after it was discovered there has been no measurable impact on health" (p. 528). In contrast, there also is research—for example, "the cost effectiveness of different incontinence pads" (R. Smith, 2001, p. 528)—that is certainly not seen as high quality by the scientific community but has immediate and important societal impact (Nightingale & Scott, 2007).

These changes in how scientific work is judged can, when viewed from the perspective of the sociology of science, be embedded in a changed research landscape which is commonly referred to as "Mode 2," "with a number of alternative accounts of current changes in scientific practice like the 'Academic capitalism,' 'Post-normal science,' 'Triple Helix,' 'Enterprise university,' 'Post-academic science,' 'Strategic research,' 'Finalization science' " (Austrian Science Fund, 2007, p. 27; see also Leydesdorff, 2012). The expression "Mode 2" was coined by Gibbons et al. (1994) to describe some of the fundamental changes occurring in the research system. Whereas "Mode 1" describes a science governed by the academic interests of a specific community (most notably, theory-building), "Mode 2" is characterized by collaboration (both within science and between the scientific world and other stakeholders), transdisciplinarity (several disciplines together studying a real-world problem; Walter, Helgenberger, Wiek, & Scholz, 2007), and (basic) research conducted in the context of its application for the users of the research. (For further characteristics of Mode 2 science in contrast to Mode 1 science, see Table 1 in Erno-Kjolhede & Hansson, 2011.) The shift from science to application and relevant knowledge also leads to a greater mix of "the market and the state, private and public sectors, science and values, producers and users of knowledge" (Barré, 2005, p. 117). While the quality of Mode 1 research is evaluated with respect to excellence and originality, the quality of Mode 2 research is assessed with respect to utilitarian values and criteria (Petit, 2004). In Mode 2, the expectation is that research will produce "socially robust" knowledge (Barré, 2005).

A series of different names have been introduced which refer to the societal impact of research: third stream activities (Molas-Gallart, Salter, Patel, Scott, & Duran, 2002), societal benefits, societal quality (van der Meulen & Rip, 2000), usefulness (Department of Education Science and Training, 2005), public values (Bozeman & Sarewitz, 2011), knowledge transfer (van Vught & Ziegele, 2011), and societal relevance (Evaluating Research in Context [ERiC], 2010; Holbrook & Frodeman, 2011). What most of these names are concerned with is the assessment of (a) social, (b) cultural, (c) environmental, and (d) economic returns (impact and effects) from results (research output) or products (research outcome) of publicly funded research (Donovan, 2011; European Commission, 2010; Lähteenmäki-Smith, Hyytinen, Kutinlahti, & Konttinen, 2006). In this context, (a) *social* benefits indicate the contribution of the research to the social capital of a nation (e.g., stimulating new approaches to social issues, informed public debate, and improved policymaking). (These and the following examples are taken from Donovan, 2008). Since social benefits are hardly distinguishable from the superior term of *societal benefits*, in much literature the term "social impact" is used instead of "societal impact." (b) *Cultural* benefits are additions to the cultural capital of a nation (e.g., understanding how we relate to other societies and cultures, contributing to cultural preservation and enrichment). (c) *Environmental* benefits add to the natural capital of a nation (e.g., reduced waste and pollution, uptake of recycling techniques). (d) *Economic* benefits denote contributions to the economic capital of a nation (e.g., enhancing the skills base, improved productivity). (Further examples of the various areas of societal impact can be found in Department of Education, 2006; Royal Society, 2011; C.H.L. Smith, 1997; United Nations Development Programme, 2010). How these definitions of the different rate-of-return areas for society show societal impact assessment is very broadly conceived: Analysis is required of all significant contributions of science to society—with the exception of recursive contributions to the science itself (van der Weijden, Verbree, & van den Besselaar, 2012).

It is not very easy to separate the aforementioned areas of societal impact from one another. "For example, improvement in the 'quality of life' may depend on a mix of social and cultural studies, environmental research, studies on food safety, health care research, etc." (Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions [SIAMPI], 2011, p. 7). Most notably, economic impact overlaps with the other three areas (Higher Education Funding Council for England, 2009): "There is a fuzzy boundary between the economic and non-economic benefits; for example, if a new medical treatment improves health and reduces the days of work lost to a particular illness, are the benefits economic or social?" (Salter & Martin, 2001, p. 510). As evidenced by the previously discussed examples of the four areas, societal impact is frequently an impact which only becomes apparent in the distant future (ERiC, 2010; Ruegg & Feller, 2003). Thus, societal impact is not a short-term phenomenon; it is mostly concerned with intermediate (e.g., partnership-based cooperation, new/improved products) or ultimate (e.g., improved industry competitiveness) returns (Lähteenmäki-Smith et al., 2006; United States Government Accountability Office, 2012). In the case of cardiovascular research, for instance, Buxton (2011) reported "an average time-lag

DOI: 10.1002/asi

between researchfunding and impacts on health provision of around 17 years" with (p. 260).

What are the other hallmarks of societal impact? It can be anticipated and unanticipated, and it can be inside and outside the target area (Lähteenmäki-Smith et al., 2006). In many cases, societal impact also cannot be limited to a certain country or region (e.g., research on climate change), having a global dimension instead (Royal Society, 2011). Spaapen, Dijstelbloem, and Wamelink (2007) identified three broad groups of stakeholders for societal impact:

> 1. policy makers at the intermediary or government level, whose goal is either to use research for their own policies, or to facilitate the transfer of knowledge from science to society; 2. professional users (profit and nonprofit); that is, industry and societal organizations that want knowledge to develop products and services (this may refer to researchers who profit from developments in other disciplines); 3. end users; that is, the public at large or individual target groups (for example farmers, aids victims). (p. 79)

Given that different stakeholders have their own interests in and expectations of research, it is virtually impossible for two measurements of the societal impact of a certain piece of research to come out the same (Spaapen & van Drooge, 2011).

Although many countries have the wish (and the will) to assess societal impact, "it is not clear how to evaluate societal quality, especially for basic and strategic research" (van der Meulen & Rip, 2000, p. 11). There is "a lack of clearly documented, empirical environmental research impact evaluations" (Bell, Shaw, & Boaz, 2011, p. 227). There is not yet an accepted framework with adequate data sets (comparable to, e.g., the *Web of Science*, Thomson Reuters), criteria, and methods for the evaluation of societal impact (Bensing, Caris-Verhallen, Dekker, Delnoij, & Groenewegen, 2003), and the conventional R&D indicators have revealed little to date (Barré, 2005). In many studies, the societal impact of research is more postulated than demonstrated (Niederkrotenthaler, Dorner, & Maier, 2011). For Godin and Doré (2005), "systematic measurements and indicators on impact on the social, cultural, political, and organizational dimensions are almost totally absent from the literature" (p. 5). When research is conducted in this field, it is primarily concerned with economic impact (Godin & Doré, 2005).

Martin (2007) cited four problems that commonly cause trouble in societal impact assessments: (a) *Causality problem*: It is not clear what impact can be attributed to what cause. (b) *Attribution problem*: Because impact can be diffuse ("impact accretion"; van der Meulen & Rip, 2000), complex, and contingent (Nightingale & Scott, 2007), it is not clear what portion of impact should be attributed to a certain research or to other inputs. (c) *Internationality problem*: R&D and innovation are intrinsically international, which makes attribution virtually impossible. (d) *Evaluation timescale problem*: Premature impact measurement may result in policies that overemphasize research bringing short-term benefits. And there are four other problems, which will be addressed here: (e) If you work with peers instead of indicators to assess societal impact, it is hard to find people to do this. "Scientists generally dislike impacts considerations" (Holbrook & Frodeman, 2011, p. 244) and have problems evaluating research in terms of its societal impact. Such evaluation concerns a wide range of issues (Spaapen et al., 2007) and "takes scientists beyond the bounds of their disciplinary expertise" (Holbrook & Frodeman, 2011, p. 244). (f) Since it can be expected that the scientific work of an engineer has a different impact than the work of a sociologist or historian and because research results affect many different aspects of society (SISOP, 2011), it will hardly be possible to have a single assessment mechanism (Martin, 2011; Molas-Gallart et al., 2002). (g) Societal impact assessment should take into account the fact that there is not just one model of a successful research institution that is valid for the whole world. Assessment should be adapted to the institution's specific focus in teaching and research, the cultural context, and the national standards (Göransson, Maharajh, & Schmoch, 2009; Molas-Gallart et al., 2002; Rymer, 2011; van der Meulen & Rip, 2000). (h) Societal impact of research is not always going to be desirable or positive (Martin, 2011). Furthermore, the same research may well lead to both positive and negative benefits:

> Environmental research that leads to the closure of a fishery might have an immediate negative economic impact, even though in the much longer term it will preserve a resource that might again become available for use. The fishing industry and conservationists might have very different views as to the nature of the initial impact—some of which may depend on their view about the excellence of the research and its disinterested nature. (Rymer, 2011, p. 6)

How should that be taken into account in the assessment?

Viewed as a whole, Godin and Doré (2005) saw research into societal impact assessment as being at the stage where the measurement of R&D was in the early 1960s. In recent years, however, it has become apparent that there is a trend in research evaluation toward giving much more emphasis to the assessment of societal impact, particularly in the health and medical field (Hanney et al., 2000; Holbrook & Frodeman, 2010; United States Government Accountability Office, 2012). For example, a special issue of *Research Evaluation* was recently published which "comprises a collection of papers initially presented to an international, two-day workshop on 'State of the Art in Assessing Research Impact' hosted by the Health Economics Research Group (HERG) at Brunel University in Spring 2011" (Donovan, 2011, p. 175). In May 2007, the Austrian Science Fund (Vienna, Austria) in cooperation with the European Science Foundation (Strasbourg, France) organized the conference "Science Impact: Rethinking the Impact of Basic Research on Society and the Economy" in Vienna (Austria). This review article intends to present existing research on and practices employed in the assessment of societal impact in the form of a literature survey. Other short overviews were published by Frank and Nason (2009; focusing on the area of health research) and Bornmann (2012). The objective is for this overview to serve as a basis for the development of a conceptual framework and the identification of a potentially fruitful research agenda for societal impact measurement in upcoming (empirical) studies.

According to Frank and Nason (2009), the best method of measuring return on investments (in health research) should be "feasible, not too labour intensive, and economically viable. It should be as accurate and responsive as possible within a reasonable evaluation budget that should represent a small percentage of the money invested in the research being assessed" (p. 531). Empirical studies searching for this method should be designed on the basis of a broad knowledge of the literature.

The literature research for this survey was conducted at the end of 2011. A systematic search of publications of all document types (journal articles, monographs, collected works, etc.) was performed using computerized literature databases (e.g., *Web of Science* and Scopus) and Internet search engines (e.g., Google). If a researched paper formulated an overview of the status quo in research (e.g., Hanney et al., 2000), the listed literature was viewed. Approximately 100 publications were considered for this article.

Societal Impact Assessment in Research and Application

This section presents a range of activities which have addressed the issue of societal impact assessment in recent years. These activities include (a) actions which have been (or are being) implemented in national evaluation systems to measure the societal impact of research, (b) actions which research-funding organizations have introduced in the interests of determining the societal impact of planned research projects, (c) research projects which have attempted to assess societal impact (e.g., in the field of health science), and (d) research projects which have been concerned with possible ways of measuring societal impact (e.g., by developing indicators). Consequently, this article presents activities conducted in the research on societal impact and in the practice of societal impact assessment. Before these activities are presented in the next four sections, a number of points will be described which should generally be taken into consideration in societal impact assessment.

Bozeman and Sarewitz (2011) defined research evaluation as "any systematic, data based (including qualitative data) analysis that seeks as its objective to determine or to forecast the social or economic impacts of research and attendant technical activity" (p. 8). As in the assessment of scientific impact, the objective of any societal impact evaluation is to end up with generally accepted and reliable indicators. The indicators should definitely exhibit a coherence between what is measured and what is supposed to be measured (Barré, 2005). For the acceptance of indicators, it is important that they can be used in a variety of contexts and are easily understood by stakeholders. It is particularly desirable to have simple indicators which are less labor-intensive yet provide meaningful assessments (Molas-Gallart et al., 2002), such as (a) relative investment in basic research (science-based R&D expenditure as a percentage of GDP/relative output of scientific papers [May, 1998]) as a measure of the relative effectiveness of different countries' investment in R&D, or (b) patent activities "for measuring the technological orientation of a public research institution or a university" (Barré, 2005, p. 128). However, since societal impact is a complex phenomenon which "involves a wide variety of direct and indirect non-linear and self-reinforcing activities" (Gregersen, Linde, & Rasmussen, 2009, p. 152), it will be difficult to develop simple indicators that apply across disciplinary boundaries "without overmuch reducing the richness and complexity" (van der Meulen & Rip, 2000, p. 12). It is unlikely to be possible (notwithstanding the proposal along these lines which is presented later) to develop a "societal" counterpart to the (Social) Science(s) Citation Index by Thomson Reuters (Bensing et al., 2003). Societal impact assessment will doubtless always need to pursue a holistic approach which examines a number of channels that bind research to the rest of society (Molas-Gallart et al., 2002). Studies have identified up to 60 different indicators as necessary to tackle the task of societal impact evaluation comprehensively (Martin, 2011).

Which additional points should be taken into consideration in societal impact assessment? For R. Smith (2001), "the instrument should (a) fit with current ways of evaluating research, (b) look to the future also, (c) be efficient for both assessors and the assessed, and (d) work in practice" (p. 528; see also van der Meulen & Rip, 2000). Godin and Doré (2005) noted that an assessment of the impact of research should refer to several precise scientific results. The people (scientists) who assess the impact should be familiar with the principal sectors and potential users of these results. Since the absence of societal impact of research is not necessarily a sign of "uselessness," it is important to determine not only the impact but also the conditions, the context, and the efforts of an institution to achieve impact (Godin & Doré, 2005). For different research institutes to be comparable in the context of societal impact measurement, normalization will be necessary, as in scientific impact measurement. Molas-Gallart et al. (2002), for example, argued the case that "indicators could be divided by the total number (full-time equivalent) of faculty employed by the university" (p. 53). As a rule, the process of indicator development will be a process "involving many stages of indicator definition, testing, piloting, validation, collection and analysis. . . . It is unlikely that a robust and standardised system of measurement to collect data could be developed immediately. Measurement requires time, investment and commitment" (Molas-Gallart et al., 2002, p. 47).

In the studies conducted in recent years which have concerned themselves with the assessment of societal impact,

DOI: 10.1002/asi

"three main methodological approaches have been adopted: 1. econometric studies; 2. surveys; and 3. case studies" (Salter & Martin, 2001, p. 513). According to Donovan (2008), case studies represent the last of the stages currently employed in the methodical approach to measuring impact. In evaluation practice, these case studies generally feature as "instances of success stories (best practices)" (SIAMPI, 2010, p. 2) which are added to a self-assessment report and so on in the evaluation of an institution. However, whereas Godin and Doré (2005) considered case studies as generally "very imperfectly" suited to impact assessment, Bell et al. (2011) described the advantages and disadvantages of case studies as follows: "While case-study approaches have been criticized as lacking objectivity and quantification, key benefits include their ability to engage with complexity . . . and the detailed, in-depth understandings gained about events or initiatives over which the researcher has little or no control" (p. 228).

National Evaluation Systems

This section presents a number of national evaluation systems in which social impact assessment is going to be applied or is already applied.

The Netherlands. According to Donovan (2008), one of the most developed examples of impact evaluation to date has occurred in the Netherlands. This is a country in which evaluation is conducted alongside quality assessment (based on self-assessments and site visits) and is primarily focused on the economic value of publicly funded research. According to van der Meulen and Rip (2000), the Netherlands has gone quite far both in thinking about societal impact assessment and in applying tools in actual evaluation practices. The authors consider the Dutch model to be robust enough to be practiced in other countries as well. As demonstrated by a study of documents from evaluation processes in the Netherlands, more than 80% of the evaluations included a societal impact assessment (van der Meulen & Rip, 2000). At best, two of the three dimensions of societal impact described later were considered in an evaluation: "The main dimensions used are the expectation that the research will contribute to socio-economic developments (relevance), the interaction with (possible) users or other societal actors and the actual use of the results" (van der Meulen & Rip, 2000, p. 19).

Common guidelines for the evaluation and improvement of research and research policy are determined—based on expert assessments—for a certain period in so-called *Standard Evaluation Protocols* (SEP) in the Netherlands (Mostert et al., 2010). SEP 2009 to 2015 is now the fourth Protocol (Royal Netherlands Academy of Arts and Sciences, 2010) also to contain the social, economic, and cultural impact of research as evaluation criteria. According to the Royal Netherlands Academy of Arts and Sciences (2010), an institute can be evaluated based on a maximum of three aspects:

- Societal quality of the work. This aspect refers primarily to the policy and efforts of the institute and/or research groups to interact in a productive way with stakeholders in society who are interested in input from scientific research. It may also refer to the contribution of research to important issues and debates in society.
- Societal impact of the work. This aspect refers to how research affects specific stake-holders or specific procedures in society (for example protocols, laws and regulations, curricula). This can be measured, for example, via charting behavioural changes of actors or institutions.
- Valorisation of the work. This aspect refers to the activities aimed at making research results available and suitable for application in products, processes and services. This includes activities regarding the availability of results and the interaction with public and private organisations, as well as direct contributions such as commercial or non-profit use of research results and expertise. (p. 10)

The three aspects are very universal, and there are scarcely specific proposals formulated for measuring societal impact.

Many Dutch organizations involved in quality assurance cooperate in the ERiC project, which has set itself the goal of developing methods for societal impact assessment (ERiC, 2010). The ERiC project is closely linked with the international SIAMPI project (funded under the European Commission's Seventh Framework Programme, discussed later). One significant result which has emerged from the ERiC project is that productive interaction is a necessary requirement for research to have a societal impact: "There must be some interaction between a research group and societal stakeholders" (ERiC, 2010, p. 10). Such interactions can be in the form of personal contact (e.g., joint projects or networks), publications (e.g., educational reports), and artifacts (e.g., exhibitions, software, or websites) and can occur during the research process or after the end of the research work. However, with this specification of productive interactions, the difference between interaction and research outcome is not completely clear (with the exception of personal contacts). In any case, it can be assumed that productive interactions inevitably lead to a societal impact of the research (ERiC, 2010). Thus, societal impact assessment could be limited to the activities of a research unit in this respect. Moreover, it transpired in the course of the ERiC project that assessors of societal impact (academic peers or stakeholders) must be able to understand both the societal context of research and the importance of research for the targeted sector (ERiC, 2010). Academic peers should not, therefore, be selected solely on the basis of their scientific expertise.

Besides the ERiC project, there have been other projects conducted in the Netherlands in recent years which have aspired to improve societal impact assessment in certain research fields by proposing certain methods. For example, a pilot study was conducted at the Leiden University Medical Center to develop methods of societal impact measurement on the research-group level. Mostert et al. (2010) "define three types of societal sectors (or stakeholders): the general (lay) public, healthcare professionals and the private sector.

DOI: 10.1002/asi

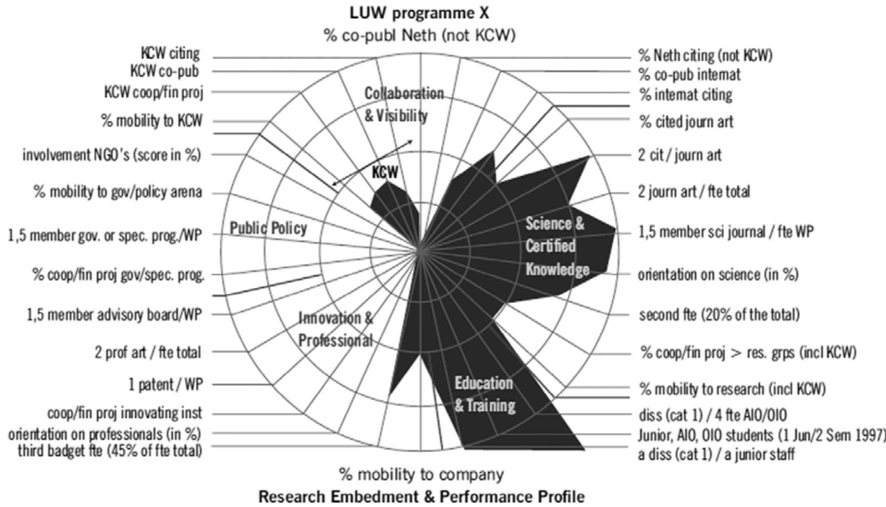

FIG. 1. Example of a Research Embedment and Performance Profile (REPP). Source: Spaapen et al. (2007, p. 67).

In these sectors respectively cultural, social and economic outputs of research will appear." This study demonstrated, for instance, that the correlation between total societal quality scores and scientific quality scores is weak. It showed that

> high scientific quality of research groups is not necessarily related to communication with society, and that in order to increase societal quality of research groups, additional activities are needed. Therefore societal quality is not simply the consequence of high scientific quality. Obviously, in a university medical centre, scientific quality prevails, and is a prerequisite, which cannot be replaced by aiming instead for high societal quality (Mostert et al., 2010).

Although in the medical area the correlation between societal and scientific impact might be small, higher correlation coefficients in other areas (e.g., engineering or social sciences) can be expected.

Spaapen et al. (2007) developed new ways of presenting research work to evaluation committees in the context of policy and societal questions. They proposed what is referred to as the *Research Embedment and Performance Profile* (REPP), in which a variety of criteria and indicators relating to a research unit can be depicted in a single graphic representation of five dimensions of research. The five dimensions are as follows: (a) science and certified knowledge, (b) education and training, (c) innovation and professionals, (d) public policy and societal issues, and (e) collaboration and visibility. Figure 1 depicts an example of an REPP.

United Kingdom. The world's best-known national evaluation system is undoubtedly the U.K. Research Assessment Exercise (RAE), which has been comprehensively evaluating research in the United Kingdom since the 1980s. Efforts are currently under way to set up the Research Excellence Framework (REF), which is set to replace the RAE in 2014 "to support the desire of modern research policy for promoting problem-solving research" (Erno-Kjolhede & Hansson, 2011, p. 140). In the scope of the REF, "the impact element will include all kinds of social, economic and cultural benefits and impacts beyond academia, arising from excellent research" (Higher Education Funding Council for England, 2011, p. 4). The approach to assessing "impact" in the REF was based on the approach developed for the Australian Research Quality Framework (RQF), which Grant, Brutscher, Kirk, Butler, and Wooding (2009) recommended as best practice in their report to the Higher Education Funding Council for England (HEFCE). To develop the new arrangements for the assessment and funding of research in the REF, the HEFCE commissioned RAND Europe to review approaches to evaluating the impact of research (Grant et al., 2009).

Not only will impact be measured in a quantifiable way in the new REF but expert panels also will review narrative evidence in case studies supported by appropriate indicators (informed peer review) (Erno-Kjolhede & Hansson, 2011; Higher Education Funding Council for England, 2009). "Case studies may include any social, economic or cultural impact or benefit beyond academia that has taken place during the assessment period, and was underpinned by excellent research produced by the submitting institution within a given timeframe" (Higher Education Funding Council for England, 2011, p. 1). The preference for a case-study approach in the REF is considered by Donovan (2011) to be "the 'state of the art' to provide the necessary evidence-base for increased financial support of university research across all fields" (p. 178). According to Erno-Kjolhede and Hansson (2011), the new REF is

> a clear political signal that the traditional model for assessing research quality based on a discipline-oriented Mode 1 perception of research, first and foremost in the form of publication in

DOI: 10.1002/asi

international journals, was no longer considered sufficient by the policy-makers. (p. 139)

The new model also will entail changes in budget allocations. The evaluation of a research unit for the purpose of allocations will be 20% determined by the societal influence dimension (60% by research output and 15% by environment; Erno-Kjolhede & Hansson, 2011). The final REF guidance (Higher Education Funding Council for England, 2012) contains lists of examples for different types of societal impact. For example, "policy debate on climate change or the environment has been influenced by research" (p. 29) and "quality of life in a developing country has improved" (p. 29) are examples of the assessment of impact on the environment. However, these lists point out that the societal impact indication is arranged with many possible, frequently less specific options to measure impact. As a result, a comparative assessment of different research institutes by reviewers will be made difficult, and subjective influences will have a greater meaning in research assessments.

Other countries. We will now present certain other nationwide activities in the sphere of societal impact assessment. The activities in the different countries show that societal impact is similarly understood, but assessed differently.

Given its influence on the design of the REF, this overview on activities in other countries starts with the Australian RQF. A detailed account of the development of "impact" assessment for the RQF is given in Donovan (2008). The RQF was designed as a panel-based exercise which aimed to assess research against the background of excellence criteria and the wider benefits on science. A qualitative and contextual approach was regarded as the best method for the evaluation of research, whereby academic peers and end-users of research played an important role. Information on research was seen as "best derived from context statements, impact statements, case studies, and (where appropriate) relevant quantitative and qualitative indicators" (Donovan, 2008, p. 51). In 2007, the government decided that the RQF should not be continued and was replaced by Excellence in Research for Australia (ERA). Under ERA, research is assessed by research evaluation committees on the basis of indicators of research quality, research volume and activity, research application, and recognition (Australian Research Council, 2011).

In 2006, the Danish Council for Research Policy presented a tool which can be used to assess the quality and relevance of *Danish* research at different levels (individual researcher, group, department, and institution). Besides quality-related indicators (e.g., publications and citation counts), it applies indicators for business-related and overall societal relevance.

> Research relevance describes the combined level of societal significance of the research. Research relevance can be described using two indicators: Business-related relevance: The scope of collaboration with private-sector companies in recent years, where relevant business-related measurements are patents, commercial success, industrial PhDs, formalised collaborations, spin-offs, etc. Overall societal relevance: Description of societal relevance, including contributions by the core area to the education of graduates and PhDs, introduction of new patient treatments, advice, authority tasks, etc. (Danish Council, 2006, p. 6)

Pålsson, Göransson, and Brundenius (2009) described two case studies of *Swedish* universities as "being anchored regionally by performing more or less specific services to the local or regional community" (p. 145). In research evaluations at *Finland's* Aalto University,

> the relevant indicators of societal impact include expert tasks, popularized works, media visibility, external funding relating to research cooperation with non-academic institutions (especially TEKES and EU funding), cooperation with the public and private sector outside academia, patents, start-up companies, etc. (European Commission, 2010, p. 92)

A consortium between five Finnish public research organizations involved in R&D activity developed methods and indicators needed to analyze the socioeconomic impacts of research and development (Lähteenmäki-Smith et al., 2006). The following five dimensions of impact were proposed with certain examples of indicators: (a) impact on economy, technology, and commercialization (e.g., patent applications, entry into new markets); (b) impact on knowledge, expertise, human capital, and management (e.g., improved research methods, strengthened expertise); (c) impact on networking and social capital (e.g., improved networking between research partners, firms, etc.); (d) impact on decision making and public discourse (e.g., participation in legislative and strategy planning); and (e) impact on social and physical environment (e.g., promotion of safety, development of infrastructure) (Lähteenmäki-Smith et al., 2006).

Criteria Used in Grant Peer-Review Processes

Research funding organizations generally employ a process of peer review to evaluate applications for research grants which are submitted (Bornmann, 2011). At the end of the 1990s, the U.S. National Science Foundation (NSF)—the U.S. basic science agency—was one of the first institutions to ask peer reviewers to evaluate not only the intellectual merit of an application but also the broader impact (Holbrook, 2012; Mervis, 2011). Broader impact is understood as follows:

> How well does the activity advance discovery and understanding while promoting teaching, training, and learning? How well does the proposed activity broaden the participation of underrepresented groups (e.g., gender, ethnicity, disability, geographic, etc.)? To what extent will it enhance the infrastructure for research and education, such as facilities, instrumentation, networks and partnerships? Will the results be disseminated

DOI: 10.1002/asi

> broadly to enhance scientific and technological understanding? What may be the benefits of the proposed activity to society? (Holbrook, 2010, p. 218)

In June 2011, the National Science Board issued new draft criteria whereby the connection to national goals is the most obvious change:

> Which national goal (or goals) is (or are) addressed in this proposal? Has the PI [Principal Investigator] presented a compelling description of how the project or the PI will advance that goal(s)? Is there a well-reasoned plan for the proposed activities, including, if appropriate, department-level or institutional engagement? Is the rationale for choosing the approach well-justified? Have any innovations been incorporated? How well qualified is the individual, team, or institution to carry out the proposed broader impacts activities? Are there adequate resources available to the PI or institution to carry out the proposed activities? (Holbrook, 2012, pp. 15–16)

The need for greater public accountability "is changing the nature of *ex ante* peer review at public science agencies worldwide" (Frodeman & Briggle, 2012, p. 3). Social concerns have been more and more included by the consideration of societal impact issues. Holbrook (2010) cited four other institutions which have since incorporated these considerations in a form similar to the NSF's: the National Institutes of Health (NIH), the National Oceanic and Atmospheric Administration, the Natural Sciences and Engineering Research Council of Canada, and the Dutch Technology Foundation. The European Commission Fifth Framework Programme used a total of five review criteria to evaluate research, "three of which concerned societal impacts: (a) community added value and contribution to EU policies; (b) contribution to community social objectives; and (c) economic development and science and technology prospects" (Holbrook & Frodeman, 2010, p. 8). The project Comparative Assessment of Peer Review (http://csid-capr.unt.edu/) examines how these agencies integrate societal impact issues into the grant peer-review process (Frodeman & Briggle, 2012). The Biotechnology and Biological Sciences Research Council (2012) funds research on plants, microbes, animals, and tools and technology underpinning biological research. The Council's perception of impact of research is much broader "than the direct and well understood impact on the economy. It includes scientific advancement, public good, influence on policy, public understanding and enthusiasm for science, and much more" (Biotechnology and Biological Sciences Research Council, 2012, p. 1). Applications for the Australian Research Council Linkage program (this program supports collaborative research projects between industry and university researchers) are not only assessed according to the level of significance and innovation but also the potential national benefits (Australian Research Council, 2012).

As the NSF's experience with assessing the societal impact in the case of applications has shown, it is a problematic feature in the grant-application process. "The concept is difficult to understand, and one that applicants and peer reviewers feel ill-equipped to address" (Holbrook & Frodeman, 2010, p. 3). Scientists often have difficulty making statements about the potential benefits of research or its longer term implications (Holbrook, 2012; Rymer, 2011; Mardis, Hoffman, & MacMartin, 2012). Since peer reviewers for the NSF are required to comment about a proposal only in areas in which they see themselves as experts, these difficulties mean that peer reviewers tend to write mostly about the intellectual impact and include hardly anything about the societal impact. "NSF has recognized this tendency and has—without altering the Broader Impacts Criterion itself—increasingly emphasized that broader impacts must be addressed in both proposals and reviews" (Holbrook & Frodeman, 2011, p. 243). Roberts (2009) conducted the first study on the NSF's broader impact criterion. The findings have suggested that "considering potential societal benefit as a broader impact may not lead to more actual societal benefits and that many potentially useful results may not be disseminated beyond the scientific community" (p. 199). A second study was published by Kamenetzky (in press). The study quantitatively compared the proposed broader impacts of 360 funded abstracts from biology, engineering, and mathematical/physical sciences. The results showed that engineering was significantly more likely to propose potential societal benefits and to partner with potential users than were mathematical/physical sciences or biological sciences.

In response to the fact that applicants face difficulties in depicting the broader impact of the research for which they are applying for a grant, even though depicting it well substantially increases the chances of obtaining funding, a company offers a CD that teaches applicants how "to successfully identify, distill, and communicate your project's broader impacts to NSF reviewers" (Mervis, 2011, p. 169). Since reviewers face difficulties in adequately assessing societal impact, the NSF (and other organizations) should think about the possibility of enhancing the pool of reviewers. For the societal impact assessment, people are required who can assess the societal context of research (e.g., people from research management groups or industry).

Studies Which Assess Societal Impact

Most studies which have assessed the societal impact of research to date have focused on the economic dimension. As far back as the 1950s, economists began to integrate science and technology in their models and study the impact of R&D on economic growth and productivity (Godin & Doré, 2005). Compared with the other dimensions (e.g., the cultural dimension), the economic dimension is certainly the easiest to measure (even though no reliable indicator has yet been developed here, either). Salter and Martin (2001) cited six types of contributions that publicly funded research makes to economic growth: (a) expanding the knowledge available for firms to draw upon in their technological activities; (b) well-educated graduates that flow to firms; (c) scientists develop

DOI: 10.1002/asi

new equipment, laboratory techniques, and analytical methods that are then available to use outside academia; (d) government-funded research is frequently an entry point into networks of expertise; (e) faced with complex problems, industry willingly draws on publicly funded research; and (f) new firms are founded out of scientific projects. It is apparent that some industries (e.g., computers) profit more from research than do others (e.g., metal products; Salter & Martin, 2001) and that some scientific disciplines (e.g., medicine) contribute more toward measurable benefits than do others (e.g., mathematics; SISOP, 2011).

Most studies which measure economic impact examine the relationship between social benefits (in the defined areas; see, e.g., Appleseed Inc., 2003) and society's costs to undertake the research (Committee on Prospering in the Global Economy of the 21st Century, 2007; Link & Scott, 2011). The results of studies conducted to date are described in a summary by Salter and Martin (2001), as follows:

> Few attempts had been made to measure the rates of return to publicly funded research and development. . . . the limited evidence gathered to date indicates that publicly funded basic research does have a large positive payoff, although this is perhaps smaller than the social rate of return on private R&D. (p. 514; see also Stephan, 2012)

Whereas Petit (2004) estimated the rate of return of public and private R&D at 20 to 30% annually, Miranda and Lima (2010) saw "the evolution of major scientific discoveries and impacting technological inventions . . . [as] exponentially correlated to the GDP" (p. 92). A survey of the estimates of rates of return on publicly funded government R&D programs compiled by Salter and Martin (2001) showed that the rates of return are between 21 and 67%. The broad range even indicates that the rate of return cannot be measured uniformly or reliably, respectively. Two extensive studies whose results were recently published found the following in relation to the medical sphere:

> The total health and GDP gains to public/charitable CVD [cardiovascular disease] research in the UK 1975–1992 give a total IRR [internal rate of return] of around 39%. In other words, a £1.00 investment in public/charitable CVD research produced a stream of benefits thereafter that is equivalent in value to earning £0.39 per year in perpetuity. (Health Economics Research Group, 2008, p. 7)
>
> The statistical analysis shows that NIH-funded basic research, potential market size, and industry R&D all have economically and statistically significant effects on the entry of new drugs. The elasticity estimate in the preferred model implies that a 1% increase in the stock of public basic research ultimately leads to a 1.8% increase in the number of new molecular entities (NMEs), an important category of new drug therapies defined by the U.S. Food and Drug Administration (FDA). (Toole, 2012, p. 2)

As a final point, this section presents some important studies which have measured societal impact with certain data (patent data or clinical guidelines) or a certain method (surveys or case studies). In a bid to measure the influence of research on industry, Narin, Hamilton, and Olivastro (1997) studied the frequency with which scientific *publications were cited in U.S. patents*. They evaluated 400,000 U.S. patents issued between 1987 and 1994. They showed that the knowledge flow from U.S. science to U.S. industry tripled in these years. Grant (1999) and Lewison and Sullivan (2008) pursued a similar objective to Narin et al. (1997) with their evaluation of *clinical guidelines*: How does knowledge flow from clinical research to clinical practice? The pilot study by Grant examined three guidelines and was able to ascertain that they contained citations of a total of 284 publications (which can be categorized by author, research institution, country, etc.). Grant's results demonstrated the usefulness of his approach to tracing the flow of knowledge from research funding into clinical practice. Lewison and Sullivan substantially expanded the data base studied by Grant and examined 43 U.K. guidelines for references to papers. They found that

> The UK papers were cited nearly three times as frequently as would have been expected from their presence in world oncology research (6.5%). Within the United Kingdom, Edinburgh and Glasgow stood out for their unexpectedly high contributions to the guidelines' scientific base. (Lewison & Sullivan, 2008, p. 1944)

Evaluating citations in patents and clinical guidelines has the advantages that (a) societal impact can be measured in a similar way to scientific impact (and a sound method of evaluating data is therefore available); (b) the fact that they are citations means that nonreactive, relatively objective, and extensive data are available; and (c) patents and guidelines are available for the evaluation in a relatively freely accessible form and—compared with other data—can be evaluated with a reasonable amount of effort.

Both of the methods (surveys and case studies), presented later for application in societal impact studies, produce data which, while demonstrating barely any of the aforementioned advantages of patents and guidelines, are advantageous in other respects. *Surveys* enable the treatment of a much larger spectrum of subjects than would be possible with patents or clinical guidelines, and they can be more flexibly put to use. Surveys can be used, for example, to canvass citizens' opinions of the impact of research (Bozeman & Sarewitz, 2011) and to poll the views of customers or other stakeholders of research-performing departments or of technology-transfer offices on the importance of research (Rymer, 2011).

The most important (and much-cited) survey in the field of societal impact measurement was presented by Mansfield (1991) in the early 1990s. He asked R&D managers from 76 U.S. firms to estimate the proportion of their products or processes that could not have been developed in the last 10 years without academic research. When analyzed, the survey data produced a figure of around 10%. In a follow-up study at the end of the 1990s involving 70 firms, Mansfield (1998)

DOI: 10.1002/asi

calculated a figure of 15% for new products and 11% for new processes. Beise and Stahl (1999) followed the approach taken by Mansfield (1991) and surveyed 2,300 companies in Germany. They "found that less than one-tenth of product- or process-innovating firms introduced innovations between 1993 and 1995 that would not have been developed without public research" (p. 397).

Case studies enable a detailed analysis of individual projects to explore the ways in which research has produced societal benefits. Case studies do not permit generalizations to be made but they do provide in-depth insight into processes which resulted in societal impact, and therefore lead to a better understanding of these processes (Rymer, 2011). Cost–benefit analyses are very similar to case studies: While they may not be as detailed, they do generally contain more quantitative data (Rymer, 2011). Ruegg and Feller (2003) reported on the results of six detailed case studies which estimated the economic benefit of Advanced Technology Program projects. Three high-performing projects produced an estimated measure of economic benefits greater than $15 billion.

Projects to Find Possible Societal Impact Indicators

Given that societal impact assessment is of ever-increasing importance in the evaluation of research, a number of projects already have been conducted, and longer term initiatives begun, which deal with the question of how societal impact can be assessed. One such project, which has been ongoing for some time now, is An Observatorium for Science in Society based in Social Models (SISOB; see http://sisob.lcc.uma.es/) and is working on defining a "conceptual model" of the social impact of science (SISOP, 2011). No reports or publications from this project have been forthcoming yet. Another project called European Indicators and Ranking Methodology for University Third Mission is a 3-year project funded by the European Commission and developed by partners from eight European countries (http://www.e3mproject.eu/). The aim of this project is to generate an instrument (including indicators) to identify, assess, and compare the societal impact activities of universities. A first report of the project (Centre for Quality and Change Management, 2011) presented the results of a Delphi study (Rowe & Wright, 2001). Using expert opinions, most-significant indicators were identified, such as number of events held by the university that were open to the general public, number of people attending/using facilities (to measure the extent of provision of services by the university), and percentage of a university's budget used for educational outreach.

One of the best known approaches developed to date in a program of assessing research impact is the so-called "Payback Framework." This framework was presented in the mid-1990s by Buxton and Hanney (1994, 1996) from the HERG at Brunel University (United Kingdom) for the sphere of health research and development. According to Donovan (2011), the framework is widely regarded as "best practice" (p. 175) to assess impact. It is a research tool which can compile data on the research impact of a unit that also facilitate comparative cross-case analyses. The framework can be used to depict the complete research process. To this end, the framework makes available a multidimensional categorization of benefits from research, which relates to the following five areas: (a) knowledge, (b) benefits to future research and research use, (c) political and administrative benefits, (d) health-sector benefits, and (e) broader economic benefits. Data for the five categories can be collected through surveys of decision makers and analysis of documents (Hanney et al., 2000). An overview of other (similar) frameworks is given by van der Weijden et al. (2012). What all these frameworks have in common is that they allow a differentiated view of societal impact, but great effort is required to apply them.

In recent years, the Payback Framework has been applied in a range of different contexts both within and beyond the health service (Donovan, 2011). It was used, for instance, for a retrospective evaluation of the benefits of research funded by a regional office of the National Health Service; in collaboration with RAND Europe, the Framework was extended to examine basic and early clinical biomedical research (Donovan & Hanney, 2011). Scott, Blasinsky, Dufour, Mandal, and Philogene (2011) described its application in an evaluation of the Mind–Body Interactions and Health Program; Nason et al. (2011) used the Framework to assess the impacts of the Health Research Board of Ireland's funding activities. Klautzer et al. (2011) tested its applicability to social sciences. As the studies have demonstrated, the Framework could successfully be used in these diverse contexts for the assessment of societal impact.

The SIAMPI project set itself the objective of developing approaches and tools for the evaluation of societal impact that are applicable in different fields and evaluation contexts (SIAMPI, 2011). Case studies were conducted in four fields in the course of the project: (a) information and communications technology, (b) nanotechnology and nanoscience, (c) health and healthcare research, and (d) social sciences and humanities. Analyses of the case studies showed that the key to the successful generation of societal impact lies in the interactions between science (scientists) and society (stakeholders): Whenever there is productive and highly professionalized interaction between stakeholders and scientists, this generally also results in societal impact. So scientists do not transfer the knowledge that they generate to society themselves; rather, societal impact happens on the basis of iterative processes among researchers and research stakeholders. Productive interactions may be (a) direct interactions (direct contacts), (b) indirect interactions through some kind of material "carrier" (e.g., exhibitions, models, or films), and (c) financial interactions (economic exchange) (SIAMPI, 2010; Spaapen & van Drooge, 2011). Molas-Gallart and Tang (2011) considered the SIAMPI approach to be highly suited to the social sciences in particular, given that research here has often originated from complex social, scientific, and political processes. However, it is doubtful

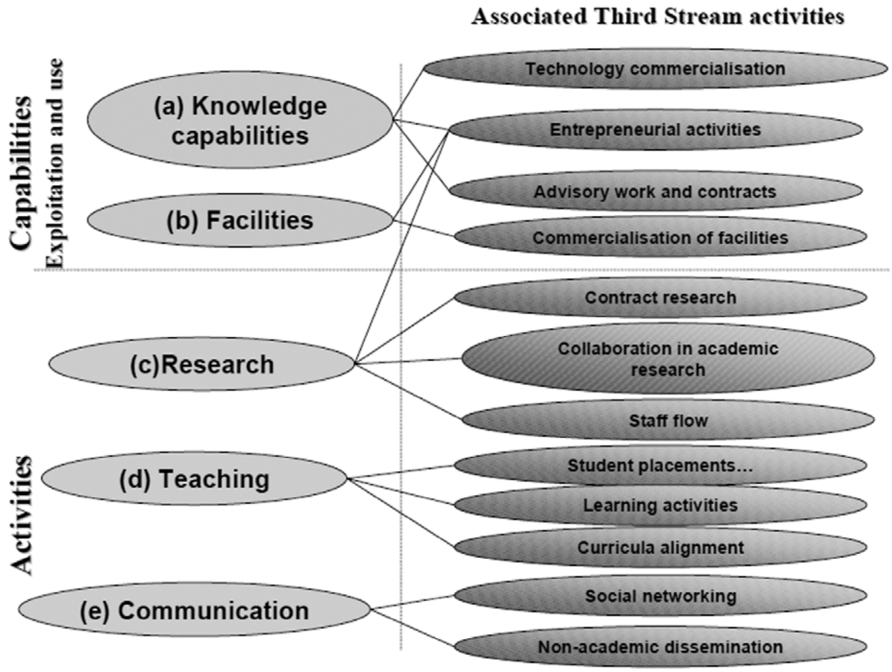

FIG. 2. Conceptual framework for analyzing third-stream activities. Source: Molas-Gallart et al. (2002, p. 21).

whether it is that simple: Further studies are needed to investigate whether productive and highly professionalized interaction between stakeholders and scientists generally result in societal impact.

Molas-Gallart et al. (2002) published an analytical framework and a comprehensive set of indicators with which to assess third-stream activities. In developing the indicators, the authors strove to make available so-called *SMART metrics*, which are simple, measurable, actionable, relevant, reliable, and reproducible as well as timely. Figure 2 contains a graphical representation of the conceptual framework of Molas-Gallart et al. (2002). On the right of the figure, all third-stream activities are listed which can be measured by indicators in an evaluation. Because they are so many in number, only a selection of these activities is further described: Activities which lead to *technology commercialization* relate "to the exploitation of intellectual property, codified in clearly identifiable information packages that can be the object of commercial transactions" (Molas-Gallart et al., 2002, p. 21). The best known activities undertaken by universities in this area are patenting activities. *Entrepreneurial activities* lead to the establishment of new firms to exploit existing university capabilities (e.g., joint ventures, spin-offs, start-ups, and incubators). The category of *advisory work and contracts* includes consulting work offered by scientists outside academia (e.g., on the grounds of scientists' familiarity with certain research tools). For each third-stream activity depicted in Figure 2, Molas-Gallart et al. (2002) specified a number of possible indicators for their assessment (see the indicators in Figure 3 as examples of the three third-stream activities mentioned). As Figure 3 shows, Molas-Gallart et al. (2002) also described how the data necessary for the indicators can be collected and what the associated costs of collection are. All in all, the framework presented by Molas-Gallart et al. (2002) appears to be very well-suited as a means of enabling the multidimensional phenomenon of societal impact to be assessed on the basis of specific indicators. But such comprehensiveness comes at a considerable cost, raising the question of whether such an assessment is actually worthwhile. Compared to usual scientific impact evaluations, the use of the framework for societal impact evaluations would mean a significantly greater effort.

Having so far presented some of the larger and better known projects dealing with societal impact assessment, we will now touch upon some of the smaller initiatives.

To identify the areas in which societal impact can be assessed and the indicators which can be applied to measure the impact, Godin and Doré (2005) conducted a series of interviews with scientists from publicly funded research centers and with actual and potential users of research results in social and economic organizations. Based on the results of their survey, they constructed a typology of societal impact with 11 dimensions to which they also add several possible indicators: (a) science (e.g., research activities), (b) technology (e.g., services), (c) economy (e.g., investments), (d) culture (e.g., attitudes), (e) society (e.g., welfare), (f) policy (e.g., national security), (g) organization (e.g., planning), (h) health (e.g., health system), (i) environment (e.g., climate and meteorology), (j) symbolic (e.g., notoriety), and (k) training (e.g., pedagogical tools). Since the list of dimensions and possible indicators is superficial, it is scarcely suitable for practical application.

DOI: 10.1002/asi

Indicator	Data collection instrument	Collection costs
Technology commercialisation		
No. of patent applications.	Technology Commercialisation Offices may gather data at some universities. At other universities such information may be held by central administration or at the departmental level.	Moderate
No. of patents awarded.		
No. of licences granted (including option agreements).		
Royalty income (including option fees).		
Median value of royalties (including option fees).		
Entrepreneurial activities		
No. of spin-offs created in the last 5 years.	Technology Commercialisation Offices may gather data at some universities. Elsewhere information may be held by central administration or by departments.	Moderate
No. of current employees in spin-offs created in the last 5 years.		
Turnover/profits from spin-offs and commercial arms.		
Development funds and loan facilities provided by universities to support start-ups	Technology Commercialisation Offices may gather data at some universities. These are metrics under development in institutes receiving funds from the Science Enterprise Challenge	Moderate
Advisory work		
No. of invitations to speak at non-academic conferences (excluding project presentations to funders).	Information could be collected as a part of the annual appraisal process or through a survey.	Medium
No. of invitations to attend meetings of advisory committee of non academic organisations.		

FIG. 3. Possible indicators to be included in a third-stream measurement model. Source: Molas-Gallart et al. (2002, p. 49).

A very practical approach to societal impact assessment was selected by Kuruvilla, Mays, Pleasant, and Walt (2006) when they developed categories for the health area that "prompt researchers to systematically think through and describe the impact of their work." The seven categories are (a) knowledge, attitude, and behavior impacts; (b) health literacy; (c) health status; (d) equity and human rights; (e) related to the economy/macroeconomic; (f) social capital and empowerment; and (g) culture and art. These categories open up possibilities for the assessment of societal impact which nevertheless have to be developed by the user.

Niederkrotenthaler et al. (2011) pursued the objective of developing a societal impact factor (SIF) which could be calculated in addition to the journal impact factor (JIF) of Thomson Reuters. (The JIF indicates the average citation impact of journals.) Contrary to the JIF, the SIF is based on an author's self-assessment of his or her publication. The questions about a publication which an author needs to answer in order to identify its SIF were developed on the basis of focus-group discussions and relate to (a) the aim of a publication (gain of knowledge, application of knowledge, and increase in awareness) and (b) "the specific efforts of the authors to translate their research results into societal action, and, if translation was accomplished, the size of the translation" (Niederkrotenthaler et al., 2011). Since the SIF rests upon subjective self-assessment, it is doubtful whether it will gain acceptance in research assessment.

To end this section, two projects will be presented which did not have the specific aim of defining indicators of societal impact at their core but which did deal with societal impact assessment in the context of university rankings. In both projects, scientists set themselves the goal of considering a more broad-based assessment of university performance beyond the conventional measurement of scientific impact as presented in rankings, such as the Shanghai ranking (http://www.arwu.org/) and the Leiden ranking (http://www.leidenranking.com/). Both projects are designed in such a way that they give an overview of possible indicators for application in university rankings. The next step would be to condense the indicators into a smaller set which could actually be used in practice.

In 2008, the European Commission put together an expert group tasked with identifying the parameters by which research produced by European universities could be evaluated (in a university ranking; European Commission, 2010). The group comprised 15 members from 12 European Union Member States, Australia, a European association, and an international organization, and formulated, among other things, proposals for how the societal impact of research could be assessed in four different areas: (a) economic benefits (e.g., adding to economic growth and wealth creation), (b) social benefits (e.g., improving people's health and quality of life), (c) environmental benefits (e.g., improvements in environment and lifestyle), and (d) cultural

DOI: 10.1002/asi

benefits (e.g., stimulating creativity within the community). The appendix to the report of the European Commission (2010) contains an extensive survey of possible indicators that can be used to assess societal impact (e.g., in the area of employability of PhD graduates or scientific partnerships and collaborations).

The U-Multirank project pursued the goal of creating a multidimensional global university ranking (http://www.u-multirank.eu/). Once a multidimensional ranking system for higher education institutions had been designed, the feasibility of the system was tested on a sample of 150 higher education and research institutions. The results of the test were presented at a conference in June 2011. When designing the ranking system, scientists also defined areas for the assessment of knowledge transfer which consider the contribution of the research to economic, social, and cultural development. For institutional rankings, these areas are (a) incentives for knowledge exchange, (b) third-party funding, (c) university–industry joint publications, (d) patents, (e) size of technology-transfer office, (f) continuous professional development courses offered, (g) co-patents, and (h) number of spin-offs (van Vught & Ziegele, 2011). For each area, van Vught and Ziegele (2011) specified a precise definition and possible ways in which it can be assessed.

Discussion

Mode 2 research is characterized by an environment of increasing financial stringency in which science (like many other areas of society) must prove that government investments in R&D have been effective and employed for the benefit of society. In Mode 2, with its requirements for counting and documenting, it is of vital importance for each research institution to come to grips with the assessment of societal impact in all areas of research (Erno-Kjolhede & Hansson, 2011): “Research organisations have to make serious efforts to gather more robust data on social impact and on research output and outcome to wider audiences” (SIAMPI, 2011, p. 7). This paper presents a number of ways in which the societal impact of research can be assessed (and a number of difficulties which can arise in so doing). If research institutions disregard this issue, they run the risk of finding that their research is not sufficiently visible and well-documented in research assessments. “The absence of readily available measures of impact can downplay the importance of some research that aims to achieve direct impact, especially when such research does not result in publications in high impact journals” (Rymer, 2011, p. 4).

The development of approaches to assess societal impact can be divided into three phases (Donovan, 2007). First, efforts have been made to record data to primarily provide information on the economic impact of research (e.g., the impact of research on industry; discussed earlier). Other types of impact, which today also are understood as societal impact, have hardly been taken into account in the first phase. Donovan (2008) refers to the first phase as *technometrics*. The second phase of societal impact assessment, which Donovan (2008) calls *sociometrics*, is not just restricted to economic impact but also covers the social impact of research in the more local regional environment. The study by Pålsson et al. (2009) is a good example of this. Not all types of societal impact are taken into account in the second phase either, and the view of research impact is focused on a regional level. The third phase of impact assessment is characterized by the case-study approach. A range of indicators (quantitative and qualitative) are typically linked in this approach to provide a comprehensive picture of all types of societal impact. The benefit of the case-study approach is that it records the complexity of societal impact in the range of academic research from the creative arts to natural sciences in a differentiated way. The disadvantages are that the assessment of the impact is very expensive and that the societal impact of different institutions can hardly be compared. For comparison purposes, a uniform approach (e.g., with the same indicators) to the assessment of impact would be required.

Approaches were presented earlier which are less expensive and ensure a uniform procedure in impact assessment. Evaluation results of two different research units based on citations in clinical guidelines or patents can be directly compared. Further advantages of these approaches are that it (a) can be deployed at various aggregation levels (larger and smaller research units) and (b) has similarities to the scientific impact evaluation and will therefore more easily gain acceptance. The disadvantage of these approaches is that it can be used only in a small number of fields and only provides very restricted information on the impact of research. However, even with extremely simple approaches such as the measurement of citation counts, the fact that a sophisticated method of statistical evaluation is required should not be overlooked. Citation counts have to be standardized to make them comparable, arithmetic averages should not be calculated because citation data are generally skewed, and there is a range of factors which influences citation counts and should be taken into account in an evaluation (e.g., the scope of a publication; Bornmann, Mutz, Neuhaus, & Daniel, 2008).

If many different indicators are used in the societal impact evaluation, similar methodological requirements as with the citation counts have to be taken into account with each indicator. Additional indicators (studies have identified up to 60 different indicators; discussed earlier) also mean that additional aspects have to be taken into account in the collection of the data. Since studies have shown that scientific and societal impact barely correlate with one another (discussed earlier), few synergy effects will be produced, and both evaluations have to be conducted alongside one another. Approaches are needed which combine the scientific and societal impact assessment in a single applicable framework. For example, Schapper, Dwyer, Tregear, Aitken, and Clay (2012) developed an evaluation methodology which is oriented to practicability to assess health research performance “across three broad categories: knowledge

DOI: 10.1002/asi

creation; inputs to research; and commercial, clinical and public health outcomes" (p. 218).

However, unlike scientific impact measurement, for which there are numerous established methods that are continually being refined in a distinct community, research into societal impact assessment is still in the early stages: There is no distinct community with its own series of conferences, journals, or awards for special accomplishments in this field. Even though robust and reliable methods for measuring societal impact have not yet been developed (Gregersen et al., 2009; Salter & Martin, 2001), budget-relevant societal impact assessments already are being conducted (or planned). This imbalance between research and practice is astonishing, given how long the methods of scientific impact measurement have had the opportunity to evolve before reaching the stage of being employed in budget-relevant practice. The field of scientific impact measurement was able to evolve gradually: Simple indicators, such as publication and citation counts, were initially employed by a few interested parties who were looking into the measurement of science, before a distinct community with its own specialists (bibliometricians) developed in this field. It was only years later that scientific impact assessments began to be used on a budget-relevant basis in a larger way.

More so than with scientific impact measurement, the assessment of societal impact research is badly needed: As the literature survey in this article has shown, societal impact is much harder to assess than is scientific impact. This research should especially address the various problems which are associated with impact evaluation (e.g., the causality and attribution problems; discussed earlier). The recently published study of van der Weijden et al. (2012) is a good example of the kind of studies required. Using a survey, they empirically investigated how the quest for societal benefits is taken up by the principal investigators. Furthermore, we need expert surveys on the importance of criteria for the evaluation of societal impact in different fields (see Montada, Krampen, & Burkard, 1999). In the societal impact area, there are unlikely to be any indicators such as publication and citation counts which can be employed across almost all disciplines and institutions and which can be relatively easily researched in databases (Martin, 2011). Societal impact of research often takes many years to become apparent, and it is in many cases difficult to identify causality between a certain piece of research and a certain impact: "The routes through which research can influence individual behaviour or inform social policy are often very diffuse" (Rymer, 2011, p. 10). Many studies which have carried out societal impact assessment to date chose to do so on the basis of case studies. Although this method is very labor-intensive and very much a "craft activity" (Martin, 2011), it appears to be the best way of measuring the complex phenomenon that is societal impact.

But it is not only the practitioners of societal impact assessment who should grapple with this form of impact by conducting their own research: The scientists (the people being evaluated) should do so as well. According to Hanney et al. (2000), many scientists see societal impact assessment as a threat to their scientific rigor and therefore reject the idea. However, given that societal impact assessment already plays an important role in some national evaluations (and other countries will surely soon follow), all scientists should be concerned with the social benefits of their research. As it turns out, scientists are often unaware that their research has a societal impact. "The case study at BRASS [Centre for Business Relationships, Accountability, Sustainability and Society] uncovered activities that were previously 'under the radar,' that is, researchers have been involved in activities they realised now can be characterized as productive interactions" (SIAMPI, 2010, p. 2). In fact, we can assume that research in many fields already is generating direct societal impact (or inducing productive interactions), but that this is not yet perceived as such by the scientists.

There already are some promising approaches toward the assessment of societal impact, such as the frameworks of Buxton and Hanney (1994, 1996) and Molas-Gallart et al. (2002), upon which further research in this sphere could build. Such research should be as broadly based as possible—similar to the approaches specified in the two frameworks—to identify appropriate indicators for different disciplines while also developing mechanisms to collect accurate and comparable data (European Commission, 2010). If the approaches in this sphere are not sufficiently broad in scope, there is a danger of readily available indicators being used for evaluations even though these fail to adequately measure societal impact (Martin, 2007). In addition, there also is a danger of scientists basing their actions on the readily available indicators: "The obvious danger is that researchers and universities intensify their efforts to participate in activities that can be directly documented rather than activities that are harder to document but in reality may be more useful to society" (Erno-Kjolhede & Hansson, 2011, p. 136). Numerous studies already have documented the fact that scientists actually do base their actions on the criteria and indicators applied in evaluations (Abbott et al., 2010; Bornmann, 2010; Erno-Kjolhede & Hansson, 2011). Until future research on societal impact has developed reliable and robust methods, it clearly makes sense to have the societal relevance of research qualitatively assessed by expert panels in the first instance: "Just as peer review can be useful in assessing the quality of academic work in an academic context, expert panels with relevant experience in different areas of potential impact can be useful in assessing the difference that research has made" (Rymer, 2011, p. 12). For example, the NIH examines grant proposals in light of the mission to improve health by using scientists from the extramural research community and public representatives (Frodeman & Briggle, 2012). Indicator-based analyses should be made available to the panels as a basis for decision making. The advantage of peers is that they can be employed at various evaluation

DOI: 10.1002/asi

levels of aggregation (e.g., peer review of grant applications, individual performance review, and national research-evaluation exercises).

The qualitative assessment of societal impact by peers should not be dominated exclusively by scientists. Experience from the field of grant peer review (discussed earlier) has shown that scientists often have trouble discerning the societal impact of research. The evaluations therefore also should involve stakeholders (particularly policymakers and professional users) with corresponding experience in the exploitation of research. The SIAMPI project (SIAMPI, 2011) has demonstrated that productive interactions between scientists and stakeholders are especially suited to guaranteeing that research has a societal impact. Having peer-review groups comprising both scientists and stakeholders at the sites being evaluated is sure to intensify these kinds of interactions. Furthermore, the efforts that researchinstitutions are expected to make in the area of societal impact in the coming years are bound to lead to the emergence of a new profession of administrative officials (as was the case with "scientific" research evaluations). These people will help to bring about productive interactions between scientists and stakeholders—with the objective of both creating and evaluating societal impact.

Although the evaluation of research is important when it comes to making investment decisions in research and ensuring the effectiveness of the national innovation system, note that research is, by its very nature, dealing with what we do not know, "so that the significance of its outputs are to a greater or lesser degree unpredictable—which is why serendipity is always important" (Rymer, 2011, p. 18). According to Molas-Gallart et al. (2002), returns to innovations are highly skewed: One should expect only 1 in 10 research projects to be successful. Research evaluations should not create a situation in which predominantly "normal" science is conducted which, while offering the scope to be easily planned and well-documented, hardly leaves any place for serendipity (Ziman, 2000). The assessment of societal impact is an additional method of scientific impact measurement to evaluate research, which further increases the administrative effort involved in evaluations. That is why we should seek above all to ensure that societal impact assessment can be undertaken with as little effort as possible for the scientists involved.

Acknowledgments

I express my gratitude to two anonymous reviewers for their helpful comments.

References

Abbott, A., Cyranoski, D., Jones, N., Maher, B., Schiermeier, Q., & van Noorden, R. (2010). Metrics: Do metrics matter? Nature, 465, 860–862.

Appleseed Inc. (2003). Engines of economic growth: The economic impact of Boston's eight research universities on the metropolitan Boston area. Boston, MA: Author.

Australian Research Council. (2011). ERA 2012 submission guidelines, Excellence in Research for Australia. Canberra, Australia: Author.

Australian Research Council. (2012). Linkage Projects funding rules for funding commencing in 2012 variation (No. 1). Canberra, Australia: Author.

Austrian Science Fund. (2007). Rethinking the impact of basic research on society and the economy. Vienna, Austria: Author.

Barré, R. (2005). S&T Indicators for policy making in a changing science–society relationship. In H. Moed, W. Glänzel, & U. Schmoch (Eds.), Handbook of quantitative science and technology research (pp. 115–131). Dordrecht, The Netherlands: Springer.

Beise, M., & Stahl, H. (1999). Public research and industrial innovations in Germany. Research Policy, 28(4), 397–422. doi:10.1016/s0048-7333(98)00126-7

Bell, S., Shaw, B., & Boaz, A. (2011). Real-world approaches to assessing the impact of environmental research on policy. Research Evaluation, 20(3), 227–237.

Bensing, J.M., Caris-Verhallen, W.M.C.M., Dekker, J., Delnoij, D.M.J., & Groenewegen, P.P. (2003). Doing the right thing and doing it right: Toward a framework for assessing the policy relevance of health services research. International Journal of Technology Assessment in Health Care, 19(04), 604–612. doi:10.1017/S0266462303000564

Biotechnology and Biological Sciences Research Council. (2012). BBSRC policy on maximising the impact of research. Swindon, United Kingdom: Author.

Bornmann, L. (2010). Mimicry in science? Scientometrics, 86(1), 173–177.

Bornmann, L. (2011). Scientific peer review. Annual Review of Information Science and Technology, 45, 199–245.

Bornmann, L. (2012). Measuring the societal impact of research. EMBO Reports, 13(8), 673–676.

Bornmann, L., Mutz, R., Neuhaus, C., & Daniel, H.-D. (2008). Use of citation counts for research evaluation: Standards of good practice for analyzing bibliometric data and presenting and interpreting results. Ethics in Science and Environmental Politics, 8, 93–102. doi:10.3354/esep00084

Bozeman, B., & Sarewitz, D. (2011). Public value mapping and science policy evaluation. Minerva, 49(1), 1–23. doi:10.1007/s11024-011-9161-7

Burke, J., Bergman, J., & Asimov, I. (1985). The impact of science on society. Washington, DC: National Aeronautics and Space Administration.

Bush, V. (1945). Science: The endless frontier. [A report to President Truman outlining his proposal for post-war U.S. science and technology policy.] Washington, DC: United States Government Printing Office.

Buxton, M. (2011). The payback of "Payback": Challenges in assessing research impact. Research Evaluation, 20(3), 259–260.

Buxton, M., & Hanney, S. (1994). Assessing payback from Department of Health Research and Development: Preliminary report: Vol. 1. The main report (Research Report, No. 19). Uxbridge, United Kingdom: HERG, Brunel University.

Buxton, M., & Hanney, S. (1996). How can payback from health services research be assessed? Journal of Health Services Research & Policy, 1(1), 35–43.

Centre for Quality and Change Management. (2011). Final report of Delphi study. E3M project—European indicators and ranking methodology for university third mission. Valencia and León, Spain: Universidad Politécnica de Valencia and the Universidad de León.

Committee on Prospering in the Global Economy of the 21st Century. (2007). Rising above the gathering storm: Energizing and employing America for a brighter economic future. National Academy of Sciences, National Academy of Engineering, Institute of Medicine. Washington, DC: National Academies Press.

Danish Council. (2006). A tool for assessing research quality and relevance. Copenhagen, Denmark: Danish Council for Research Policy.

Department of Education, Science and Training. (2005). Research quality framework: Assessing the quality and impact of research in Australia (Issue paper). Canberra: Commonwealth of Australia.

Department of Education, Science and Training. (2006). Research quality framework: Assessing the quality and impact of research in Australia. Research Impact (Report by the RQF Development Advisory Group). Canberra: Commonwealth of Australia.

Donovan, C. (2007). The qualitative future of research evaluation. Science and Public Policy, 34(8), 585–597. doi:10.3152/030234207X256538

Donovan, C. (2008). The Australian Research Quality Framework: A live experiment in capturing the social, economic, environmental, and cultural returns of publicly funded research. New Directions for Evaluation, 2008(118), 47–60. doi:10.1002/ev.260

Donovan, C. (2011). State of the art in assessing research impact: Introduction to a special issue. Research Evaluation, 20(3), 175–179.

Donovan, C., & Hanney, S. (2011). The "Payback Framework" explained. Research Evaluation, 20(3), 181–183.

Evaluating Research in Context (ERiC). (2010). Evaluating the societal relevance of academic research: A guide. Delft, The Netherlands: Delft University of Technology.

Erno-Kjolhede, E., & Hansson, F. (2011). Measuring research performance during a changing relationship between science and society. Research Evaluation, 20(2), 131–143. doi:10.3152/095820211x12941371876544

European Commission. (2010). Assessing Europe's university-based research. Expert group on assessment of university-based research. Brussels, Belgium: Publications Office of the European Union.

Frank, C., & Nason, E. (2009). Health research: Measuring the social, health and economic benefits. Canadian Medical Association Journal, 180(5), 528–534. doi:10.1503/Cmaj.090016

Frodeman, R., & Briggle, A. (2012). The dedisciplining of peer review. Minerva, 50(1), 3–19. doi:10.1007/s11024-012-9192-8

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). The new production of knowledge: The dynamics of science and research in contemporary societies. London, United Kingdom: Sage.

Godin, B., & Doré, C. (2005). Measuring the impacts of science; beyond the economic dimension, INRS Urbanisation, Culture et Société. Paper presented at the HIST Lecture, Helsinki Institute for Science and Technology Studies, Helsinki, Finland. Available at: http://www.csiic.ca/PDF/Godin_Dore_Impacts.pdf

Göransson, B., Maharajh, R., & Schmoch, U. (2009). New activities of universities in transfer and extension: Multiple requirements and manifold solutions. Science and Public Policy, 36(2), 157–164. doi:10.3152/030234209x406863

Grant, J. (1999). Evaluating the outcomes of biomedical research on healthcare. Research Evaluation, 8(1), 33–38.

Grant, J., Brutscher, P.-B., Kirk, S., Butler, L., & Wooding, S. (2009). Capturing research impacts. Cambridge, United Kingdom: RAND Europe.

Gregersen, B., Linde, L.T., & Rasmussen, J.G. (2009). Linking between Danish universities and society. Science and Public Policy, 36(2), 151–156. doi:10.3152/030234209x406818

Hanney, S., Packwood, T., & Buxton, M. (2000). Evaluating the benefits from health research and development centres: A categorization, a model and examples of application. Evaluation, 6(2), 137–160.

Health Economics Research Group, Office of Health Economics, RAND Europe. (2008). Medical research: What's it worth? Estimating the economic benefits from medical research in the UK. London, United Kingdom: Evaluation Forum.

Higher Education Funding Council for England. (2009). Research Excellence Framework. Second consultation on the assessment and funding of research. September 2009/38. Bristol: Higher Education Funding Council for England.

Higher Education Funding Council for England. (2011). Decisions on assessing research impact. Bristol: Higher Education Funding Council for England.

Higher Education Funding Council for England. (2012). Panel criteria and working methods. Bristol: Higher Education Funding Council for England.

Holbrook, J.B. (2010). The use of societal impacts considerations in grant proposal peer review: A comparison of five models. Technology & Innovation, 12(3), 213–224. doi:10.3727/194982410x12895770314078

Holbrook, J. (2012). Re-assessing the science–society relation: The case of the US National Science Foundation's broader impacts merit review criterion (1997–2011). Retrieved from http://www.scienceofsciencepolicy.net/system/files/attachments/Holbrook_BIC_2.0_final.pdf

Holbrook, J.B., & Frodeman, R. (2010). Comparative Assessment of Peer Review (CAPR). EU/US Workshop on Peer Review: Assessing "broader impact" in research grant applications. Brussels, Belgium: European Commission, Directorate-General for Research and Innovation.

Holbrook, J.B., & Frodeman, R. (2011). Peer review and the ex ante assessment of societal impacts. Research Evaluation, 20(3), 239–246.

Kamenetzky, J.R. (in press). Opportunities for impact: Statistical analysis of the National Science Foundation's broader impacts criterion. Science and Public Policy. doi:10.1093/scipol/scs059

Klautzer, L., Hanney, S., Nason, E., Rubin, J., Grant, J., & 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(3), 201–209.

Kuruvilla, S., Mays, N., Pleasant, A., & Walt, G. (2006). Describing the impact of health research: A Research Impact Framework. BMC Health Services Research, 6.

Lähteenmäki-Smith, K., Hyytinen, K., Kutinlahti, P., & Konttinen, J. (2006). Research with an impact evaluation practises in public research organisations, Kemistintie, Finland: VTT Technical Research Centre of Finland.

Lamm, G.M. (2006). Innovation works. A case study of an integrated pan-European technology transfer model. B.I.F. Futura, 21(2), 86–90.

Lewison, G., & Sullivan, R. (2008). The impact of cancer research: How publications influence UK cancer clinical guidelines. British Journal of Cancer, 98(12), 1944–1950.

Leydesdorff, L. (2012). The Triple Helix of university–industry–government relations. Amsterdam, The Netherlands: Amsterdam School of Communications Research.

Link, A.N., & Scott, J.T. (2011). The theory and practice of public-sector R&D economic impact analysis. Gaithersburg, MD: National Institute of Standards and Technology.

Macilwain, C. (2010). What science is really worth. Nature, 465(7299), 682–684.

Mansfield, E. (1991). Academic research and industrial innovation. Research Policy, 20(1), 1–12. doi:10.1016/0048-7333(91)90080-a

Mansfield, E. (1998). Academic research and industrial innovation: An update of empirical findings. Research Policy, 26(7–8), 773–776. doi:10.1016/s0048-7333(97)00043-7

Mardis, M.A., Hoffman, E.S., & McMartin, F.P. (2012). Toward broader impacts: Making sense of NSF's merit review criteria in the context of the National Science Digital Library. Journal of the American Society for Information Science and Technology, 63(9), 1758–1772. doi:10.1002/asi.22693

Martin, B.R. (2007). Assessing the impact of basic research on society and the economy. Paper presented at the Rethinking the impact of basic research on society and the economy (WF-EST International Conference, 11 May 2007), Vienna, Austria.

Martin, B.R. (2011). The Research Excellence Framework and the "impact agenda": Are we creating a Frankenstein monster? Research Evaluation, 20(3), 247–254.

May, R.M. (1998). The scientific investments of nations. Science, 281(5373), 49–51.

Mervis, J. (2011). Beyond the data. Science, 334(6053), 169–171. doi:10.1126/science.334.6053.169

Miranda, L.C.M., & Lima, C.A.S. (2010). On trends and rhythms in scientific and technological knowledge evolution: A quantitative analysis. International Journal of Technology Intelligence and Planning, 6(1), 76–109.

Molas-Gallart, J., Salter, A., Patel, P., Scott, A., & Duran, X. (2002). Measuring third stream activities. Final report to the Russell Group of universities. Brighton, United Kingdom: Science and Technology Policy Research Unit, University of Sussex.

Molas-Gallart, J., & Tang, P. (2011). Tracing "productive interactions" to identify social impacts: An example from the social sciences. Research Evaluation, 20(3), 219–226.
Montada, L., Krampen, G., & Burkard, P. (1999). Personal and social orientations of psychology college teachers on evaluative criteria for own job performances: Results of an expert survey in German graduate psychology college teachers. Psychologische Rundschau, 50(2), 69–89.
Mostert, S., Ellenbroek, S., Meijer, I., van Ark, G., & Klasen, E. (2010). Societal output and use of research performed by health research groups. Health Research Policy and Systems, 8(1), 30.
Narin, F., Hamilton, K.S., & Olivastro, D. (1997). The increasing linkage between US technology and public science. Research Policy, 26(3), 317–330. doi:10.1016/s0048-7333(97)00013-9
Nason, E., Curran, B., Hanney, S., Janta, B., Hastings, G., O'Driscoll, M., & Wooding, S. (2011). Evaluating health research funding in Ireland: Assessing the impacts of the Health Research Board of Ireland's funding activities. Research Evaluation, 20(3), 193–200.
Niederkrotenthaler, T., Dorner, T.E., & Maier, M. (2011). Development of a practical tool to measure the impact of publications on the society based on focus group discussions with scientists. BMC Public Health, 11, 588.
Nightingale, P., & Scott, A. (2007). Peer review and the relevance gap: Ten suggestions for policy-makers. Science and Public Policy, 34(8), 543–553. doi:10.3152/030234207x254396
Pålsson, C.M., Göransson, B., & Brundenius, C. (2009). Vitalizing the Swedish university system: Implementation of the "third mission." Science and Public Policy, 36(2), 145–150.
Petit, J.C. (2004). Why do we need fundamental research? European Review, 12(2), 191–207.
Roberts, M.R. (2009). Realizing societal benefit from academic research: Analysis of the National Science Foundation's broader impacts criterion. Social Epistemology, 23(3–4), 199–219. doi:10.1080/02691720903364035
Rowe, G., & Wright, G. (2001). Expert opinions in forecasting: The role of the Delphi technique. In J.S. Armstrong (Ed.), Principles of forecasting: A handbook for researchers and practitioners (pp. 125–144). Norwell, MA: Kluwer Academic.
Royal Netherlands Academy of Arts and Sciences, Association of Universities in the Netherlands. (2010). Standard Evaluation Protocol 2009–2015. Protocol for research assessment in the Netherlands. Amsterdam, The Netherlands: Author.
Royal Society. (2011). Knowledge, networks and nations: Global scientific collaboration in the 21st century. London, United Kingdom: Author.
Ruegg, R., & Feller, I. (2003). A toolkit for evaluating public R&D investment: Models, methods, and findings from ATP's first decade. Gaithersburg, MD: National Institute of Standards and Technology.
Rymer, L. (2011). Measuring the impact of research—The context for metric development. Turner, Australia: The Group of Eight.
Salter, A.J., & Martin, B.R. (2001). The economic benefits of publicly funded basic research: A critical review. Research Policy, 30(3), 509–532.
Schapper, C.C., Dwyer, T., Tregear, G.W., Aitken, M., & Clay, M.A. (2012). Research performance evaluation: The experience of an independent medical research institute. Australian Health Review, 36, 218–223. doi:10.1071/AH11057
Scott, J.E., Blasinsky, M., Dufour, M., Mandal, R.J., & Philogene, G.S. (2011). An evaluation of the Mind–Body Interactions and Health Program: Assessing the impact of an NIH program using the Payback Framework. Research Evaluation, 20(3), 185–192.
SISOP. (2011). Conceptualizing the social impact of science. Retrieved from http://sisobproject.wordpress.com/2011/07/20/conceptualizing-the-social-impact-of-science-2/
Smith, C.H.L. (1997). What's the use of basic science? Retrieved from http://wwwnew.jinr.ru/section.asp?sd_id=94
Smith, R. (2001). Measuring the social impact of research. British Medical Journal, 323(7312), 528. doi:10.1136/bmj.323.7312.528
Social Impact Assessment Methods for Research and Funding Instruments Through the Study of Productive Interactions (SIAMPI). (2010, December). SIAMPI Workshop, Brussels. Retrieved from http://www.siampi.eu/Content/SIAMPI/Report%20SIAMPI%20workshop.pdf
Social Impact Assessment Methods for Research and Funding Instruments Through the Study of Productive Interactions (SIAMPI). (2011). Final report on social impacts of research. Retrieved from http://www.siampi.edu/
Spaapen, J., & van Drooge, L. (2011). Introducing "productive interactions" in social impact assessment. Research Evaluation, 20(3), 211–218.
Spaapen, J.B., Dijstelbloem, H., & Wamelink, F. (2007). Evaluating research in context: A method for comprehensive assessment. The Hague, The Netherlands: Consultative Committee of Sector Councils for Research and Development.
Stephan, P. (2012). How economics shapes science. Cambridge, MA: Harvard University Press.
Toole, A.A. (2012). The impact of public basic research on industrial innovation: Evidence from the pharmaceutical industry. Research Policy, 41(1), 1–12. doi:10.1016/j.respol.2011.06.004
United Nations Development Programme. (2010). The millennium development goals. New York, NY: Author.
United States Government Accountability Office. (2012). Designing evaluations. Washington, DC: Author.
van der Meulen, B., & Rip, A. (2000). Evaluation of societal quality of public sector research in the Netherlands. Research Evaluation, 9(1), 11–25.
van der Weijden, I., Verbree, M., & van den Besselaar, P. (2012). From bench to bedside: The societal orientation of research leaders: The case of biomedical and health research in the Netherlands. Science and Public Policy, 39(3), 285–303. doi:10.1093/scipol/scr003
van Vught, F., & Ziegele, F. (Eds.). (2011). Design and testing the feasibility of a multidimensional global university ranking. Final report: Consortium for Higher Education and Research Performance Assessment, CHERPA-Network.
Walter, A.I., Helgenberger, S., Wiek, A., & Scholz, R.W. (2007). Measuring societal effects of transdisciplinary research projects: Design and application of an evaluation method. Evaluation and Program Planning, 30(4), 325–338. doi:10.1016/j.evalprogplan.2007.08.002
Ziman, J. (2000). Real science: What it is, and what it means. Cambridge, United Kingdom: Cambridge University Press.