

Research Article

Strategies towards Evaluation beyond Scientific Impact. Pathways not only for Agricultural Research

Birge Wolf ^{1,*}, Anna-Maria Häring ² and Jürgen Heß ¹

¹ University of Kassel, Faculty of Organic Agricultural Sciences, Organic Farming & Cropping Systems, Nordbahnhofstr. 1a, 37214 Witzenhausen, Germany; E-Mail: jh@uni-kassel.de (JH)

² Eberswalde University for Sustainable Development, Department Policy and Markets in the Agro-Food Sector, Schicklerstr. 5, 16225 Eberswalde, Germany; E-Mail: anna.haering@hnee.de

* Corresponding Author: E-Mail: birge.wolf@uni-kassel.de; Tel.: +49 5542981536; Fax: +49 5542981568

Submitted: 21 July 2014 | In revised form: 20 February 2015 | Accepted: 23 February 2015 |

Published: 15 April 2015

Abstract: Various research fields, like organic agricultural research, are dedicated to solving real-world problems and contributing to sustainable development. Therefore, systems research and the application of interdisciplinary and transdisciplinary approaches are increasingly endorsed. However, research performance depends not only on self-conception, but also on framework conditions of the scientific system, which are not always of benefit to such research fields. Recently, science and its framework conditions have been under increasing scrutiny as regards their ability to serve societal benefit. This provides opportunities for (organic) agricultural research to engage in the development of a research system that will serve its needs. This article focuses on possible strategies for facilitating a balanced research evaluation that recognises scientific quality as well as societal relevance and applicability. These strategies are (a) to strengthen the general support for evaluation beyond scientific impact, and (b) to provide accessible data for such evaluations. Synergies of interest are found between open access movements and research communities focusing on global challenges and sustainability. As both are committed to increasing the societal benefit of science, they may support evaluation criteria such as knowledge production and dissemination tailored to societal needs, and the use of open access. Additional synergies exist between all those who scrutinise current research evaluation systems for their ability to serve scientific quality, which is also a precondition for societal benefit. Here, digital communication technologies provide opportunities to increase effectiveness, transparency, fairness and plurality in the dissemination of scientific results, quality assurance and reputation. Furthermore, funders may support transdisciplinary approaches and open access and improve data availability for evaluation beyond scientific impact. If they begin to use current research information systems that include societal impact

data while reducing the requirements for narrative reports, documentation burdens on researchers may be relieved, with the funders themselves acting as data providers for researchers, institutions and tailored dissemination beyond academia.

Keywords: interdisciplinary; research evaluation; societal impact; transdisciplinary

1. Introduction

A crucial aim of agricultural research is to address sustainable development. Global challenges like climate change [1] or the degradation of ecosystem services have fundamental negative impacts on human health and well-being [2]. Agriculture is both driving and being affected by those developments ([2] p. 98), [3]. Such challenges require immediate and adequate action on the part of the whole of society, but also the contribution of relevant knowledge through research ([3] p. 3; [4] p. 322). However, whether research is able to make that contribution depends primarily on the conditions and incentives within the scientific system.

In this article, the focus will be on research evaluation, which can be an important driver for developing science in the direction of scientifically robust, societally relevant and applicable knowledge production. Currently, scientific quality assurance is mainly performed through peer review of papers and project proposals, while scientific impact is evaluated based on publication output in peer-reviewed journals and citation-based performance indicators (detailed in Section 2.3). Citations of a publication are a measure of the acknowledgement by the respective researcher's peers. Citations are counted by and in peer-reviewed journals that are indexed for citation counting. Furthermore, a researcher's publication output and citation rates can be subsumed in an index, e.g. the h-index [5]. Citations are also used as a measure of the recognition of journals, where all citations of a journal within other journals are counted, e.g. the Journal Impact Factor (IF) used by Thompson Reuters [6]. Accordingly, scientific impact is associated with high publication output in high-impact journals and high citation rates in other highly ranked journals. These measures assess, at best, the impact of research on science itself. However, they neither assess societal impact nor serve as proxies for it [7]. As a result, research which similarly targets audiences outside academia may not be adequately appreciated in research evaluation. The term societal impact is used here to sum up all the practical, social, environmental, economic and other 'real-world' impacts research may have for its target groups and society as a whole.

To overcome shortcomings in current research evaluation practices, several alternative evaluation concepts which take societal impacts into account have been developed over the past few years (see Section 3.2). However, such an evaluation of societal impact

faces some inherent challenges, including time and attribution gaps. The term 'time gap' describes the problem that if impact occurs, it is in most cases with some delay after completion of the research. Secondly, the 'attribution gap' means that impacts are not easily attributed to a particular research activity like a project or publication. For example, the adoption of a particular agricultural innovation may be the result of several research activities combined with policy changes and other influences. Accordingly, the state of the art of societal impact assessment focuses on the contribution of research in complex innovation systems, instead of attributing the impacts linearly in terms of cause and effect [8]. Furthermore, proxies are often employed, instead of direct measures of impact. One example is the concept of 'productive interactions', defined as direct, indirect or financial interactions with stakeholders that support the use of research results and make an impact likely [9].

With bibliometric data it is possible to analyse interdisciplinary publications via references from and citations in different fields [10], as well as interactions between basic and applied research. By contrast, the assessment of societal impact (or corresponding proxies) cannot be built on bibliometric analysis, and in most cases there are no other sources with easy-to-use data available either. Thus the effort involved in data assessment for documentary analysis or interviews, for example, inhibits the frequent use of such evaluation approaches.

Starting from these observations, the aim of this paper is to discuss two possible strategies to facilitate research evaluation that is more balanced, both with regard to scientific quality and impact, and to societal relevance and applicability. The first strategy is to strengthen general support for such evaluation beyond scientific impact; the second is to reduce the effort of societal impact evaluations by improving data availability.

Section 2 below introduces the relevant movements and focuses on shared interests as a base for broader support of evaluation beyond scientific impact. Section 3 then provides concrete measures for such support, including possibilities for improving data availability for evaluation beyond scientific impact. In each section the paper shows how agricultural research that is oriented towards sustainability and real-world impact, with a special focus on organic agricultural research, could be involved in these developments in order to create good conditions for its fields of research. We will

conclude with an overview of the actions that may be undertaken jointly by various actors.

2. Multiple Voices Call for Changes in Knowledge Production and Research Evaluation

Various societal groups are demanding changes in knowledge production and research evaluation, for example researchers and funding agencies engaged in sustainability, global challenges and transdisciplinary approaches, the open access movements, and researchers who scrutinise current research evaluation systems for their ability to serve scientific quality.

2.1. Research Engaged in Sustainability, Global Challenges and Transdisciplinary Approaches

2.1.1. Sustainable Development Requires the Support of Interdisciplinary and Transdisciplinary Research Approaches

Several international assessments synthesise scientific and non-scientific knowledge via multiple-stakeholder processes involving science, governments, NGOs, international organisations and the private sector, for example the Millennium Ecosystem Assessment (MA) [2], the International Assessment of Agricultural Knowledge, Science and Technology for Development (IAASTD) [3] and the World Health Summit ([11] pp. 86–87). These assessments, and some scientific groups that give policy advice, such as the WBGU (German Advisory Council on Global Change) [4], point out that there is considerable pressure on society to tackle pressing challenges adequately, which in turn requires knowledge to be produced, accessed and used in ways that assist such adequate action and are conducive to sustainable development.

However, the transfer of existing knowledge and technologies faces several challenges. On the one hand, the balance of power and conflicting interests impede the use of research evidence ([2] p. 92). The reduction in greenhouse gas emissions, for example, is still not sufficient, although the IPCC has been transferring the state of the art regarding climate change to politics for 20 years now. [1]. On the other hand, the need to increase access, clarity and relevance of research evidence for politics has been discussed [12]. Furthermore, concepts for the transfer of knowledge and technology should reflect on possible risks. Instead of merely assuming the superiority of external knowledge and novel technologies, they should be tested beforehand under actual conditions of use ([3] p. 72) or evaluated in sustainability assessments [13].

The challenges in knowledge transfer also lead to a demand for changes in knowledge production in order to increase the applicability and sustainable benefits of knowledge. The reasons for such demands are firstly that technological development is fast and may have deep, in some cases irreversible impacts on our

ecological, economic or social environment ([14] pp. 87–93). Secondly, post-modern societies consist of complex subsystems that function according to their own inherent rules and often fail to deal with impacts that occur in more than one of them at the same time ([14] pp. 61–63, 87–93). Thus, knowledge production also needs to cut across specialised areas and societal subsystems ([15] p. 544; [4] p. 322) and should support transformative processes ([4] p. 322), [11]. Thirdly, true participation of stakeholders in research processes is required to support practical applicability, ownership of solutions and sustainable impact of knowledge ([2] p. 98; [3] pp. 72–73; [4] p. 322). Accordingly, recommendations cover enhanced knowledge exchange among disciplines, between basic and applied research ([4] p. 322) and between science and politics [12], ([16] p. 9) and the involvement of stakeholders, including the integration of traditional and local knowledge ([2] p. 98; [3] pp. 72–73; [4] p. 322). Such transdisciplinary processes may also be supported by involving 'knowledge brokers' as intermediaries to facilitate knowledge exchange [12], ([17] p. 17). Additionally, joint agenda setting, including science, politics, the economy and in particular civil society organisations is recommended for research regarding sustainability ([4] p. 322) and agriculture ([17] p. 17) and is, in some cases, already practised [18–20]. This corresponds to the aim of civil society organisations to strengthen their influence in research policy, for example [21].

The recommendations specified in this section are well subsumed in the terms co-design, co-production, co-delivery and co-interpretation used by the project VisionRD4SD [22]. These recommendations show that concepts for inter- and transdisciplinary research (e.g. [23–26]) and approaches of 'systems of innovation', understanding innovation as a set of complex processes involving multiple actors beyond science (e.g. [27]), are now well accepted in policy advice. Likewise, several research funders have started to support sustainability and transdisciplinarity explicitly in research programming ([14] pp. 202–214), [28,29].

2.1.2. Current Incentive Systems Are Criticised

Apart from the promising developments mentioned above, current incentive systems are considered inappropriate for encouraging researchers to focus their research on sustainable development.

Reputation-building processes based on publications in high-ranking scientific journals and third-party funding are often governed by disciplinary perceptions and fail to acknowledge interdisciplinary and systemic approaches ([4] p. 351). Interdisciplinary research usually has to match the standards of different disciplines in peer review processes, which adversely affects publication success [10], ([15] p. 547) and the evaluation of multidisciplinary institutions [30]. Audits based on bibliometric performance indicators [15] and, explicitly, the use of journal rankings [10] have been shown to

be biased negatively against inter- and multi-disciplinary research.

Some authors discuss consequences such as poorer career prospects, orientation of research away from complex social questions, reduction in cognitive diversity within a given discipline or the entire science system [10], and an increasing relevance gap between knowledge producers and knowledge users [15]. Similarly, Schneidewind et al. highlight the diversity of the sciences in objectives and theories as a base for societal discussion processes ([14] pp. 30–33) and good scientific policy advice ([14] p. 63).

Thus, researchers, institutions and funding agencies that move towards joint knowledge production for sustainable development may often feel contradicted by the current incentives within scientific reputation systems. Accordingly, the indication is that it is necessary to improve current evaluation practices in general and apply evaluation criteria beyond scientific impact.

2.1.3. Opportunities for (Organic) Agricultural Research

Broader support for changes in knowledge production and research evaluation provides multifarious opportunities for agricultural research. As organic and sustainable farming addresses and works within the complexity of ecological systems, and farmers' knowledge and practices are key to building resilient agricultural production systems, the approaches highlighted in Section 2.1.1 have, since their early days, been advocated in agroecology [31] and organic agricultural research ([19] pp. 15–16), [32,33]. Agricultural researchers are often already in contact with actors along the whole value chain of agriculture, and approaches are reflected in diverse concepts for transdisciplinarity e.g. [34–36], and systems of innovation e.g. [37]. Researchers' experiences, and their awareness of the challenges posed by such approaches e.g. ([19] p. 61), [38], promote their adequate advancement via mutual learning with other research communities. Furthermore, the competence of (organic) agricultural research to develop applicable solutions with substantial value in the context of some pressing social and ecological challenges may become more visible.

Research evaluation that goes beyond conventional performance indicators and involves stakeholders is seen as necessary for agricultural research too ([3] pp. 72–73; [17] pp. 81–84; [19] p. 56). Such research evaluation may facilitate the application of transdisciplinary and related research approaches without disadvantages for researchers' reputations. The necessity of such incentive effects is supported by various statements, e.g. "European agricultural research is currently not delivering the full complement of knowledge needed by the agricultural sector and in rural communities" ([19] p. 57). Similarly, the evaluation of an organic agricultural research programme in Sweden re-

sulted in the verdict 'excellent' by scientific peers, while the agricultural advisors indicated too little relevance to pressing problems [39]. The DAFA position paper "Assessment of applied research" considers it necessary to build a consensus about possible indicators, make a commitment to their rigorous application and improve documentation for practice impact [40]. Thus, (organic) agricultural research may use its commonalities with sustainability research in order to jointly advance inter-disciplinary and transdisciplinary research approaches and to advocate their adequate support in funding and appreciation in research evaluation.

2.2. Open Access with Focus on Benefit for Society

Open access movements also aim to increase the benefit of research results for science and society. More than ten years ago, the Berlin declaration called for open access for original research results, raw data, metadata, source materials, digital representations of pictorial and graphical materials and scholarly multimedia [41]. Arguments in favour of open access are for example a) to regard publicly funded knowledge as public property, b) to enhance the transfer, visibility and benefit of knowledge, which is now easily possible via digital technologies and reasonable because of the increased scientific literacy of the public, and c) to support participation in democratic societies [41,42].

Furthermore, the open access movements provide concepts for increased collaboration and interaction in the creation of research results and pluralisation and transparency in the evaluation of publications, and support the full use of technological developments in data processing (see Section 3.1).

However, the inadequate exchange, use, relevance and ownership of scientific knowledge in politics, practice and society indicate that open access alone does not suffice to create benefits of knowledge. Thus co-design, co-production, co-interpretation and co-delivery are necessary on one hand to serve societal benefit, whilst on the other the dissemination of openly accessible research outputs tailored to target groups within and beyond science is also a requirement. Such a comprehensive view of the benefits of research for society increases the credibility of the arguments and supports the view that the corresponding changes in evaluation criteria can be promoted jointly by open access movements and research that is concerned with sustainable development. In our view, (organic) agricultural research is well placed to become a proficient actor in the process of combining the tasks of these two groups. The (organic) agricultural research community is experienced in knowledge transfer and inter- and trans-disciplinary approaches within the diverse agricultural sector and is aware of 'open-access issues', for example interrelations between agriculture and public goods ([3] pp. 24, 30, 73).

2.3. Improve Current Scientific Impact Evaluation Procedures

In general, evaluation procedures that support scientific quality are required for both basic and applied research as foundations for evidence-based decisions. However, as detailed below, current scientific impact evaluation procedures are shown to have potential negative consequences for scientific quality. Knowledge of these consequences and possibilities for improvement is helpful for strengthening scientific quality, increasing awareness of the general effects of evaluation processes, and generating some 'open space' to introduce criteria related to societal impact.

2.3.1. Challenges of Peer Review as a Socially Embedded Process

Several criteria are used by the scientific community to assess scientific quality. The most common are the novelty and originality of the approach, the rigour of the methodology, the reliability, validity and falsifiability of results and the logic of the arguments presented in their interpretation. Peer review processes are broadly perceived as functioning self-control of the scientific community towards scientific quality in publications and third-party funding. Correspondingly, reviewers trust the fairness and legitimacy of their own review decisions [43].

Nevertheless, peer review processes also reflect hierarchy and power within science as a social system. Editors and peers appear as 'gatekeepers', who not only maintain quality but also uphold existing paradigms and decide which of the many high-quality research papers submitted will be allowed to enter the limited space available in the journal concerned [44, 45]. Evaluative processes are found to involve not only expertise, but also interactions and emotions of peers [46] in ([43] p. 210). Instead of erroneously assuming that a "set of objective criteria is applied consistently by various reviewers", it is necessary to focus on what factors promote fair peer review processes ([43] p. 210).

Undesired decision processes such as strategic voting may occur on peer review panels; it has been suggested that fairness is improved if peers rate rather than rank proposals and give advice to funders instead of deciding about funding [43]. Furthermore, in single-blind reviews, knowledge of the author's person, gender and institutional affiliation may influence peer review [43,47–50]. Double-blind and triple-blind reviews, the latter including editor-blindness, partly reduce bias [45], but advantages for native speakers, preferences for the familiar and insufficient reliability of reviewer recommendations do remain ([43] p. 210), [48,50]. For example, the agreement between peers with and without experience in organic agricultural research has been found to be poor with regard to reviewers' assessment of scientific quality in

organic farming research proposals [51]. In some cases peer review fails to identify fraud, statistical flaws, plagiarism or repetitive publication [47,50]. Recently, trials on the submission of fake papers have revealed alarmingly high acceptance rates, in high-ranked subscription journals [52] and open access journals [53]. The latter study includes some publishers who were already on Beall's list of 'predatory publishers', which identifies open access publishers of low quality [54], [55].

Accordingly, further possibilities for improving peer review processes are being discussed. They focus on increasing efficacy and transparency in research dissemination and quality assurance via the full use of technological developments in connection with open access (see Section 3.1).

2.3.2. Self-Reinforcing Dynamics of Bibliometric Indicators

Bibliometric indicators (Table 1) are also results of socially embedded processes because, firstly, publication in a certain journal reflects the decisions of reviewers and editors, and secondly, citation-based performance indicators subsume the decisions of many scientists as to whether to cite or not. In general, the publication of research evidence is influenced by researcher bias (the observer expectancy effect), which results in a higher likelihood of false positive findings and publication bias, meaning that "surprising and novel effects are more likely to be published than studies showing no effect" ([56] p. 3). Accordingly, "the strength of evidence for a particular finding often declines over time". This is also known as the decline effect ([56] p. 3). Moreover, non-significant results often remain unpublished. This phenomenon, known as the file-drawer effect, distorts the perception of evidence and reduces research reliability and efficacy [57].

The fact that peer decisions are often influenced by metrics also has to be taken into account: Merton describes the cumulative processes of citation rates as the Matthew effect, which follows the principle that "success breeds success" and results in higher citations being overestimated and lower citations underestimated [58]. Such dynamics are enforced by increasing scarcity of time resources and an augmented need to filter a large amount of accessible information [59]. Evidence of the Matthew effect, also called accumulative advantage, is frequently detected in science [60] and considered by scientists to be the major bias in proposal evaluation ([48] pp. 38–39).

A further interaction occurs between metrics and strategic behaviour: as person-related indicators of productivity (publication output) and impact (citation-based indicators) influence funding or career options [61], dividing results into the 'least publishable unit' [62], increasing the number of authors, or citing 'hot papers' are strategies for boosting scientists' performance indicators [45].

Furthermore, indices may hide information. The popular h-index combines publication output and citation rates in one number. It reduces the disproportionate valuation of highly cited and non-cited publications, with the result that researchers with quite different productivity and citation patterns may obtain the same h-index. This has been criticised, and the recommendation is to use several (complementary) indicators to measure scientific performance, in particular separate ones for productivity and impact [63].

The relevance and use of journal-related metrics are also subjects of intense debate. A review of several empirical studies about the significance of the Journal Impact Factor (IF) concluded that "the literature contains evidence for associations between journal rank and measures of scientific impact (e.g. citations, importance and unread articles), but also contains at least equally strong, consistent effects of journal rank predicting scientific unreliability (e.g. retractions, effect size, sample size, replicability, fraud/misconduct, and methodology)" ([56] p. 7). For example, a correlation was detected between decline effect and the IF: initial findings with a strong effect are more likely to be published in journals with a high IF, followed by replication studies with a weaker effect, which are more likely to be published in lower-ranked journals [56].

Moreover, the IF and other journal-based metrics are increasingly considered inappropriate for comparing the scientific output of individuals and institutions. This is indicated by the San Francisco Declaration on Research Assessment (DORA), currently signed by nearly 500 notable organisations and 11,000 individuals [64]. DORA substantiates this statement with findings which show that a) citation distributions within journals are highly skewed; b) the properties of the IF are field-specific: it is a composite of multiple, highly diverse article types, including primary research papers and reviews; c) IFs can be manipulated (or 'gamed') by editorial policy; and d) data used to calculate the IF are neither transparent nor openly available to the public [65]. Gaming of the IF is, for example, possible by increasing the proportion of editorials and news-and-views articles, which are cited in other journals although they do not count as citable items in the calculation of the IF [66].

Thus, journal-based metrics are not only found to be unreliable indicators of research quality; the pressure to publish in high-ranked journals may also com-

promise scientific quality. Furthermore the latter "slows down the dissemination of science (...) by iterations of submissions and rejections cascading down the hierarchy of journal rank" ([56] p. 5) which also enormously increases the burden on reviewers, authors and editors [67].

In agricultural research, some scepticism about journal-related metrics is already evident: the Agricultural Economics Associations of Germany and Austria, for example, perform 'survey-based journal ranking', because this was perceived to be more adequate than using the IF [68].

Apart from current criticism, efforts in indicator development should be acknowledged. In article-based metrics, the weighting of co-authoring and highly cited papers, excluding self-citations, leverage of time frames and inclusion of the citation value (rank of the citing journal) aim to assess scientific impact more precisely. Similarly, the further development of journal-based metrics (see Table 1) involves the exclusion of self-citations and inclusion of citation value, the weighting of field-specific citation patterns, the inclusion of network analyses of citations or weighting the propinquity of the citing journals to one another [69]. Nevertheless, the self-reinforcing dynamics of bibliometric indicators and their interactions with the credibility of science are not taken into account in these indicator variations. For example, the weighting of citation value may even increase accumulative advantage.

To sum up, it seems appropriate to improve peer-review processes, to reject certain indicators, and crucially, to apply a broad set of indicators, because scientific performance is a multi-dimensional concept and indicators always contain the risk that scientists will respond directly to them rather than to the value the indicator is supposed to measure ([10] p. 7). Explicitly, DORA recommends that funding agencies and institutions should "consider a broad range of impact measures including qualitative indicators of research impact, such as influence on policy and practice" [65]. As societal benefit requires scientific quality as a base for evidence, but also goes beyond it, needing a high degree of applicability and positive application impacts, these are in fact supplements, not opponents. Therefore, enriching scientific performance with societal impact indicators can result in decisions and incentives in the scientific system that are more reliable and more beneficial to society.

Table 1. Indicators that are frequently used for scientific impact evaluation.

Citation count	In general, the number of citations received by a paper is counted. They can be summed up for all publications of an institution or person, or calculated relative to the average citation rate of the journal or respective field over a certain period (usually three years) [70,71]. Citation data are counted (except examples provided in Section 3.1) for and in journals indexed in the Journal Citation Report by Thompson Reuters or in the SCImago database by Elsevier [69]. Citations are generally assessed in papers, letters, corrections and retractions, editorials, and other items of a journal.
----------------	--

h-index	The h-index combines publication output and impact in one index: $h = N$ publications with at least N citations, (where the time span for calculation can be selected). For the h-index, there are some derivatives that include the number of years of scientific activity, excluding self-citations, and weighting co-authoring and highly cited papers [5].
IF and journal-based metrics built on Thompson Reuters database	<p>The Journal Impact Factor (IF) is calculated by dividing the number of current-year citations to the source items published in that journal during the previous two years by the number of citable items. It can also be calculated for five years and exclude journal self-citations [6]. Example:</p> $IF = \frac{\text{Number of citations} \in 2014 \text{ for articles of journal } A \text{ published} \in 2012 \wedge 2013}{\text{Number of citeable items} \in \text{journal } A \text{ published} \in 2012 \wedge 2013}$ <p>Another metric is Article Influence, in which the citation time frame is five years, journal self-citation is excluded and the citation value (impact factor of the citing journal) is weighted [69].</p>
Eigenfactor	Eigenfactor also uses Thompson Reuters citation data to calculate journal importance with several weightings. It includes network analysis of citations, weighting citation value and field-specific citation patterns [72].
Journal-based metrics built on Elsevier's Scopus database	All indicators are calculated within a citation time frame of three years. The Source Impact Normalized per Paper (SNIP) is calculated in a similar way to the IF. The Scimago Journal Ranks (SRJ and SJR2) limit journal self-citation and weight citation value. SJR2 includes a closeness weight of the citing journals, meaning that citation in a related field is calculated as being of higher value, because citing peers are assumed to have a higher capacity to evaluate it [69].

3. Concrete Strategies to Support Evaluation beyond Scientific Impact

While Section 2 introduced relevant movements and pointed to shared interest as a base for further cooperation, this section will describe concrete measures for facilitating evaluation beyond scientific impact. As seen in the previous section, evaluation beyond scientific impact may introduce criteria for various aspects of knowledge production (Figure 1).

3.1. Open Access and Technical Development Provide some Solutions to Improve Current Evaluation Practices

Although the quality of peer reviews and self-reinforcing dynamics affect open and subscribed publication models, several possibilities for increasing efficacy in dissemination and quality assurance via digital communication technologies are discussed in the context of open access. For peer review processes, increased transparency is the core issue [73]. Open review, meaning that reviews are published with the preprint or the final paper, is possible with different degrees of openness and interactivity [42], though some aspects are discussed controversially. Disclosure

of authors' identities entails the risk of increasing bias as in single-blind reviews [74], while disclosure of reviewers' identities is shown to preserve a high quality of reviews [75], though suspicions do remain that this may inhibit criticism and make it more difficult to find reviewers [47,76]. However, the publishing of reviews, enabling interactions between reviewers and authors and increasing the basis of feedback and valuation via comment, forum and rating functions for readers, is commonly expected to increase transparency, fairness and scientific progress [44,67,73]. Some applied examples are the Journal BMJ [42], Peerevaluation.org [77] or arXiv.org. At arXiv.org the publication of manuscripts accelerates dissemination and reduces the filedrawer effect; in case of revisions and publication in a journal, the updated versions are added [44,78]. Another possibility is to guarantee publication (except in cases of fraud), but not until there has been a double-blind review of the manuscript focusing solely on scientific quality [67]. Reviews and revised versions may be used for suggested new publication concepts with a modified role for editors [67] or even without journals [56], but also for the current system, where they can serve to assist in publication decisions made on the editorial boards of individual journals.

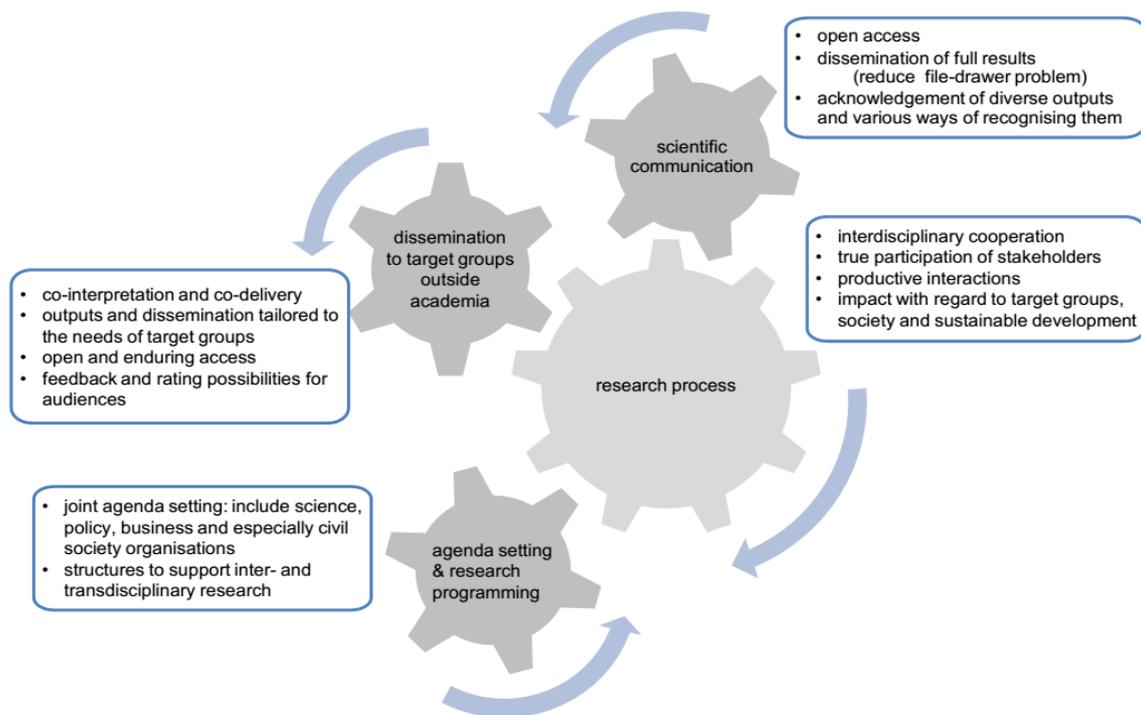


Figure 1. Possible criteria for evaluation beyond scientific impact regarding various aspects of knowledge production.

Additionally, review approaches should allow the engagement of peers in research evaluation to be rewarded [67] and the quality of peer review activities to be assessed [77].

Open access to data is supported by several actors [79]. It enables verification, re-analysis and meta-analysis and reduces publication bias, thus safeguarding scientific quality and societal benefit [80]. Accordingly, it is suggested that the full dissemination of research and re-use of original datasets by external researchers should be implemented as additional performance metrics [80].

Diverse citation and usage data can be accessed via the Internet for all objects with a digital object identifier (DOI) or other standard identifiers [81]. Thus, citation counting beyond Thomson Reuters or Scopus databases is possible, e.g. via Google Scholar, CrossRef, or within Open Access Repositories [42]. Furthermore, responses to papers can be filtered with various Web 2.0 tools (e.g. Altmetrics.com [82]), which are often combined with platforms to share and discuss diverse scholarly outputs (e.g. Impactstory.org). Such data are also tested for the evaluation of the societal use of research [83]. Consequently, the call for open metrics includes open access to citation data in existing citation databases and all upcoming metrics that record citations and utilisation data [42].

In conclusion, there are many opportunities for increasing transparency and interaction in review processes, facilitating and acknowledging cooperative behaviour and including a higher diversity of scientific products and ways of recognising them in research

evaluation processes. This may help to improve current evaluation systems. Until now, these approaches have mostly been restricted to scientific outputs, but they may likewise be used to disseminate outputs and implement feedback functions tailored to diverse user communities outside academia. For example, enhanced data assessment and communication tools are also found to support the concept of citizen science [84], where citizens carry out research or collect data as volunteers [85].

3.2. Science Politics towards Changed Incentive Systems

Science politics, funding procedures and applied evaluation criteria are important drivers of research focuses, and therefore determine what knowledge will exist to face future societal challenges. As seen already in Section 2.1.1, research funders are increasingly interested in supporting transdisciplinarity and related research approaches and they also support open access. For example, the most recent European research programme, "Horizon 2020" [86,87] highlights the need for multi-stakeholder approaches and the support of "systems of innovation" via European Innovation Partnerships [88]. It also makes open access to scientific peer-reviewed publications obligatory and tests open data approaches in certain core areas [89].

Adequate measures to support "Research and Development for Sustainable Development" via research programming are provided by VisionRD4SD, a collaboration process between European research fund-

ers. It identifies measures for the whole programme cycle, presents them in a prototype resource tool and recommends a European or international platform to support networking, dialogue and learning processes on this subject [90]. Likewise, a guide for policy-relevant sustainability research is directed at funding agencies, researchers and policymakers [91].

Institutions and funders who are interested in applying concepts of research evaluation beyond scientific impact (see criteria in Figure 1) can build on existing approaches. Evaluation concepts are developed for interdisciplinary and transdisciplinary research and for societal impact assessment used by research agencies, research institutions or for policy analysis (reviews may be found in [92–94]). Examples of regularly applied evaluation procedures including societal outputs are the Standard Evaluation Protocol for Universities in the Netherlands ([95] p. 5) (see below) and the Research Excellence Framework in the UK [96].

In the section that follows, we will suggest measures to ensure, that evaluation beyond scientific impact is effectively. First, steps should be taken to ensure that societal impact criteria are applied by reviewers, although these indicators may be felt to be outside of reviewers' realm of disciplinary expertise [97] or of lesser importance to them ([48] pp. 32–35). Interestingly, in one study ([48] pp. 32–35), societal impact indicators such as relevance for global societal challenges or citizens' concerns, public outreach, contribution to science education and usefulness for political decision-makers were ranked higher in agricultural research than in other fields, and they were ranked higher by students than by professors. Such results suggest that not only peers, but also knowledge users ([15] p. 548), [97] should be involved in evaluation. To increase the ability of scientists and others to judge societal impacts, data on the societal impact of research and their proxies (hereinafter subsumed as societal impact data) could provide a transparent and reliable basis for such judgement.

Furthermore, the experiences documented in Section 2.3 suggest avoiding narrow indicator sets and their use for competitive benchmarking or metrics-based resource allocation. Instead, broad indicator sets and fair and interactive processes which support organisational development [30] or learning processes [98] need to be applied. One example is the above-mentioned Standard Evaluation Protocol in the Netherlands, where "the research unit's own strategy and targets are guiding principles when designing the assessment process" ([95] p. 5).

However, when funders or institutions begin to apply evaluation beyond scientific impact, they should focus on increasing the acknowledgement of societal impact within the scientific reputation system in general. This is necessary to ensure that their incentives are effective and do not merely increase researchers' trade-offs between contributing to scientific and societal impact. Adequate measures adopted by funders could be

additional funding or distinctions of particularly successful projects as "take-home values" for researchers.

Moreover, research institutions and research funders should become active in improving data availability. Only with reliable and easy-to-use data beyond scientific impact can balanced research evaluation be conducted frequently enough to provide the desired incentives within the scientific system.

Until now, research funding agencies have often demanded detailed reporting on the dissemination and exploitation of results. In German federal research, exploitation plans are required as text documents for proposals and reports [99]. Proposals for Horizon 2020 include plans for dissemination and exploitation ([100] p. 17), but the need to improve digital data assessment for evaluation purposes is also emphasised ([101] p. 47). However, texts with societal impact descriptions cannot be analysed with ease, and the facilities they offer in terms of filtering and cross-referencing are also poor, so they have little value for research evaluation or for the sharing of the information within the scientific system. Likewise, the use of digital systems is only valuable if they allow multiple reuse of data.

4. Improve Data Availability for Evaluation beyond Scientific Impact

To improve the availability of data for societal impact evaluation, we recommend uniting the interests of institutions and funders in such data and giving them more leverage by making use of the current state of interoperability in e-infrastructures, especially research information systems and publication metadata.

Interoperability, in general, enables the exchange, aggregation and use of information for electronic data processing between different systems. Its functionality depends on system structures and exchange formats (entities and attributes), federated identifiers (for persons, institutions, projects, publications and other objects) and shared (or even mapped) vocabularies and semantics [102]. Thus, interoperability includes, besides technical aspects, cooperation to reach agreement.

The interests of institutions and funders in societal impact data may be served by the possibilities of Current Research Information Systems (CRIS). These are used increasingly by research institutions as a tool to manage, provide access to and disseminate research information. Standardisation of CRIS aims to enable automated data input, e.g. via connection to publication databases, and ensure it is only necessary for data to be input manually once but can be used many times (e.g. for automated CVs, bibliographies, project participation lists, institutional web page generation, etc.) [103]. Standardisation is promoted by euroCRIS via the CERIF standard (Common European Research Information Format) [103] and CASRAI (Consortia Advancing Standards in Research Administration Information) via the development of data

profiles and semantics [104], and is embedded in diverse collaborations with initiatives related to interoperability and open access [105].

The CERIF standard is explicitly convenient for enabling interoperability between research institutions and funders, because research outputs can be assigned to projects, persons and organisational units. In the UK, interface management between the research councils and higher education institutions is already established, and societal outputs and impacts are part of the data assessment [106,107]. The aim is to develop these systems further by applying the current CERIF standard in order to increase interoperability with institutional CRIS. It has been shown that output and impact types used in the UK can be implemented in the current CERIF standard [108].

Accordingly, research funders should engage in the development and use of CERIF-CRIS that (a) include data related to interactions with, and benefit for, practice and society, and (b) partly replace written documents in the process of application and reporting. They should (c) act as data providers by making data available, e.g. via interface management with research institutions, file transfer for individual scientists and re-use of data for subsequent proposals and reports. Thus, funders can contribute to the provision of comprehensive societal impact data without increasing the documentation effort for scientists. In doing so, they also help to corroborate and ensure the quality of such data.

To facilitate these aims, several measures can be applied. Regarding (a), it is necessary to develop shared vocabularies for societal impact related to outputs and outcomes. Compiling societal impact data (based on existing evaluation concepts and documentation tools) and structuring them in coherence with CRIS standards (e.g. CERIF, CASRAI) is one task in the project 'Practice Impact II' [109]. Furthermore, funders, researchers and their associations that are interested in societal impact could formulate a mandate to CASRAI and euroCRIS to further develop shared vocabularies for types and attributes of output, outcome and impact towards society and stay involved in this process. Such a commitment would also facilitate the integration of societal impact data in their CRIS by different providers, and this would create a base for data transfer between funders and institutions with regard to (c).

Regarding (b), it is necessary to build a closer connection between those data and the documentation requirements in proposals and reports. The above-mentioned research project, "Practice Impact II", is developing this with a focus on German federal research in the realm of organic and sustainable agriculture. The project integrates the user perspectives of scientists, research funding agencies and evaluators in its development and testing [109,110], in order to achieve the required usability and reduction in effort, with regard to (c), above.

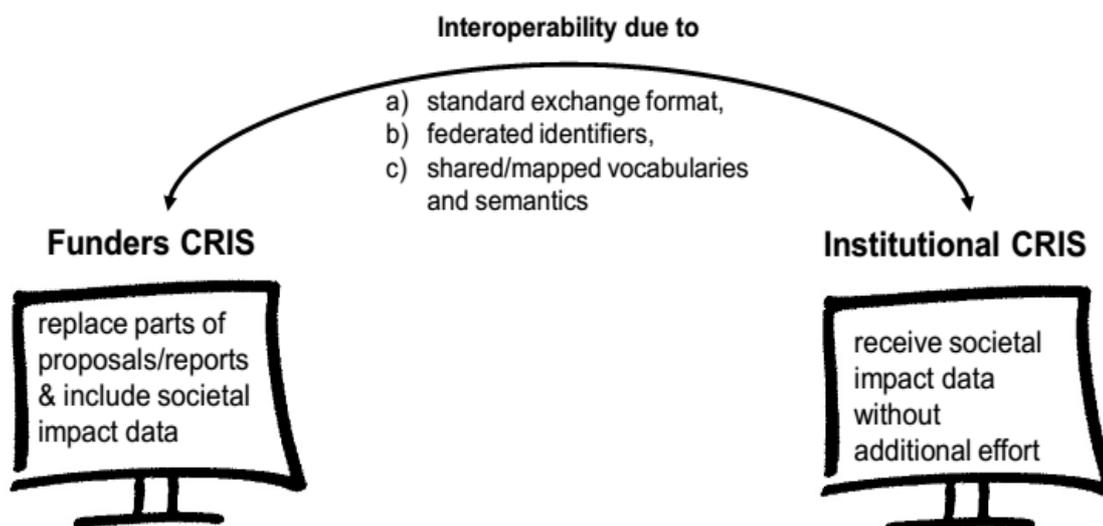


Figure 2. Possibilities for using and developing Current Research Information Systems (CRIS) for interoperable data transfer between funders and institutions to assess and use societal impact data without additional effort.

Regarding (c), there are further possibilities besides the interoperability between funders and institutions. CRIS, with their function as repositories, are also tools for presenting research results to the public. Research funders could use them to support open access dissemination tailored to specific target groups within and beyond academia. Furthermore, closer con-

nections between societal impact data and scientific publications might be established.

For bibliographic metadata of publications, such as authors, title, year, interoperability has already been developed further than it has for other research outputs. Common vocabularies for publication types, advancement of standards and mapping between dif-

ferent standards of metadata are being pushed ahead by libraries [111] and open access repositories [112,113] in order to aggregate machine-readable metadata from multiple systems to create new platforms or services [114]. Furthermore, linked data standards (like the Resource Description Framework, RDF) help to apply the full benefit of web applications for bibliographic metadata. The RDF, for example, allows classical standards-based metadata to be complemented with socially constructed metadata, e.g. user tags, comments, reviews, links, ratings or recommendations [115]. Furthermore, in future, closer links between data and publications will evolve. For example, in 2013, the research data alliance (RDA) started to build social and technical bridges to enable open sharing and interoperability of research data and make them citable, also with an agricultural section [79]. The practice of linking scientific publications with their associated data with the aim of increasing reliability is a recent innovation [80].

Accordingly, the development of systems that link scientific publications via the project to research outputs for audiences outside academia, and to the inter-

actions and impacts of this research as an indication of their societal relevance and applicability is a promising opportunity. Such an increase in the visibility of knowledge tailored towards specific target groups can increase the real-world impact of research and record that impact via feedback functions.

5. Conclusion: Argumentation for Evaluation beyond Scientific Impact

Joint interests of the actors introduced in this paper can be built on the basis that science needs to generate greater societal benefit, and that high scientific quality is a precondition for that. Higher societal benefit is then associated both with open access and with tailored knowledge production and dissemination for audiences beyond academia. Furthermore, evaluation beyond scientific impact can be given some leverage by the full use of digital communication technologies and progress in interoperability. The possible measures suggested in this paper assume close cooperation among various actors (Figure3).

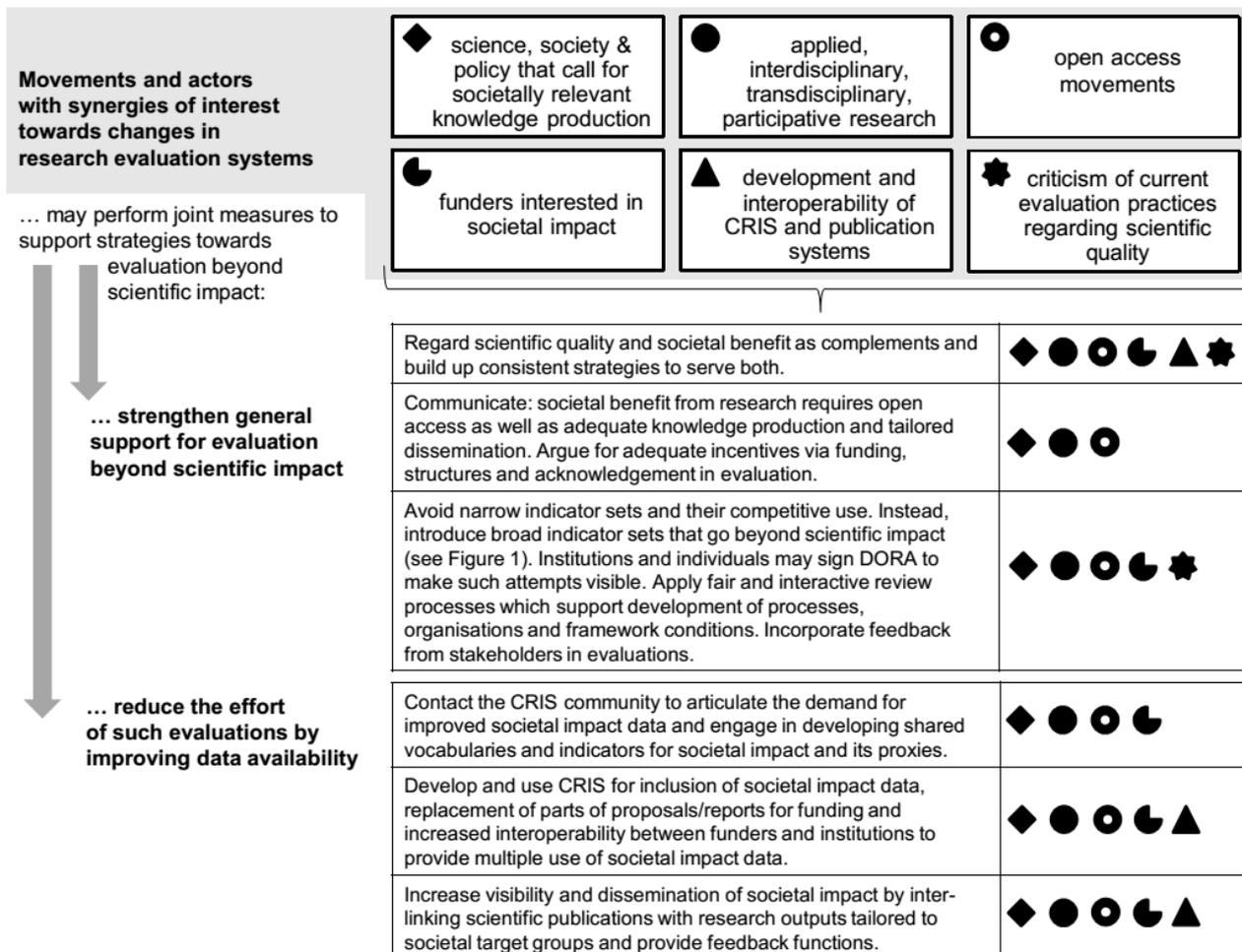


Figure 3. Supporting movements and joint measures to facilitate evaluation beyond scientific impact.

Research funders in particular may support changes in knowledge production because they perform programme design, define funding criteria, and may provide easy-to-use data related to societal impact, for example if research institutions aim to be evaluated with a balance of scientific and societal impact.

As argued in this paper, the measures summarised above are also valid for organic agricultural research and related fields. In the section that follows, some measures and opportunities will be specified.

- Being small, the (organic) agricultural research community may focus on commonalities with other movements. For example, it may benefit from critical voices in scientific impact evaluation, statements of sustainability research and open access movements, which provide the base for introducing criteria beyond scientific impact in research evaluation.
- The (organic) agricultural research community has several synergies with the sustainability (research) community. One is the potential for mutual learning to further develop transdisciplinary research concepts and their proficient application. Another is to organise more powerful support for those research approaches via adequate funding and acknowledgement of societal impact indicators in research evaluation.
- Building up a closer connection between open access and knowledge production tailored to societal needs as two complementary aspects of the societal benefit of science corresponds well with the self-conception of (organic) agricultural research.

If agricultural research funders intend to improve the capabilities for agricultural research to contribute to

References

[1] Intergovernmental Panel on Climate Change (IPCC). Summary for Policymakers. In: Edenhofer O, Pichs-Madruga R, Sokona Y, Farahani E, Kadner S, Seyboth K, Adler A, Baum I, Brunner S, Eickemeier P, Kriemann B, Savolainen J, Schlömer S, Stechow C von, Zwickel T, Minx J, editors. *Climate Change 2014, Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY, USA: University Press.

[2] Reid WVC. *Ecosystems and human well-being. General synthesis: a report of the Millennium Ecosystem Assessment*. Washington, DC, USA: Island Press; 2005.

[3] McIntyre BD. *Synthesis report. A synthesis of the global and sub-global IAASTD reports*. Washington, DC, USA: Island Press; 2009.

[4] Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen (WBGU). *World in transition. A social contract for sustainability*. Berlin, Germany: WBGU; 2011.

real-world impact and sustainable development, they should engage in improving access to societal impact data for supporting evaluation beyond scientific impact within the scientific system. Use-cases for CRIS that integrate societal impact data, reveal funders' needs and reduce scientists' efforts towards proposals and reports may be developed successful in agricultural research. This is because funders and the research community in agricultural research are well connected to jointly develop a use-case with effective feedback loops. Furthermore, they may share their experiences in assessment of societal impact data with other research fields and funders. This may lead to further involvement in processes that support the standardisation and interoperability of those societal impact data.

To conclude, the range of interest groups and viable measures is such that there is no need to accept the deficits in current research evaluation systems, it is possible to change them!

Acknowledgements

Our research has been supported by the German Federal Ministry of Food and Agriculture through the Federal Agency of Agriculture and Food within the Federal Programme of Organic Farming and Other Types of Sustainable Agriculture; project title: Development and Testing of a Concept for the Documentation and Evaluation of the Societal Impact of Agricultural Research. Furthermore we would sincerely like to thank Donal Murphy Bokern, Thorsten Michaelis and Hansjörg Gaus for their very helpful recommendations, and Thomas Lindenthal for his collaboration in our initial work on this topic.

[5] Bornmann L, Daniel H. The state of h index research. Is the h index the ideal way to measure research performance? *EMBO reports*. 2008;0(1):2–6.

[6] Garfield E. *Web of Science. The Thompson Reuters Impact Factor*. Available from: <http://wokinfo.com/essays/impact-factor/>.

[7] Wooding S, Pollitt A, Hanney S, Buxton M, Grant J. *Projekt Retrosight Understanding the returns from cardiovascular and stroke research*. Santa Monica, CA, USA: 2011. Available from: http://www.rand.org/content/dam/rand/pubs/research_briefs/2011/RAND_RB9573.pdf.

[8] Donovan C. State of the art in assessing research impact: introduction to a special issue. *Research Evaluation*. 2011;20(3):175–179.

[9] Spaapen J, van Drooge L. Introducing 'productive interactions' in social impact assessment. *Research Evaluation*. 2011;20(3):211–218.

[10] Rafols I, Leydesdorff L, O'Hare A, Nightingale P, Stirling A. How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management. *Research Policy*. 2012;41(7):1262–1282.

- [11] Ganten D, Bonk M, Gerber T, Heldmann J. World Health summit Yearbook 2013. 2014. Available from: http://www.worldhealthsummit.org/fileadmin/downloads/2013/WHS_2013/Publications/WHS_Yearbook2013_web.pdf.
- [12] Oliver K, Innvar S, Lorenc T, Woodman J, Thomas J. A systematic review of barriers to and facilitators of the use of evidence by policymakers. *BMC Health Services Research*. 2014;14(2):1–12.
- [13] United Nations Environment Programme (UNEP). Application of the Sustainability Assessment of Technologies (SAT) Methodology. 2. February 2012. Available from: http://www.unep.org/ietc/Portals/136/Publications/Waste%20Management/IETC_SAT%20Manual_Full%20Doc_for%20web_Nov.2012.pdf.
- [14] Schneidewind U, Singer-Brodowski M. Transformative Wissenschaft. Klimawandel im deutschen Wissenschafts- und Hochschulsystem. Marburg, Germany: Metropolis; 2013.
- [15] Nightingale P, Scott A. Peer review and the relevance gap: ten suggestions for policy-makers. *Science and Public Policy*. 2007;34(8):543–553.
- [16] United Nations (UN). Report of the United Nations Conference on Sustainable Development. 28. August 2012. Available from: <http://www.uncsd2012.org/content/documents/814UNCS2012%20REPORT%20final%20revs.pdf>.
- [17] European Commission, Standing Committee on Agricultural Research (EC SCAR). Agricultural knowledge and innovation systems towards 2020. An orientation paper on linking innovation and research. Luxembourg, Luxembourg: Publications Office; 2014.
- [18] Future earth. Future earth initial design: Report of the Transition Team. Paris, France: International Council for Science, Future Earth, Interim Secretariat; 2013.
- [19] TPorganics Technology Platform. Implementation Action Plan for organic food and farming research. 19. January 2011. Available from: http://www.tporganics.eu/upload/TPOrganics_ImplementationActionPlan.pdf.
- [20] Deutsche Agrarforschungsallianz (DAFA). Available from: <http://www.dafa.de/>.
- [21] VDW (Vereinigung Deutscher Wissenschaftler) eng. FGS (Federation of German Scientists). Zivilgesellschaftliche Plattform Forschungswende. Available from: <http://www.forschungswende.de/index.php>.
- [22] VisionRD4SD. Vision and Principles for Harnessing Research and Development for Sustainable Development (RD4SD). Available from: http://visionrd4sd.eu/?wpfb_dl=3.
- [23] Hirsch Hadorn G, Hoffmann-Riem H, Biber-Klemm S, Grossenbacher-Mansuy W, Joye D, Pohl C, Wiesmann U, Zemp E. Handbook of transdisciplinary research. Dordrecht, The Netherlands and London, UK: Springer; 2008.
- [24] Frodeman R. The Oxford handbook of interdisciplinarity. Oxford, UK and New York, NY, USA: Oxford University Press; 2010.
- [25] Bergmann M, Klein JT, Faust RC. Methods for transdisciplinary research. A primer for practice. Frankfurt, Germany: Campus-Verlag; 2012.
- [26] Jahn T, Bergmann M, Keil F. Transdisciplinarity: Between mainstreaming and marginalization. *Ecological Economics*. 2012;79:1–10.
- [27] Wicczorek AJ, Hekkert MP. Systemic instruments for systemic innovation problems: A framework for policy makers and innovation scholars. *Science and Public Policy*. 2012;39(1):74–87.
- [28] Bundesanstalt für Landwirtschaft und Ernährung (BLE). Bundesprogramm Ökologischer Landbau und andere Formen nachhaltiger Landwirtschaft (BÖLN). Available from: <http://www.bundesprogramm-oekolandbau.de/>.
- [29] Griebhammer R, Jahn T, Korbun T, Kraemer AR, Leggewie C, Renn O, Scheidewind U, Zahrnt A. Understanding—Evaluating—Shaping. Transdisciplinary Knowledge for a Sustainable Society. Memorandum on the Development of Social-Ecological Research in Germany. 2012. Available from: http://www.fona.de/mediathek/pdf/soef-Memorandum_2012_en.pdf.
- [30] Simon D, Knie A. Can evaluation contribute to the organizational development of academic institutions? An international comparison. *Evaluation*. 2013;19(4):402–418.
- [31] Méndez EV, Bacon CM, Cohen R. Agroecology as a Transdisciplinary, Participatory, and Action-Oriented Approach. *Agroecology and Sustainable Food Systems*. 2013;1(37):3–18.
- [32] Lindenthal T, Vogl CR, Heß J. Forschung im Ökologischen Landbau. Integrale Schwerpunktthemen und Methodikkriterien. Sonderausgabe der Zeitschrift Förderungsdienst. 1996;(2c).
- [33] Aeberhard A, Rist S. Transdisciplinary co-production of knowledge in the development of organic agriculture in Switzerland. *Ecological Economics*. 2009;68(4):1171–1181.
- [34] Hoffmann V, Probst K, Christinck A. Farmers and researchers: How can collaborative advantages be created in participatory research and technology development? *Agriculture and Human Values*. 2007;3(24):355–368.
- [35] Barbercheck M, Kiernan NE, Hulting AG, Duiker S, Hyde J, Karsten H, Sanchez E. Meeting the 'multi-' requirements in organic agriculture research: Successes, challenges and recommendations for multifunctional, multidisciplinary, participatory projects. *Renewable Agriculture and Food Systems*. 2012;27(2):93–106.
- [36] Zscheischler J, Rogga S. Transdisciplinarity in land use science—A review of concepts, empirical findings and current practices. *Futures*. 2015;65(0):28–44.
- [37] Le Gal P, Dugué P, Faure G, Novak S. How does research address the design of innovative agricultural production systems at the farm level? A review. *Agricultural Systems*. 2011;104(9):714–728.
- [38] Lang DJ, Wiek A, Bergmann M, Stauffacher

- M, Martens P, Moll P, Swilling M, Thomas CJ. Transdisciplinary research in sustainability science: practice, principles, and challenges. *Sustainability Science*. 2012;7(S1):25–43.
- [39] Formas. Evaluation Report 2006—Evaluation of research on organic production in Sweden. Ljunglöfs, Sweden: EDITA; 2007.
- [40] Deutsche Agrarforschungsallianz (DAFA). Forschungsstrategisches Fachforum 2012. Agrarforschung—zwischen bewerten und verwerten. Impulspapier Bewertung angewandter Forschung. 29. April 2013. Available from: http://www.dafa.de/fileadmin/dam_uploads/images/Fachforen/FSFF/FSFF_2012/DAFA-2013-04-29-Impulspapier%20Bewertung%20angewandter%20Forschung.pdf.
- [41] BDOA. Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. Available from: <http://openaccess.mpg.de/Berlin-Declaration>.
- [42] Herb U. Offenheit und wissenschaftliche Werke: Open Access, Open Review, Open Metrics, Open Science & Open Knowledge. In: Herb U, editor. Open initiatives. Offenheit in der digitalen Welt und Wissenschaft. Saarbrücken, Germany: Universaar; 2012. pp. 11–44.
- [43] Lamont M, Huutoniemi K. Comparing customary rules of fairness. *Evaluative Practices in Various Types of Peer Review Panels*. In: Camic C, Gross N, Lamont M, editors. *Social Knowledge in the Making*. Chicago, IL, USA: The University of Chicago Press; 2011. pp. 209–232.
- [44] Nosek BA, Bar-Anan Y. Scientific Utopia: I. Opening Scientific Communication. *Psychological Inquiry*. 2012;23(3):217–243.
- [45] Fröhlich G. Anonyme Kritik: Peer Review auf dem Prüfstand der Wissenschaftsforschung. *Medizin-Bibliothek-Information*. 2003;3(2):33–39.
- [46] Lamont M. *How Professors Think: Inside the Curious World of Academic Judgment*. Cambridge, MA, USA: Harvard University Press; 2009.
- [47] Twaij H, Oussedik S, Hoffmeyer P. Peer review. *The bone & joint journal*. 2014;96-B(4):436–441.
- [48] Otsus O, Must Ü, Denton M, Mustajoki A, Tiirik T. Peer Review Practices. 27. December 2012. Available from: http://research-acumen.eu/wp-content/uploads/Del-1.3_27Dec2012.pdf.
- [49] Eyre-Walker A, Stoletzki N, Eisen JA. The Assessment of Science: The Relative Merits of Post-Publication Review, the Impact Factor, and the Number of Citations. The assessment of science: the relative merits of post-publication review, the impact factor, and the number of citations. *PLoS Biology*. 2013;11(10):1–8/e1001675.
- [50] Ludbrook J. Peer review of manuscripts. *Journal of clinical neuroscience: official journal of the Neurosurgical Society of Australasia*. 2002;9(2):105–108.
- [51] Rasmussen J, Langer V, Alrøe HF. Bias in Peer Review of Organic Farming Grant Applications. *Agriculture and Human Values*. 2006;23(2):181–188.
- [52] van Noorden R. Publishers withdraw more than 120 gibberish papers. *Nature News*; 2014.
- [53] Bohannon J. Who's Afraid of Peer Review? *Science*. 2013;342(6154):60–65.
- [54] Herb U. Unzutreffend, aber schmerzhaft: Der Open-Access-Sting der Zeitschrift *Science*. *Telepolis* (Print). 14. October 2013. Available from: <http://www.heise.de/tp/artikel/40/40056/1.html>.
- [55] Beall. Scholarly Open Access. Bealls list of predatory publishers. Available from: <http://scholarlyoa.com/publishers/>.
- [56] Brembs B, Button K, Munafò M. Deep impact: unintended consequences of journal rank. *Frontiers in Human Neuroscience*. 2013;7:1–12.
- [57] Pautasso M. Worsening file-drawer problem in the abstracts of natural, medical and social science databases. *Scientometrics*. 2010;85(1):193–202.
- [58] Merton RK. The Matthew Effect in Science. *Science*. 1968;159(3810):56–63.
- [59] Xie Y. "Undemocracy": inequalities in science. *Science*. 2014;344(6186):809–810.
- [60] Perc M. The Matthew effect in empirical data. *Journal of the Royal Society, Interface. The Royal Society*. 2014;11(98):1–15.
- [61] van Dijk D, Manor O, Carey LB. Publication metrics and success on the academic job market. *Current Biology*. 2014;24(11):R516.
- [62] Dupps W Jr., Randleman JB. The perils of the least publishable unit. *Journal of refractive surgery*. 2012;28(9):601–602.
- [63] Bornmann L, Daniel H. The state of h index research. Is the h index the ideal way to measure research performance? *EMBO reports*. 2008;10(1):2–6.
- [64] American Society for Cell Biology. Homepage DORA. Available from: <http://www.ascb.org/dora/>.
- [65] American Society for Cell Biology. San Francisco Declaration on Research Assessment. 07. May 2013. Available from: <http://www.ascb.org/dora-old/files/SFDeclarationFINAL.pdf>.
- [66] The PLoS Medicine Editors. The impact factor game. It is time to find a better way to assess the scientific literature. *PLoS medicine*. 2006;3(6):e291.
- [67] Kravitz DJ, Baker CI. Toward a New Model of Scientific Publishing: Discussion and a Proposal. *Frontiers in Computational Neuroscience*. 2011;5:1–12.
- [68] Herrmann R, Berg E, Dabbert S, Pöchtrager S, Salhofer K. Going Beyond Impact Factors: A Survey-based Journal Ranking by Agricultural Economists. *Journal of Agricultural Economics*. 2011;62(3):710–732.
- [69] Guerrero-Bote VP, Moya-Anegón F. A further step forward in measuring journals' scientific prestige: The SJR2 indicator. *Journal of Informetrics*. 2012;6(4):674–688.
- [70] Colledge L. *Snowball Metrics Recipe Book*. 2014. Available from: <http://www.snowballmetrics.com/wp-content/uploads/snowball-metrics-recipe-book-upd.pdf>.
- [71] Thompson Reuters. Summary Metrics. Available from: http://incites.isiknowledge.com/common/help/h_summary_metrics.html#jxc_ratio.

- [72] Bergstrom CT, West JD, Wiseman MA. The Eigenfactor metrics. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2008;28(45):11433–11434.
- [73] Kriegeskorte N, Walther A, Deca D. An emerging consensus for open evaluation: 18 visions for the future of scientific publishing. *Frontiers in computational neuroscience*. 2012;6:1–5.
- [74] Lee M, Om K, Koh J. The Bias of Sighted Reviewers in Research Proposal Evaluation: A Comparative Analysis of Blind and Open Review in Korea. *Scientometrics*. 2000;47(1):99–116.
- [75] van Rooyen S, Delamothe T, Evans SJW. Is open peer review the fairest system? Yes. *BMJ*. 2010;341:c5729.
- [76] Khan K. Is open peer review the fairest system? No. *BMJ*. 2010;341:c6425.
- [77] Birukou A, Wakeling JR, Bartolini C, Casati F, Marchese M, Mirylenka K, Osman N, Ragone A, Sierra C, Wassef A. Alternatives to Peer Review: Novel Approaches for Research Evaluation. *Frontiers in Computational Neuroscience*. 2011;5(issue):1–12.
- [78] Boldt A. Extending ArXiv.org to Achieve Open Peer Review and Publishing. *Journal of Scholarly Publishing*. 2011;42(2):238–242.
- [79] Research Data Alliance (RDA). Homepage RDA. Available from: <https://rd-alliance.org/>.
- [80] Chan A, Song F, Vickers A, Jefferson T, Dickersin K, Gøtzsche PC, Krumholz HM, Ghersi D, van der Worp HB. Increasing value and reducing waste: addressing inaccessible research. *The Lancet*. 2014;383(9913):257–266.
- [81] Piowar H. Value all research products. *Nature*. 2013;493:159–159.
- [82] Altmetrics Manifesto. Available from: <http://altmetrics.org/manifesto/>.
- [83] Meijer I, Calero-Medina C, Abma-Schouten R. Societal Use of Research (SURE): stakeholder perceptions and data method development. In: Hinze S, Lottmann A, editors. *Translational twists and turns: Science as a socio-economic endeavor*. Proceedings of STI 2013. Berlin, Germany. pp. 237–241.
- [84] European Commission (DG Research, DG Communications Networks Content and Technology). Background Document Public Consultation 'Science 2.0': Science in Transition. 02. July 2014. Available from: <http://ec.europa.eu/research/consultations/science-2.0/background.pdf>.
- [85] Bird TJ, Bates AE, Lefcheck JS, Hill NA, Thomson RJ, Edgar GJ, Stuart-Smith RD, Wotherspoon S, Krkosek M, Stuart-Smith JF, Pecl GT, Barrett N, Frusher S. Statistical solutions for error and bias in global citizen science datasets. *Biological Conservation*. 2014;173:144–154.
- [86] European Union (EU). Regulation (EU) No 1291/2013 of the European Parliament and of the Council of 11 December 2013. Establishing Horizon 2020—the Framework Programme for Research and Innovation (2014–2020) and repealing Decision No 1982/2006/ECText with EEA relevance. 20. December 2013. http://inea.eu.europa.ec/download/legal_framework/regulation_12912013_establishing_h2020.pdf.
- [87] European Commission (EC). Homepage Horizon 2020. Available from: <http://ec.europa.eu/programmes/horizon2020/en>.
- [88] European Commission (EC). Homepage Innovation Union EIP. Available from: http://ec.europa.eu/research/innovation-union/index_en.cfm?pg=eip.
- [89] European Commission (EC). Fact sheet: Open access in Horizon 2020. Available from: https://ec.europa.eu/programmes/horizon2020/sites/horizon2020/files/FactSheet_Open_Access.pdf.
- [90] VisionRD4SD. Project Results. Available from: <http://visionrd4sd.eu/project-results/>.
- [91] Jahn T, Keil F. Policy Relevant Sustainability Research—Requirements Profiles for Research Funding Agencies, Researchers and Policymakers Regarding Improving and Ensuring Quality of Research—A Guide. 16. September 2013. Available from: http://www.umweltbundesamt.de/sites/default/files/medien/378/publikationen/guide_naf0_20130826_komplett.pdf.
- [92] Wolf B, Lindenthal T, Szerencsits M, Holbrook JB, Heß J. Evaluating Research beyond Scientific Impact—How to Include Criteria for Productive Interactions and Impact on Practice and Society. *GAIA—Ecological Perspectives for Science and Society*. 2013;22(2):104–114.
- [93] Butz B, Hirsch Hadorn G, Joye D, Lawrence R, Nentwich M, Paulsen T, Perrig-Chiello P, Pohl C, Rossini M, Truffer B, Wastl-Walter D, Wiesmann U, Zinsstag J. Questions to evaluate inter- and transdisciplinary research proposals. Berne, Switzerland: td-net for Transdisciplinary Research; 2010.
- [94] Grant J, Brutscher P, Kirk SE, Butler L, Wooding S. Capturing Research Impacts: A review of international practice. 25. January 2010. Available from: http://www.rand.org/content/dam/rand/pubs/documented_briefings/2010/RAND_DB578.pdf.
- [95] Association of Universities in the Netherlands (VSNU), Netherlands Organisation for Scientific Research (NWO), Royal Netherlands Academy of Arts and Sciences (KNAW). Standard Evaluation Protocol 2015–2021. 20. March 2014. Available from: <https://www.knaw.nl/shared/resources/actueel/publicaties/pdf/standard-evaluation-protocol-2015-2013-2021>.
- [96] Education Funding Council for England (HEFCE). Submission system user guide. 2013. Available from: <http://www.ref.ac.uk/media/ref/content/pub/userguide/REF2014%20live%20submission%20system%20user%20guide%20-%20April%2013.pdf>.
- [97] Holbrook JB, Frodeman R. Peer review and the ex ante assessment of societal impacts. *Research Evaluation*. 2011;20(3):239–246.
- [98] Rogers JD, Jordan G. New research evaluation frameworks and methods for systems level learning: introduction to a special section. *Research Evaluation*. 2010;19(4):235–237.

- [99] Bundesministerium für Bildung und Forschung (BMBF). Muster Schlussbericht BMBF-Anlage 2 zu Nr. 3.2 BNBest-BMBF 98. 20. November 2008. Available from: https://foerderportal.bund.de/easy/module/easy_formulare/download.php?datei=1021.
- [100] European Commission (EC). Horizon 2020 Research and Innovation Actions (RIA) Innovation Actions (IA) Administrative forms (Part A), Research proposal (Part B). Available from: http://ec.europa.eu/research/participants/data/ref/h2020/call_ptef/pt/h2020-call-pt-ria-ia_en.pdf.
- [101] European Commission (EC). Commission Staff Working Paper: Impact Assessment. Brussels, Belgium: 30. November 2011. Available from: http://ec.europa.eu/justice/data-protection/document/review2012/sec_2012_72_en.pdf.
- [102] Baker D, Simons E, Brown J. The various aspects of Interoperability: A strategic partnership driving interoperability in research information through standards. Available from: <http://hdl.handle.net/11366/354>.
- [103] euroCRIS. The European Organisation for International Research Information. Available from: <http://www.eurocris.org>.
- [104] CASRAI. The Casrai Dictionary. Available from: <http://casrai.org>; <http://dictionary.casrai.org/>.
- [105] Laliu H. euroCRIS Newsflash. Issue 65; April 2014. Available from: <http://www.eurocris.org/Uploads/Web%20pages/newsflash/Newsflash%2065.pdf>.
- [106] Research Councils UK (RCUK). Research Outcomes Project Invitation to Tender Research Outcome Types. 21. April 2010. Available from: <http://researchoutcomes.files.wordpress.com/2010/07/summary-of-output-types-issue-1-01.pdf>.
- [107] Research Councils UK (RCUK). Research Outcomes System (ROS). Available from: <http://www.rcuk.ac.uk/research/researchoutcomes/>.
- [108] Jörg B, Waddington S, Jones R, Trowell S. Harmonising Research Reporting in the UK—Experiences and Outputs from UKRISS. *Procedia Computer Science*. 2014;33:207–214.
- [109] Wolf B, Szerencsits M, Gaus H, Müller CE, Heß J. Developing a Documentation System for Evaluating the Societal Impact of Science. *Procedia Computer Science*. 2014;33:289–296.
- [110] Blanc J, Aföldi T, Bellon S, Niggli U. How to promote innovation and interdisciplinarity in organic food and farming research evaluation. Available from: <http://orgprints.org/view/projects/conference.html>.
- [111] Deutsche Initiative Netzwerkinformation e.V. (DINI). Gemeinsames Vokabular für Publikations- und Dokumenttypen. 02. July 2010. Available from: <http://edoc.hu-berlin.de/series/dini-schriften/12/PDF/12.pdf>.
- [112] Schirrwagen J. Briefing Paper about the COAR Interest Group "Controlled Vocabularies for Repository Assets". Available from: https://www.coar-repositories.org/files/Briefing_paper_coar_ig_controlled_vocabularies_13012014.pdf.
- [113] Confederation of Open Access Repositories (COAR). The Current State of Open Access Repository Interoperability (2012). Available from: <https://www.coar-repositories.org/files/COAR-Current-State-of-Open-Access-Repository-Interoperability-26-10-2012.pdf>.
- [114] Lagoze C, Van de Sompel H, Warner S. Homepage Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). Available from: <http://www.openarchives.org/pmh/>.
- [115] Alemu G, Stevens B, Ross P, Chandler J. Linked Data for Libraries: Benefits of a Conceptual Shift from Library-Specific Record Structures to RDF-based Data Models. World Library and information Congress. Available from: <http://conference.ifla.org/past-wlic/2012/92-alemu-en.pdf>.