Criteria and indicators for the societal effects of IASS activities

Version: Final Version, October 2018

Author: Janina Schirmer

Final Version, approved by Board of Directors, External Evaluation Group and Advisory Board


IASS
POTSDAM

Table of Contents

The IASS is due to be evaluated once again by the German Council of Science and Humanities (hereafter, the Science Council) in 2020. The results of this evaluation will determine the future of the Institute and have a decisive influence on the future focus, institutional structure and financial resources available to the IASS. A coherent set of criteria and indicators must therefore be developed in preparation for this evaluation. These criteria and indicators must be similar enough to those applied by the Science Council in order to meet its requirements and, at the same time, specific enough to reflect the Institute's defining characteristics. In a first step toward this, a quality model was developed in 2017, utilizing a set of criteria and indicators derived from the Institute's profile. This was submitted to the Advisory Board for discussion in October of that year and any deficits or significant issues in this model were pointed out to the Advisory Board at this time.

The Advisory Board welcomed the development of this model, emphasized the importance of impact indicators, and suggested that steps be taken to further develop the criteria for societal effects and their integration within a comprehensive, science-based framework.[1]

This further development has been implemented here as follows: the first part of this document offers an overview of the criteria and indicators applied to measure societal effects. In a first step, the criteria and indicators used in existing quality assessment procedures are presented and explained. This overview serves to shape the expectations of peers in the coming evaluation and, at the same time, to provide examples of best practice. Following this, an assessment of the current academic discussion of impact criteria is offered. The main focus here is on identifying suitable indicators and criteria as well as methods of documentation, collection and analysis.

Building on this and the findings noted in the memorandum of October 2017, a set of quality criteria and indicators for societal effects is proposed that reflects the specific transdisciplinary research approach applied at the IASS. In doing so, the Institute's profile provides a starting point for the elaboration of a comprehensive evaluation framework, within which selected criteria, indicators and methods have been integrated in response to the relevant requirements.

In the fourth section, the methods of collection, documentation and analysis necessary for the evaluation framework are presented and the necessary resources and changes to the existing system are discussed. A brief discussion of the outstanding issues and future steps concludes this document.

[1] The Advisory Board also welcomed the decision to develop internal process-oriented quality assurance mechanisms along with the effect assessment. The criteria for these mechanisms will be deduced from the IASS methodology, which is still subject to discussion. Even though this document has not been finalized, the links to process quality evaluation are pointed out in this paper where necessary.

1 An overview of current practice: Criteria and indicators for societal effects applied in relevant evaluation processes

1.1 Criteria and indicators applied in the evaluations of the German Council of Science and Humanities (WR), the Leibniz Association (WGL), and the Helmholtz Association of German Research Centres (HGF)

In the case of evaluations conducted by the Science Council, an expert working group is convened to identify suitable criteria and determine their weighting. The criteria selected generally reflect the primary focus of the respective institution – for example, research, the supply of scientific services and infrastructure, or science-based advice.

The guidelines set out in **"Tasks, criteria and procedures of the Evaluation Committee of the German Council of Science and Humanities"** (Wissenschaftsrat 2014) provide a range of criteria that can applied by the working group in its evaluation of the respective institution, including criteria for the evaluation of research performance,[2] organization and (material and financial) endowment,[3] and infrastructure, services and consultancy. The range of criteria applied within this scope addresses the **accessibility, relevance and user orientation** of research infrastructures and/or services as well as the **independence, research basis, quality assurance, transparency, strategic thinking and action** and **user orientation** of consulting services.

Data relevant to these criteria are collected using a standardized questionnaire covering a wide range of indicators. The societal effects of research, on the other hand, are only dealt with to a limited extent in this questionnaire. The areas covered within this questionnaire include "classical" technology transfer (quantitative data on **patents and licensing revenues**), user orientation (user groups, specifically)[4], and policy consultancy outputs (policy papers, internal position papers per year). **Qualitative indicators** (e.g. the existence and quality of a transfer strategy and/or strategies for communicating with users and the public) are also assessed using the questionnaire. **Examples of successful transfer efforts** are also collected. However, the questionnaire does not set out a pre-defined format for these case studies or specify how they are to be assessed.[5]

In 2007, the Science Council developed criteria and indicators especially for **departmental research institutions of the Federal Government and the Länder** to measure the delivery of "Science-based services and transfer" (Wissenschaftsrat 2007): Here too, the accessibility, relevance and user orientation of infrastructures and/or services as well as independence, a sound research basis, adequate quality assurance processes, sufficient transparency, strategic focus and the user orientation of consultancy services are identified as criteria. The "Professionalism of public-oriented services"

[2] Functional and flexible internal governance (bodies and management structures), adequacy of personnel numbers and organization, quality assurance with regard to personnel, gender equality, flexibility in resource management, and adequacy of funding, facilities and infrastructure.

[3] Coherence of the research programme, quality and innovative nature of research, adequate quality assurance, (national and international) cooperation, commitment to the promotion of young researchers, appropriate practical orientation

[4] "Are user satisfaction surveys carried out at regular or irregular intervals? If so, please summarize the results of the last user survey."

[5] "Please provide examples from the last three years (2011-2013) of the successful transfer into practice of research or development services and other activities (e.g. spin-offs, etc.)."

(ibid., p. 13) is also noted as a relevant measure. In addition to the set of indicators mentioned above, the following are also noted: the **inclusion of the institution** in political inquiries, legislative, regulatory and harmonization projects; the existence and quality of **rules and procedures for ensuring "good policy advice"** and as indicators of the professionalism of public/policy-oriented services; the **topicality of topics**; the coverage of all important **target groups** and the **scope** and **appropriateness of the media used** (ibid., p. 11ff.). Based on an assessment of several evaluations, the recommendations for the further development of the departmental research institutions of the Federal Ministry of Food and Agriculture (BMEL) provides guidance on the use of these indicators in defining and evaluating the performance area "Science-based services and transfer". To this end, data is collected (and classified by policy level) on the number of scientific opinions provided within the context of legislative, approval, evaluation and harmonization processes and/or official decisions; similarly, the number of inquiries and requests is classified by client/inquirer and communication channel. A comparative assessment of regulations and quality assurance processes for policy consulting was also conducted in this context. A list of different formats and pathways reflects the diversity of transfer formats (Wissenschaftsrat 2017: p. 41ff.).

Further criteria for the quality of societal effects can also be derived from the **Statement of the German Council of Science and Humanities on Science and Technology Transfer** (Wissenschaftsrat 2016) The Science Council attaches particular importance to transfer as one of the four dimensions of performance within the science system. In its statement of 2016, the Science Council defines transfer in a broad and recursive sense as "interactions of scientific actors with partners from outside science, society, culture, business and policymaking" (Wissenschaftsrat 2016, *op. cit.*, p. 4) and discusses the particular challenges associated with its successful implementation on the basis of three exemplary areas: communication, consultancy and application. As part of a transfer strategy to be developed by each institution, the Science Council recommends the establishment of **institution-specific criteria** and the development of a corresponding **institution-specific evaluation concept** for transfer. The Science Council further recommends that **input, transfer activity** and **output** be recorded in a manner that takes into account the purpose and context of the evaluation. In certain contexts, the Science Council notes, documenting the **outcomes**[6] and **impacts**[7] of transfer activities can also provide valuable insights. Such criteria have been subject to criticism, however, and this should be taken into account.

This **criticism** is concerned with the **problematic attribution** of impacts in light of a lack of causal links, challenges in pinpointing specific sources, accounting for delayed effects, and tracking the **unintended consequences** of scientific activities (ibid., p. 46). According to the Science Council, these challenges make it inappropriate to evaluate the dimension of performance solely on the basis of a few (quantitative) indicators/data – here, even more so than in the evaluation of research itself. Attempts to do justice to the heterogeneity of transfer services by expanding the number of indicators applied in evaluations are unlikely to prove successful, particularly in the case of organizations with a highly diversified transfer portfolio. Instead, the Council notes, transfer services should be assessed by experts on the basis of qualitative information **(informed expert review**) in addition to the quantitative input and output variables directly related to them. The evaluation standards applied should also reflect the expectations and perspectives of the interacting partners (ibid., p. 46).

[6] Outcomes are defined here as "the results (...) generated by transfer partners of a scientific institution through follow-up actions to the transfer activity. Examples include: new products or processes in a company, the establishment of new businesses, changes to legal norms by state actors or new media products" (ibid., p. 19).

[7] "Changes in society (e.g. a change in carbon emission levels) induced by transfer activities and innovations based on them" (ibid.).

Conversely, the **Core Data Set on Research Activities**, the implementation of which has been recommended by the Science Council, identifies and defines data that can be used as indicators of (societal) impact. However, as with the classical evaluation process implemented by the Science Council, its scope does not extend beyond traditional technology transfer and **only** includes **patents and spin-offs** (Wissenschaftsrat 2016) While plans exist to broaden this dimension, this is not likely to occur prior to the forthcoming evaluation of the IASS.

The Science Council has in the past made specific proposals with respect to the treatment and evaluation of societal effects. In its **evaluation of the Wuppertal Institute** (Wissenschaftsrat 2012), the Science Council used the standard criteria (see above), but emphasized additional criteria relevant to transfer (research basis, relevance to target groups, successful application of results, cf. ibid., 13, 43), aligning the evaluation criteria with the institute's mission. The scientific basis of research was qualitatively evaluated on the basis of the existence of and the factual correlation with the research results, while the **number of spin-offs**, **the number of applied instruments** and measures, and the **number of target group-specific publications** were used to measure the relevance of research to target groups and the successful application of research results. In addition to this, process quality was treated as an indirect indicator of societal effects. The evaluation took a positive view of efforts to ensure process quality by applying the concept of "transition cycles" (ibid., p. 9). Internal quality assurance processes at the Wuppertal Institute include a qualitative approach in which impact stories, developed in line with the REF model applied in the United Kingdom, are evaluated using the productive interaction approach (see chapters 1.2. and 2).

The Federal Environment Agency (UBA) was evaluated by the Science Council with a focus on transfer and science-based consultancy (Wissenschaftsrat 2015). Here too, user orientation and the scientific basis of consultancy products were applied as criteria. **Memberships in committees** and the **number of contacts with members of the public** (ibid., p. 56) also served as indicators of societal impact, alongside those previously noted. The lack of guidelines for good policy advice was criticized, confirming the importance afforded to process quality (see above).

Science-based consultancy and the criteria used in its evaluation also formed a particular focus in the recent evaluation of the **Institute for Social-Ecological Research (ISOE)** (Wissenschaftsrat 2016). The evaluation took a positive view of the methodological concept and its role in assuring the quality of research processes as well as internal process-oriented quality assurance mechanisms that facilitate reflection on the impacts targeted. (p. 12).

The evaluation for admission to the Leibniz Association consists of an assessment of the scientific quality of the respective institution, which is carried out by the Science Council and incorporates the criteria and indicators noted above. A scale of four possible grades[8] is applied for the purposes of this assessment. As a result, the assessment leaves less scope for interpretation. In addition to considering the scientific quality of the respective institution/measure, particular attention is paid to its supra-regional significance and its relevance for national science policy. (Wissenschaftsrat 2014).

In the second part of the evaluation process for admission to the Leibniz Association, a panel of experts convened by the Leibniz Association makes a recommendation in the form of a policy statement on the respective institution based on the evaluation report. (Leibniz-Gemeinschaft 2014)

[8] Excellent, very good, good, not sufficient.

Societal effects considered within this context include an institution's **relevance to target groups**[9]**, target group reach**, and **maximization of reach**. Several quantitative[10] and qualitative[11] indicators are referred to in the guidelines of the Leibniz Association.

The **regular evaluations** of Leibniz Institutes with similarities to the IASS are also of interest with respect to their ability to provide potential criteria and indicators for the quality of IASS work that oscillate between the scope of traditional technology transfer and science.

First of all, there are the **museums in the Leibniz Association**. As a considerable proportion of their work does not address science directly, these institutions must operationalize their specific form of "transfer". However, two recent evaluations (*Schifffahrtsmuseum Bremerhaven* and *Deutsches Museum München*) provide only a small number of criteria and indicators that could be applied to the IASS or offer further impetus. Indeed, besides classical quantitative indicators such as publication figures (differentiated according to categories attributable to different target groups), the only other quantitative data collected are visitor figures for the transfer/service area ("mediation") of the *Schiffahrtsmuseum* and the *Deutsches Museum*. Qualitative aspects of interest noted in these evaluations include the existence and quality (in terms of their coherence and factual completeness) of strategic documents as well as the expansion of visitor research activities (Leibniz-Gemeinschaft 2017), p. B6). The area of infrastructure is also addressed primarily from a qualitative standpoint, without an explicit structure in terms of its documentation, assessment and evaluation (on the basis of examples).

The **major science associations** have recognized that the inadequate representation of (societal) effects is a fundamental and widespread problem and have set up a pilot project to address this deficit. The Fraunhofer Society, Helmholtz Association of German Research Centres and the Leibniz Association have carried out a BMBF-funded joint research project on "Sustainability Management in Non-University Research Organisations (LeNA)". Quality criteria for socially sustainable research were developed within this framework in the sub-project "Socially responsible research". However, these criteria are exclusively process-oriented and are accordingly not relevant[12] to the aspect under consideration here (Daedlow, Podhora et al. 2016, Ferretti, Daedlow et al. 2016, Helming, Ferretti et al. 2016).

In addition, the Leibniz and Helmholtz Associations have internal working groups: In 2016, the **Leibniz Association's Knowledge Transfer Working Group** proposed a definition of "transfer" that could be used and elaborated upon by member institutions (Leibniz-Gemeinschaft 2016).

In 2015, the **Helmholtz Association** launched an internal organizational discussion process titled "Strategic development of knowledge transfer within the Helmholtz Association". In a first step, the General Assembly of the Helmholtz Association adopted a position paper on knowledge transfer in June 2015 (Helmholtz-Gemeinschaft 2015) A working group was subsequently established, which presented its findings in a paper published in 2016. In their paper, the working group recommended that impact-focused evaluation processes for knowledge transfer activities also be developed. In place

[9] Is the institution's work of relevance outside of academia, e.g. commercially, politically, culturally, or in any other areas of society?

[10] The number of commercial property rights and patents; the number of consulting contracts and expert reviews; the amount of third party funds raised for research, consulting, services, etc.; income from commercial activity. (ibid. p.)

[11] Existence of especially remarkable research results or other outcomes, existence of in-house quality assurance measures, user orientation, adequacy of research-based services and consulting, suitability of the forms of consulting offered.

[12] They are, of course, highly relevant for the future development of the internal process-oriented quality assurance mechanisms.

of a set of indicators and key performance figures, the working group has proposed a framework to facilitate a tailored approach to measuring success in research centers and programs. The proposal identifies areas for the collection of impact and output indicators, and requires that the individual research areas, centers and programs further specify the nature of relevant impacts. The exact wording: "A further specification or description of outcome and impact indicators using case studies should only be pursued when the initial findings reveal the extent to which knowledge transfer in the area of input or activities/processes can be measured and evaluated and compared across the Helmholtz Association."(Gemeinschaft and Hansjürgens 2016)

1.2 Criteria and indicators in select non-German quality assessment procedures

The integration of impact measurements into national evaluation frameworks has only recently become an issue abroad, but efforts to address this have gained pace since the turn of the millennium. In this regard, the national systems of Great Britain and the Netherlands stand out in particular.

1.2.1 Great Britain: Research Excellence Framework (REF)

The UK has had a national performance-based funding system for its scientific institutions since 1986. Quality assessments conducted as part of an "informed peer review" system determine the allocation of funding. The "Research Excellence Framework" (REF) (REF 2011, REF 2012), which was introduced in 2014, measures and evaluates "societal impacts"[13] as well as the quality and vitality of research: To this end, institutions were called upon to complete an impact template, detailing their respective approach to/strategies for achieving impacts as well as the institutional resources and infrastructure made available for this purpose. In addition to this, they are required to submit one "impact case study" (REF 2011, p. 52) per 15 FTEs.

Each of these case studies, which are also developed using a template, must be underpinned by an "excellent" research output[14] and supported by evidence.[15] Evidence and/or indicators must meet certain conditions,[16] and a list of examples is provided. The main criteria for evaluating these case studies are "Reach", understood as "the extent and diversity of the communities, environments, individuals, organizations or any other beneficiaries that have benefited or been affected"[17] (REF 2012, 74) and "Significance", understood as "the degree to which the impact has enriched, influenced, informed or changed policies, opportunities, perspectives or practices of communities, individuals or organizations" (ibid.).

[13] "Impact is defined as an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia" (REF 2011, p. 26)

[14] "Excellent research means that the quality of the research is at least equivalent to two star: quality that is recognized internationally in terms of originality, significance and rigour." (REF 2011, p. 29).

[15] Critics have argued that this approach is mechanistic and fails to take into account alternative sources of impacts and transfer pathways. (Stern, N. (2016). Building on success and learning from experience: an independent review of the Research Excellence Framework , p. 52).

[16] "The main panel will consider any appropriate evidence that is verifiable. Wherever possible, quantitative indicators should be included. (…) The main panel does not welcome testimonials offering individuals' opinions as evidence of impact; however, factual statements from external, non-academic organizations would be acceptable as sources to corroborate claims made in a case study." (REF 2012: p. 30).

[17] (…) Reach will not be assessed in purely geographic terms, nor in terms of absolute numbers of beneficiaries, but rather based on the spread or breadth to which the potential constituencies have been affected." (REF 2012, p. 34).

1.2.2 The Netherlands: "Standard Evaluation Protocol" (SEP) and "Evaluation of Research in Context" (ERIC)

In the Netherlands, academic research is evaluated using the Standard Evaluation Protocol (SEP). This system, developed and implemented by the national association of universities, has existed since 2003 and is updated with a new version every six years. The current evaluation phase runs from 2015 to 2021 and adheres to the updated protocol (SEP 2014).

The protocol is used to assess and confirm the quality of research and its relevance to society. The results of these evaluations do not have an immediate redistributive effect. Evaluations are carried out at the research unit level by expert committees. The three main criteria[18] considered in these evaluations are the quality of research, its relevance to society, and the viability of research units. "Relevance to society" is assessed on the basis of the "quality, scale, and relevance of contributions" (ibid., p. 7).

In addition to a SWOT analysis and reports on strategies and plans, the evaluated units must also identify indicators for the criteria from a list of options as part of their self-reporting. These indicators are divided across the categories "Output"[19], "Use"[20] and "Recognition"[21], and their focus varies in accordance with the respective criterion. Examples of evidence/indicators are provided for each category and quality domain. The reporting units can choose from these suggestions (or put their own proposals forward) and decide for themselves whether they wish to provide quantitative or qualitative evidence. At least one item of narrative evidence (a case study) must be provided in the case of indicators of societal relevance.[22] As with the REF model, these case studies are commonly used to measure the "societal impact" of individual units and compiled using a template.

The "Evaluation of Research in Context" protocol was published in 2010 and serves as a supplementary protocol to the SEP. The protocol provides a framework for scientific institutions to analyze and evaluate the impacts of their (research) activities beyond the academic system (Eric 2010). The protocol uses a definition of "Societal Relevance"[23] that is consistent with the SEP, but has a strong prospective character in addition to its retrospective components. This prospective evaluation perspective is implemented in the protocol with a focus on "productive interactions" between researchers and non-researchers, applying an approach developed in the 7th Framework Programme for Research "Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society" (SIAMPI) The protocol assumes that

[18] Productivity originally figured as a separate criterion, but was removed for the current version in response to criticisms that the focus on the quantity of output entailed adverse effects for quality and relevance.

[19] Units wishing to document the "Relevance to society" of products under the category "Output", for example, must provide evidence relevant to the indicator "Research products for societal target groups" such as: "(policy) reports, articles in professional journals, outreach activities, public lectures, exhibitions, other output (instruments, infrastructure, datasets, software tools, designs)...", cf. p. 25.

[20] To document "Relevance to society" under the category "User", for example, evidence relevant to the indicator "Use of research products by societal groups" must be provided, such as: "Patents/licences, use of research facilities by societal partners, projects with societal partners, contract research", see Ibid.

[21] To document "Relevance to society" under the category "Recognition", for example, evidence relevant to the indicator "Marks of recognition by societal groups" must be provided, such as: "Public prizes, valorization funding, number of appointments/positions paid for by societal parties, membership of civil society advisory bodies", cf. ibid.

[22] Cf. ibid., p. 14.

[23] "For the purposes of this guide, societal relevance is defined as the degree to which research contributes to and creates an understanding of the development of societal sectors and practice (such as industry, education, policymaking, health care) and the goals they aim to achieve, and to resolving problems and issues (such as climate change and social cohesion); a well-founded expectation that the research will provide such a contribution in the short or long term.", cf. ibid. p. 10.

these productive interactions occur throughout the knowledge production process in different phases and are a necessary, if not vital, prerequisite of societal effects and thus a meaningful proxy for their measurement: "A summary of instances of such interaction is therefore an essential element of the information on a research group's performance. If productive interaction exists between research groups and stakeholders, there is more reason to expect that the research will sooner or later have a societal impact." (ibid., p. 11). Within the four steps for the evaluation of societal relevance proposed in this document, these productive interactions figure both in the self-description and as indicators to be used.

2 An overview of current theory: Impact criteria and indicators from (applied) science studies and evaluation research

The practice of measuring performance across the dimensions of input, output, outcome and impact is well established within the field of quality assurance. However, the distinction between outcome and impact contained therein is both theoretically ambiguous and difficult to implement.[24] This issue is addressed in more recent approaches in German research on transdisciplinarity (Bergmann, Schäfer et al. 2017) as well as older Anglo-American approaches (Walter, Helgenberger et al. 2007, Wiek, Talwar et al. 2014) through the adoption of the term "effect".

The adoption of this terminology has consequences for how impacts/effects are measured that could be instructive for the development of indicators at the IASS. On the one hand, the term "effect" offers an alternative to the problem of causal attribution, which otherwise looms large in impact measurement, by jettisoning the assumption that there must be a clear and exclusive linear causality between "output" and "outcome/impact".

A current inventory of the effects of transdisciplinarity distinguishes between four different types of effects (conceptual impact, capacity building, network effects, improvement of the situation), each of which correlates specifically with different types of results and process properties and displays different efficiencies (Bergmann, Schäfer et al. 2017) Similarly, Wiek et al have proposed that effects be divided into four categories (Wiek, Talwar et al. 2014), one of which (usable products) is in fact an output category that serves as an indirect indicator for the effects expected from these material outputs.

These and similar distinctions can be found in much of the literature reviewed for this paper: a common feature is their differentiation between:

- discursive/cognitive (e.g. conceptual impact, enhanced capacity),
- behavioral (e.g. network effects, action, instrumental impact) and
- structural (e.g. improvement of the situation, structural change) effects.

The range of proposed criteria and indicators for these effects and material output is very diverse. Categorized across the three dimensions noted above, it includes the following[25]:

[24] The OECD (OECD 2002), for example, distinguishes between outcome as a direct effect of an output on an object and impact as a long-term, indirect effect that can also affect adjacent target groups, whereas in applied research on transdisciplinarity (e.g. Walter et al 2007), an impact is an intermediate effect connecting output and outcomes and an outcome a long-term effect resulting from an impact.

[25] The following overview was developed on the basis of various papers that address the ex-post evaluation of the impacts/effects of transdisciplinary research and develop criteria and indicators for this purpose. Papers reviewing the implementation of previously developed criteria/indicators were excluded from this selection together with studies focussed on

Level	Category	Indicators	Source
Discourse/ Cognition	Enhanced capacity: acquired knowledge, understanding, anticipatory competence	Co-produced knowledge or scenarios, trainings for stakeholders or researchers, common language, enhanced communication skills, generating attention/excitement…	(Wiek, Talwar et al. 2014, Belcher, Rasmussen et al. 2016)
	Trust in others	Increase of willingness to cooperate, in readiness to share knowledge, in readiness to identify cooperation, in readiness to leave important tasks in a joint project to others	(Walter, Helgenberger et al. 2007)
	Development of new ideas	Citations (outside of academia) and documentation	(Penfield, Baker et al. 2014)
	Contribution to solving a practical, real-world problem	Broader acceptance of new technologies	(Carew and Wickson 2010)
	Gain and distribution of knowledge (system, goal, transformation)	Number of mentions in private or public discussion, increase of knowledge about system, goals, transformation strategies; number of citations of a report outside of science	(Walter, Helgenberger et al. 2007, Bornmann and Marx 2014)
	Change in public opinion, knowledge exchange	Change of opinion and user perceptions	(Penfield, Baker et al. 2014)
	Change in the attitude of policymakers/decision-makers	Change in the knowledge, understanding and attitudes of policymakers and practitioners, A new interest or attitude toward questions of public interest involving S&T	(Godin and Doré 2007, Meagher, Lyall et al. 2008)
	Impact on society on organizational level	The appearance of new discourses on S&T, the appearance of new styles of intervention or the solution to social problems in speeches, interventions and actions	(Godin and Doré 2007)
Behavior/ Practice	Network effects: networks created or expanded, community created or expanded, trust, accountability	New contacts, boundary-crossing collaboration, expanded participant networks…	(Wiek, Talwar et al. 2014)
	(adapted) application of research results, concepts, products	Extent of use: user groups, area, sales volume etc., extent of benefit for target group, environment and society resp. in sustainability categories	(Wolf, Lindenthal et al. 2013)
	Contacts with actors in practice and society	Contacts in general, dissemination/presentations	(Wolf, Lindenthal et al. 2013)
	Network-building	Number and quality of new ties (survey)	(Walter, Helgenberger et al. 2007)

processes (of interest for process-oriented quality assurance) and purely economic impacts as well as studies drawn from other fields that operate with similar concepts (development studies, public health, etc. – as these are unlikely to be of use in the development of criteria/indicators relevant to the IASS).

	Attempt at behavioral change	Face-to-face communications, number of researchers holding dual posts, number of memberships in advisory committees, number of presentations for lay audiences	(Spaapen, van Drooge et al. 2011)
		Contacts mediated by different information channels (texts, expert reports, guidelines), Transfer of artefacts and services (models, movies, exhibitions, trainings, teaching etc.)	(Krainer and Winiwarter 2016)
	Use of research in making a specific decision or in defining the solution to a specific problem.	Policymakers' self-assessments of use, occurrence of research results in policy documents	(Meagher, Lyall et al. 2008)
Structure/ Organization	Structural changes: economic benefits, policies, decisions made, landscape shift (norms), solutions implemented	Policies/laws passed, new public discourse, new social norm, infrastructure changes, shift in rules of engagement/interaction, new jobs, new business models	(Godin and Doré 2007, Wiek, Talwar et al. 2014)
	Organizational changes	Shift in organizational expectations (roles, responsibilities), new strategic orientations, missions or objectives, an administrative restructuring, the number of people affected by the restructuring	(Godin and Doré 2007, Wiek, Talwar et al. 2014)
Output/ Outcome	Presentation of results in non-scientific settings	Quantity of articles/references in the mass media, articles of practical relevance, workshops/presentations for non-scientific audiences	(Molas, Salter et al. 2002, Bergmann, Schäfer et al. 2017)
	Advisory work	Quantity of lectures at non-academic events and involvement in advisory boards	(Molas, Salter et al. 2002)
	Guidelines/instruments	Recommendations for action in the form of consultations, guidelines, manuals, etc.	(Kaufmann-Hayoz, Defila et al. 2016, Bergmann, Schäfer et al. 2017)
		Number of participating trainees in non-academic training	(Kaufmann and Kasztler 2009)
	Usable and used Products	Intellectual property and patents, software, innovative technologies, consumer goods and services, media articles, handbooks and training manuals, access count	(Kaufmann and Kasztler 2009, Wiek, Talwar et al. 2014, Kaufmann-Hayoz, Defila et al. 2016)
		recognition of usability by users	(Schuck-Zöller, Cortekar et al. 2017)
		Response in practice and society: reaching relevant target groups/multipliers, size/characteristics of target groups	(Wolf, Lindenthal et al. 2013)

Table 1: Proposed criteria and indicators for societal effects and material output

It is notable that to a large degree (applied) research on the effects of transdisciplinary, co-productive or participatory research does not focus directly on impacts but on the process of knowledge production or its application. Within this context, the properties and quality of these processes are frequently viewed as adequate markers and predictors of societal effects and thus the process quality is used as a proxy or indirect indicator for the effect.[26]

The SIAMPI evaluation approach focuses on the bridge between process and effect in the form of "productive interactions" (see above). Here too, it is assumed that these productive interactions are indicative of societal effects and thus offer a meaningful proxy for their measurement. Within this approach, productive interactions are defined as "exchanges between researchers and stakeholders in which knowledge is produced and valued that is both scientifically robust and socially relevant" (Spaapen and van Drooge 2011), p. 212) and divided into direct, indirect (publications, exhibitions etc.) and financial interactions (ibid., p. 113). An interaction is deemed successful "when it leads to efforts by stakeholders to somehow use or apply research results or practical information or experiences" (ibid., p. 112).

3 Evaluation Framework of the IASS

The development of evaluation standards requires a clear picture of the object to be evaluated – i.e. the desired effect of the work of the IASS. In addition to this, efforts to identify proxies and indirect indicators can be supported by the development of a model depicting the interactions between research, transfer, and the resulting effects. This (implicit) image of the object of evaluation and the intended effects of research activities has been outlined in a number of IASS documents (for example, IASS Methodology 2018).

The IASS is positioned as a research institute that seeks to explore, understand, and shape processes of societal transformation with the aim of fostering sustainable development. Its work spans two different but interrelated modes of research. **Transformation research** addresses the conditions of as well as the obstacles to and drivers of societal change. This research is primarily of a descriptive or analytical nature. **Transformative research**, on the other hand, represents an explicit intervention: this research seeks to promote and support processes of societal change by developing concrete approaches and contributing to their implementation. Adopting a self-reflexive perspective, this research also considers the effectiveness and application horizons of its knowledge outputs. As a hybrid think tank and institute for advanced studies, the IASS engages with and in processes of both societal change and the systematic, self-critical reflection of the epistemological and ethical foundations of these forms of knowledge generation. IASS research activities address three forms of knowledge in particular, which structure transdisciplinary research processes at the institute and perspectives on the objects of this research. These are: *system knowledge*, which deals with the conditions, obstacles and drivers for sustainability transformations; *orientation knowledge*, which deals with the processes by which ideas of sustainable development are negotiated within societies and the ethical underpinnings of these ideas; and *transformation knowledge*, which deals with the design and support of concrete transformation processes. The Fellow Program supports both of these research modes. Firstly, by bringing outstanding researchers to the Institute from all over the world as part of its program for academic junior and senior fellows and integrating them into the aforementioned processes. Complementing this, the IASS seeks to initiate processes of co-creation by inviting Practice Fellows to join ongoing knowledge production processes at the institute in order to transfer knowledge acquired at the IASS into practice when they return to their fields of work.

[26] These criteria will be used for the internal process-oriented quality assurance mechanisms, which will be developed after the methodological approach has been finalized.

Following the Science Council's stipulation (Wissenschaftsrat 2014) the criteria for the evaluation can be distilled from this self-description. Accordingly, the overarching evaluation questions shall be:

Are efforts to support transformations toward greater sustainability successful across politics, society, and the private sector?

This criterion for the societal impact of IASS research activities can be broken down into four sub-criteria:

- Have research activities improved the understanding of sustainability in politics, society and business?
- Do the results of IASS research activities inform political decision-making and/or public debates and discussion processes?
- Did research make a positive contribution to existing decision-making and implementation routines?
- Has research resulted in the successful development of collaborative design processes for transformations toward sustainability?

In order to answer these evaluative questions, the criteria discussed in this section must be translated into appropriate indicators for both functions – in other words, for research on and research for transformations. In addition, the criteria and matching indicators that the scientific peers and the Science Council are likely to apply need to be considered and integrated. Finally, the indicators must be organized in a logical and theoretically grounded framework. The next section introduces our proposal for this integrated framework.

3.1 Theoretical reformulation of the impact assessment model

In order to create a comprehensive evaluation framework with a solid theoretical foundation (as desired by the Advisory Board), it would be advisable to revise the description of the mechanisms of action in light of relevant theoretical concepts and, within this framework, to select and systematize the above-mentioned proposals for criteria and appropriate indicators.

3.1.1 Theoretical reformulation of the overall effect model: "Multi-level transition theory".

The **multi-level transition perspective** developed by Geels (Geels 2002) offers a suitable tool with which it is possible to reformulate the IASS overall effect model. The multi-level transition perspective emphasizes the importance of **interactions between innovations, policy regimes and socio-cultural landscapes** in which complex processes of change are embedded. The multi-level transition theory postulates that change in socio-technical systems occurs at the intersection of two processes: a first process in which innovative ideas and new knowledge-integrating insights gather and pressure for action builds on the prevailing policy regime, and a second process in which changes in the socio-cultural landscape (e.g. change in cultural attitudes) open windows of opportunity for the implementation of new movements and innovative processes on a broader level.

In this theoretical framework, the **IASS** can be conceived **as an innovator** in German (and international) policy regimes. As such, the IASS is subject, on the one hand, to the constraints of the socio-cultural landscape as it analyzes and identifies existing barriers to transformation. On the other hand, operating in the mode described above, the IASS can act as a catalyst on multiple levels.

Williams (Williams 2017) has applied this theoretical framework to the concepts of impact evaluation developed by Wiek et al. (see above). In his framework he adopts the suggestion of Wiek et al. that first-order effects such as products, cognitive capacities/new narratives and discourses, and networks contribute to second-order effects (such as structural changes and actions) through application, coordination and reinforcement (Wiek, Talwar et al. 2014). The combination of first- and second-order effects in turn contributes to changes in social practice. At the same time, broadly shared alternative ideas support structural change, because innovative policies appear to policymakers to be less risky if they are in line with public perceptions. This effect is captured by Taylor (Taylor, 2004) in his concept of "social imaginaries", which provide defined "repertoires of possibilities" for society. If an imaginary changes, the repertoires of possibilities change with it.

Research Opportunity 1a: Theoretical Modeling of Policy Effects

Different theoretical frameworks have already been merged, roughly adapted to the IASS and expanded by or specified in individual aspects. However, this adjustment is not yet complete and needs further research. Individual aspects should be further developed to better reflect the specific features of the IASS: The topics dealt with by the IASS – air quality, the climate, the oceans and the Arctic, and the energy transition – are by no means merely innovative fields of application, but are broadly anchored cross-cutting topics in research, politics and society. Here, the innovation activities of the IASS would have to be more strongly modeled in terms of development lines, scenarios and visions for the future. Unlike the multi-level perspective of transition theory, the IASS focuses less on technical innovations that are replacing the mainstream. Rather, it focuses on the economic, political or social practices that need to be reoriented towards sustainable development. Other approaches from the tradition of Science & Technology Studies (Berkhout, Smith et al. 2004, Loorbach and Rotmans 2010) place greater emphasis on these socio-economic elements and thus offer a helpful theoretical supplement to the specifics of the IASS.

Translated into the IASS evaluation framework and viewed within the context of the multi-level transition perspective, the societal effects generated by IASS research activities would occur across three levels:

- First-order effects generated by IASS activities would influence the contents and scientific underpinnings of emerging narratives and imaginaries, strategic discourses, existing networks, and existing capacities (networks extend to the regime level).
- Second-order effects generated by IASS activities would include effective process innovations toward policy and organizational change located at the regime level and contribute to efforts to foster transformations toward sustainability.
- In the multi-level transition approach, effects of the third order refer to the long-term and permanent diffusion within social imaginaries and lived social practice or, in Wiek's words, within the "fabrics of socio-cultural landscapes".

The interplay of these levels is illustrated in the following diagram based on Williams:

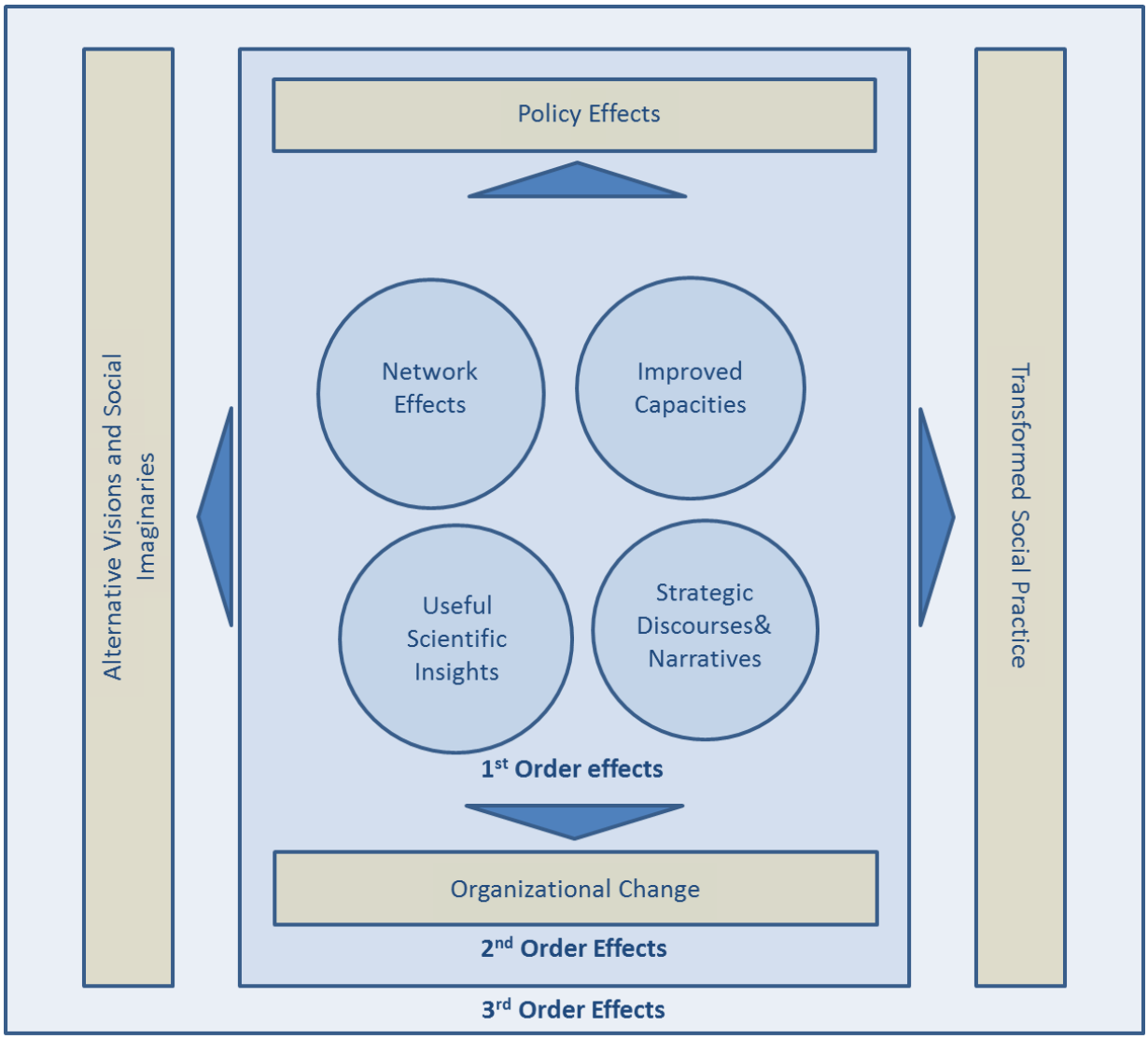

Figure 1: Societal effects (based on Williams 2017, op. cit., p. 6)

The following overview integrates the indicators developed on the basis of the quality model (IASS 2017) and some indicators generated from the above evaluations within this perspective:

	Category	Effect	Indicators
1. Order	Useful scientific insights	• Technologies & societal innovations • Products/services • Instrumental knowledge • Publications	• **Scientific advice** (based on own research and/or evaluation of currently available knowledge; ability to conduct advisory activities independently; transparent procedures, relevance of advisory activities to target groups) • **Publications** for non-scientific audiences • **Events** for non-scientific audiences • **Products/materials** for non-scientific audiences
	Improved capacities	• Increased cognitive capacity among relevant non-scientific partners • Improved anticipative capacity among relevant non-scientific partners	• improved **understanding** among non-scientific partners • new **knowledge** among relevant non-scientific partners • improved **communication capacities** • improved ability to anticipate future developments
	Network effects	• Improved networks and communities	• Quantity of **networks** created or extended • Cross-system collaborations and partnerships • New **contacts** • Enhanced **trust** within networks • **Strengthened identity** within communities • Willingness to share with other networks
	Strategic discourses and narratives	• New/altered strategic discourses • New/altered public narratives	• Quantity of **new narrative elements** within strategic discourse • **Quality of changes** to strategic discourse

2. Order	Policy effects	• knowledge-based policies • knowledge-based decisions • implemented solutions • instrumental or conceptual application of knowledge • economic benefits	• Quantity/quality of knowledge-based **policies & laws** adopted • Quantity of knowledge-based decisions • Change in **policy-related discourse**
	Organizational change	• Changed organizational context for current and future work • New organizations • New standards • New partnerships • New business models	• Changes in **responsibilities** and **roles** • Changes in the rules of participation/organization of processes • New **institutional framework** • Changes in **investment strategies**
3. Order	Alternative visions and social imaginaries	• Changes in collective goals and visions • Greater social cohesion across groups and beyond	• Quality/resilience of new **social imaginaries** • Quality/resilience of new **visions**
	Transformed social practice	• Changes in the socio-cultural landscape • Adjustment of or change in standards • Inclusion of new actors or themes in public spaces and discourses • Changes in practices of participation • New spaces for innovation and experimentation	• **Changes in the behavior of collective actors** • Increased **capacity for collective action** • New **forms of participation** • Deeper networks and **solidarity**

Table 2: Overview Criteria and Indicators

3.1.2 Theoretical Reformulation of the interface between process and effect: "Productive Interaction".

In addition to the effect indicators outlined by Williams, the evaluation approach put forward by Wiek et al. focuses on process properties as an important prerequisite for non-scientific effects. According to Wiek et al., empirical research has shown that certain properties are conducive to the generation of societal effects. These include, on the one hand, the nature of the participatory research process (e.g. number, type and sequence of participatory events), and on the other hand the quality of the participatory research process (e.g. adequate representation of the full spectrum of opinions and perspectives (cf. Wiek et al 2014, *op. cit.*, p. 124).

> ***Research Opportunity 2: Process-Effect link***
>
> The mutual relationships between the quality and methodology of the research process (e.g. co-production and transdisciplinary research methods) and its societal effect are not yet fully explored. The simultaneous adjustment of the research process according to the novel methodology on the one hand and the continuous measurement of the effects on the other hand offer a unique opportunity to contribute to the research concerning these interactions.

The SIAMPI approach, which is also applied in the Dutch evaluation system and the internal quality assurance system of the Wuppertal Institute, complements this process perspective and uses the process property of "productive interaction" as a proxy for non-scientific effects[27]. According to Spaapen and Drooge (Spaapen, van Drooge et al. 2011), this compensates for issues arising in connection with the large time lag and the unclear attribution of effects.[28]

These issues in the area of impact measurement are particularly pertinent to the IASS: if the reporting period for the 2020 evaluation spans 2015–2019, it will still be far too early to measure most effects (especially third-order effects) due to the anticipated time lag. In order to offset this, on the one hand an internal process-oriented quality assurance mechanism implementing the novel methodology (IASS 2018, see 4.1.1.) should be developed, which can then serve as an indirect indicator of societal effects (see above in the examples of the Science Council Evaluations of the WI and the ISOE).

On the other hand and in addition, in light of the **"productive interaction" model's** sound theoretical and empirical basis, its proven track record in evaluations at the regional and organizational level, and its compatibility with the approach of the IASS, this process-related approach should be directly integrated into the framework.

Within this approach, productive interactions are defined as "exchanges between researchers and stakeholders in which knowledge is produced and valued that is both scientifically robust and socially relevant" (Spaapen, van Drooge 2011, p. 212). These are distinguished from "normal" interactions by the effects which they successfully bring about: "when it [the interaction, js] leads to efforts by stakeholders to apply research results to social goals, i.e. when it induces behavioral change" (Spaapen, van Drooge et al. 2011).

Productive interaction involves contact between researchers and a social group using a typology of facilitating tools. This approach distinguishes between three main types of productive interactions pitched toward particular anticipated effects:

---

[27] Thus, this claim clearly goes beyond the one formulated by Wiek et al., who state that the studies on which the synopsis of the process properties is based "(…) also indicate to be cautious in attributing effects exclusively to the participatory features of the research process – there are other factors, including financial incentives that influence what societal effects the research process yields." (Wiek et al. 2014, *op. cit.*, p. 123).

[28] This approach employs the concepts of "contribution" and "uptake" in place of the problematically linear concepts of "attribution" and "temporality" (see Spaapen, van Drooge et al, 2011, *op. cit.*, p. 31).

Productive interaction	Social Impact	Stakeholder	Indicators (e.g.)
Direct, personal	Behavioral Change	One-to-one, personal and professional networks	Face-to-face communications, number of researchers holding dual posts, number of memberships in advisory committees, number of presentations for lay audiences
Indirect, media	Uptake, use	Different audiences	Quantitative indicator "Contextual Response"
Financial or in-kind support	Collaboration	Joint Projects	Contracts, licenses, projects grants, sharing of facilities, personal sponsorships

Table 3: Overview of productive interaction types, abridged version from Spaapen, van Drooge (2011), a.a.O., p. 2 and Spaapen, van Drooge et al. 2011, a.a.O., p. 217

By qualifying "productive interactions" in terms of their real impact on "stakeholders", the proposed approach can be precisely integrated into the framework proposed by Williams and supplemented by the process properties developed by Wiek et al. without extending the subject area too much – namely to all interactions. In addition, much of the data necessary for this purpose can be collected and interpreted in other contexts.

Applied to the IASS evaluation framework developed here, process-oriented indicators for measuring the quality of interactions would be added in order to demonstrate first- to third-order effects (which are far removed in time). This would result in the following addition to the diagram developed above:

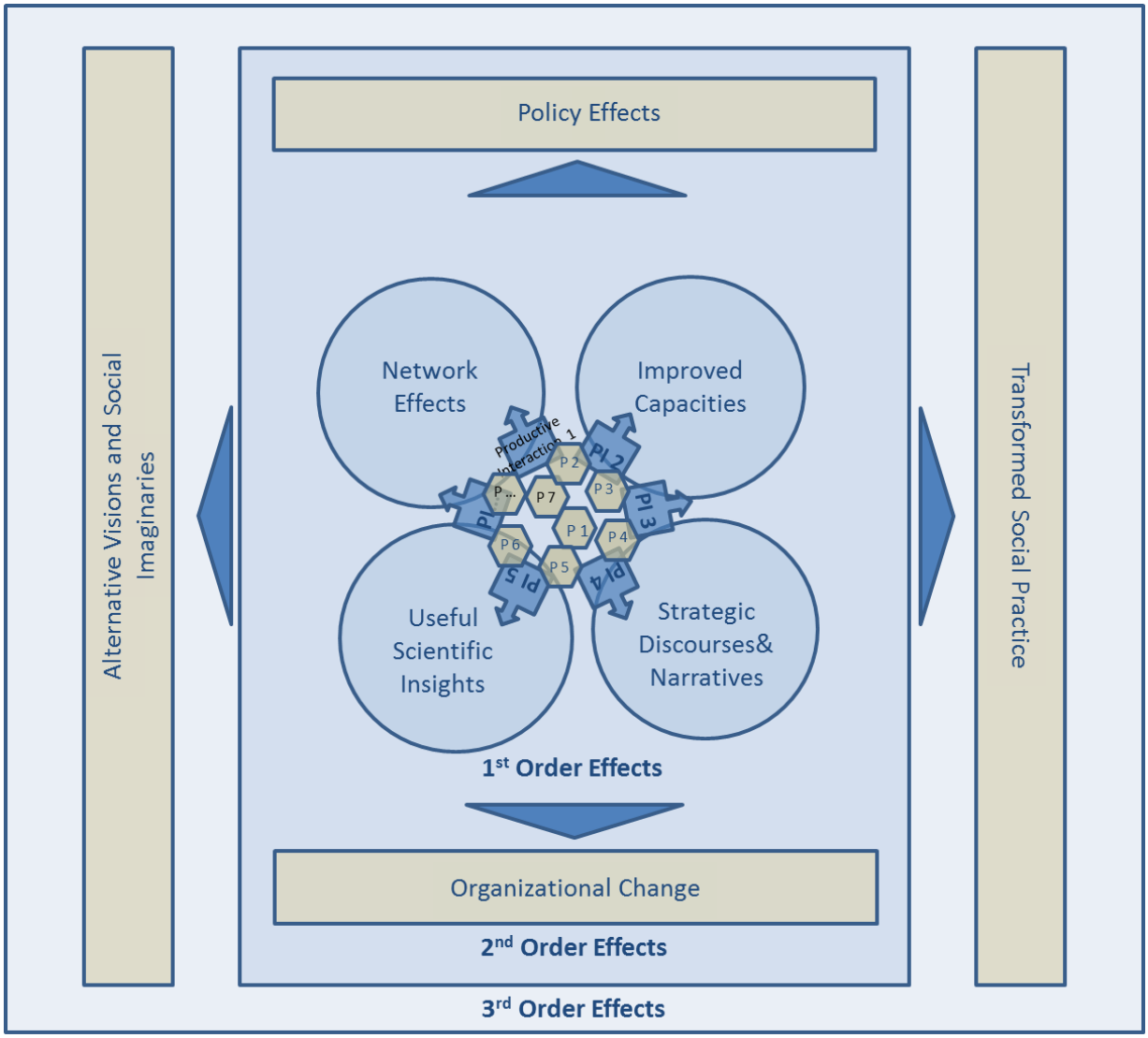

Figure 2: Societal effects and productive interactions, based on Williams 2017, ibid. and Spaapen, van Doogen et al. 2011, ibid.

Accordingly, the framework laid out above needs to be extended (for a complete version of the framework, see appendix 1) to include indicators that cover both "processes" and "productive interactions":

	Category	Indicators
Process	**Nature of the participatory research process**	• **Participative events** (quantity, type, phase, sequence) • **Stakeholders' motivation** in participating • **Adequacy of the role** of stakeholders in events • Perceived importance of events
	Quality of the participatory research process	• Proper **representation** of the range of opinions and perspectives • Fulfillment of the most important **roles for participation** • Sufficient level of **interaction** • Adequate appreciation and **processing of stakeholder input** • Successful representation and resolution of dissent and conflict • **Diversity** of participation activities
Prod. interaction	**Direct productive interaction**	• Individual discussions, participation in **advisory boards**, **lectures** to non-scientific audiences
	Indirect productive interaction	• "Contextual response"
	Productive interactions of a financial nature	• Contracts, **licenses**, level/source of **project funding**

Table 4: Processes and Productive Interactions: Extension to the Criteria and Indicators

3.1.3 Linking Evaluation Questions to Effects Categories

To clarify the links from the categories developed with the help of the theoretical framing and the guiding questions deduced from the self-description pointed out in the introduction to the Chapter, the following table can be helpful. Note that some categories apply to multiple evaluation questions. We have identified the primary link for simplicity. Further development of this evaluation framework will align data collection and analysis methods to questions and effects categories (See Section 4 for more detail on methods)

Evaluation Question	Effect Category	
Are efforts to support transformations toward greater sustainability successful across politics, society, and the private sector?	• Alternative visions and social imaginaries • Transformation of social practice	*Productive Interactions of different sort as indirect indicators*
Have research activities improved the understanding of sustainability in politics, society and business?	• Useful scientific insights • Improved capacities	
Do the results of IASS research activities inform political decision-making and/or public debates and discussion processes?	• Strategic discourses and narratives • Useful scientific insights • Improved capacities	
Did research make a positive contribution to existing decision-making and implementation routines?	• Policy effects • Organizational change	
Has research resulted in the successful development of collaborative design processes for transformations toward sustainability?	• Organizational change • Network effects	

4 Data collection and analysis

The impact studies from the field of science studies and evaluation research evaluated in this context propose a variety of methods for the collection and analysis of data, which could be applied to the criteria and indicators developed for this framework. In most cases, however, these studies focus on the aggregation level of a single project or process. Applying these methods to the IASS with its research program of currently 30 projects would necessitate the collection and analysis of an enormous quantity of data. For example, Williams (2018, *op.cit.*) lists eight different methods for data collection alone in his concept paper for the evaluation of one, albeit large project/process (ibid., p. 29). These methods include participant observation at events, actor network tracing, document analysis and prototype tracking as well as structured interviews with more than 50 people (ibid., p. 30). Even accounting for overlap across projects and related policy fields, implementing these methods at the IASS would require that an unmanageable number of interviews be carried out, processed and analyzed.

In light of this, the IASS evaluation framework should draw upon the survey and analysis methods applied in the evaluation systems of England and Holland, which also function well with larger samples, and then systematically supplement these with further selected elements. In addition, the large number of IASS projects and minimal capacity for accompanying research requires a pre-structured approach to data collection, the use of existing data or data generated for other purposes and processes, as well as the involvement of project personnel in the acquisition of information. An overview of how indicators might be assigned to various data collection and evaluation methods can be found in the appendix (Appendix 1). A detailed explanation of the individual methods is provided in chronological order in the following sub-chapters.

4.1 Step 0: Externalizing requirements/secondary utilization

4.1.1 Process-oriented quality assurance

As previously noted in an outline of the comprehensive quality model (IASS 2017, *op. cit.*, p. 7), many of the process quality indicators put forward in scientific discourse were developed to facilitate internal steering and necessary formative functions – this is particularly the case in transdisciplinary contexts (e.g. Bergmann (2005), Jahn and Keil (2015)). The methodology currently in development at the IASS also has its basis in process-related quality criteria. This reflects the framework outlined above, in which process quality assurance mechanisms serve as a proxy for second- and third-order effects (see 3.2.1).

In the case of the IASS, the internal process-oriented quality assurance mechanism for projects should include these criteria and indicators and implement them, for example, by means of guidelines (Jahn and Keil 2015), impact strategies at the institute and project levels, or by means of reporting requirements such as project storyboards (see Wiek et al. 2014, p. 128). As the examples of the ISOE and Wuppertal Institute show, this type of impact assurance can be used to measure process quality indirectly. This approach is also compatible with the expectations of the Science Council.

[29] Moreover, Williams employs five evaluation methods, each of which is quite complex in itself (actor network tracing, outcome harvesting, process tracing, most significant change, contribution analysis).

[30] Interviews with 38 participants, Bellwether interviews with 10 participants and an as yet undetermined number of interviews with (process) participants and team members.

Individual instruments within this quality assurance system could also be used in the context of the ex-post evaluation of societal effects if they were configured and adapted accordingly. Thus, impact strategies should be developed at the institute and project levels, in which, in addition to a prospective presentation of the Theories of Change, the concrete desired effects are also listed, including the target groups and their implementation over time. These would then inform the selection of Bellwether interview partners as part of the in-depth analyses of individual cases, and contribute to efforts to attribute individual effects to IASS activities.

4.1.2 External media analysis

Every year, the IASS commissions an external service provider to conduct a media evaluation, the individual parts of which can be used for the ex-post evaluation of the Institute's effects. It covers individual indicators directly, but can also be adapted to meet individual requirements – most recently for the Climate Engineering Conference CEC 2017. Insight is also provided into the effectiveness of messages within selected user groups.

4.2 Step 1: Continuous data collection using pre-defined formats

4.2.1 Case studies of impacts/Impact stories using pre-defined templates

The REF and the Dutch SEP frameworks both make use of structured effect stories to collect data. However, while the REF requires just one impact story per 15 FTE, each project at the IASS should contribute at least one effect story per reporting period.

The effect stories should be structured in a manner that reflects the frameworks described above and the individual stories should be linked to specific effect categories[31] in order to embed them within the framework (multiple answers possible). A proposal to this effect can be found in the Appendix 2. It corresponds to the overview of the assignment of indicators and data collection and evaluation methods (see Appendix 1).

Each of these case studies, which are also developed using a template, must be underpinned by an "excellent" research output (fulfilling the WR's requirement for research-based consulting) and supported by evidence. These certificates or indicators must meet certain conditions[32] borrowed from the REF.

4.2.2 Output and impact reporting in the research database

Last year, the IASS introduced a system for reporting effects of the first, second and third orders. Building on the databases already existing at the Institute (project database, publication database, event database), a comprehensive research information system was set up and associated reporting channels defined. In accordance with a memorandum approved by the Board of Directors on the subject of "Activity, awards and impact reporting", numerous effects in terms of outputs and outcomes have already been defined in this system, which have been expanded, refined and adapted

[31] 1. Productive interaction: a) direct, b) indirect, c) financial); 2. Capacities: a) understanding, b) knowledge, c) communication; 3. Networks: a) expansion, b) crossing borders, c) contacts, d) trust, e) identity, f) sharing; 4. New discourses/narratives: a) narrative element, b) qualitative change discourse; 5. Policy effect: a) legal/regulatory, b) decision, c) discourse; 6. Organizational change: a) role/responsibility, b) process/rules of participation, c) institutional framework, d) investment strategy; 7. Alternative visions and imaginaries: a) imaginaries, b) visions; 8. Social practice: a) behaviour, b) capacity

[32] Verifiable, if possible quantitative, not individual testimonials but factual statements, cf. REF (2014).

on the basis of consultations with the IASS Advisory Board. In its current form, a large part of the data required for the framework is already collected by the researchers by means of structured self-entry.

However, the existing tables and definitions of the database must be extended and modified in order to better meet the requirements of the comprehensive evaluation framework. The structured effect stories should also be recorded and stored in the database. This requires an adjustment of the schema. Furthermore, the definitions still in force must be adapted to the requirements developed here. This will affect numerous details, but also more fundamental considerations such as the cross-references to outputs necessary in the effect stories and an expansion of process and product descriptions to include target groups and time sequences (e.g. target group reference of publications). A detailed overview of the required adjustments can be found in Appendix 1.

4.3 Step 2: Identification of effects of particular interest

It is evident that the IASS lacks the personnel to process all of the effect categories presented in this evaluation framework across all of its projects. Individual cases will have to be selected as an intermediate step. To support this, information collected in the research database on the effects of individual projects should be evaluated at regular intervals.

In addition to the strategic value for the overall mission of the institute, the (potential) relevance and scope, pragmatic criteria (data availability, diversity of effects and topics, etc.) should also be applied as selection criteria for the evaluation of exemplary individual cases.

Research Opportunity 3: Case Study Comparisons

Development and use of this evaluative framework opens the possibility of cross-case comparisons with other institutions and projects working towards sustainability transition and transformation. Much could be learned by comparing IASS process, methods, and effects with transdisciplinary research centres such as Mistra Urban Futures based in Gothenburg, Sweden (focusing on multi-city urban sustainability transition), Sustainability Transition and Intervention Research Lab at Arizona State University, Institute for Resources, Environment and Sustainability at the University of British Columbia, or the Dutch Research Institute for Transitions (DRIFT) based in Rotterdam.

4.4 Step 3: Further collection and analysis of select cases

4.4.1 Contextual Response Analysis (CRA)

The literature proposes a procedure for the collection of data for effects in the category "Productive interaction – indirect" in which a structured Internet search is used to identify relevant observers to change processes (Spaapen, van Drooge 2011, *op.cit.*, p. 217). These "testimonials" can be categorized by user group, policy area, effect type and more. This procedure facilitates the development of a quantitative indicator – "contextual response analysis (CRA)" – which can also be assessed through an external service provider. This option should be carefully considered and trialed in order to ascertain whether significant benefits can be derived from this approach when compared to the findings generated by an annual media analysis; or, whether a similar indicator can be generated for selected outcomes through minor modifications to the analysis.

4.4.2 Context Monitoring (Williams)

Williams has proposed the use of context monitoring to measure effects in the category "Alternative Visions and Imaginaries" (Williams 2017, *op.cit.*, p. 16). This approach also makes use of computer-aided content analysis. The focus here, however, is on a defined subject area rather than outputs. In the case presented by Williams, the data sets comprised press releases and speeches issued by public

authorities, which were analyzed using a keyword search and linked to the project under investigation. According to Williams, this delivers insights into the developmental dynamics of the discourse in certain subject areas, which can then be triangulated with other data (interviews or other instruments such as Issue Crawler, GoogleScraper, etc.) to ascertain their validity (Chilvers and Kearnes 2016) p. 273). The use of triangulation mirrors the approach taken by Spaapen and van Drooge, and provides a comprehensive picture of unintended and parallel effects that can serve as evidence of an effect on the development of alternative visions and imaginaries.

Various questions arise for the IASS evaluation framework when applying this method. For example, it is not clear which "topics" would provide suitable starting points. It would also be necessary to clarify which data sets could be selected and to what extent they are in fact accessible. In light of this, this method should be trialed, focusing on an individual subject area. This could be accomplished in cooperation with Steve Williams, who will be a Fellow at the Institute in the second half of the year.

4.4.3 Interviews (Bellwether interviews, interviews with participants, interviews with team members)

In the case of a number of effects, interviews should be used to collect relevant data. In addition to semi-structured interviews with project/process participants and interviews with team members and project partners, in which mainly data on first- and second-order effects are collected, Williams proposes that so-called "Bellwether interviews" be utilized. These interviews are conducted with people who are regarded as leading figures in certain policy areas and are therefore able to offer relevant observations on effects. In the context of this evaluation framework, these individuals could be identified in advance by project leaders as part of a future process quality assurance system. In addition to this, further interviews should be conducted with stakeholders, team members and cooperation partners on a case-by-case basis.

Bellwether interviews should be explored extensively in an expert workshop with regard to both the concrete interview technique and the identification mechanism.

4.4.4 Surveys

Numerous authors suggest that standardized written surveys of stakeholders and/or other parties be used to collect information on attitudes, opinions, knowledge, behaviors and changes therein.

For example, Walter et al (Walter, Helgenberger et al. 2007) conducted a multi-attribute utility analysis in a controlled pre-post design with a sample of 188 stakeholders based on a written, standardized survey. Wiek at al. (2014, *op. cit.*) also put a standardized, written survey of a large number of stakeholders at the center of their data collection and through which they gain information that serves as indicators for various effect categories (such as growing trust in networks).

Within the IASS evaluation framework, this form of data collection would provide a means to generate direct evidence (e.g. of increased trust) that would otherwise be extremely difficult to source. However, surveys of this type could only be carried out on an exemplary basis in connection with key topics and/or large projects. Here too, experts should be consulted in order to ascertain that data generated in this manner could not be otherwise obtained with less effort (e.g. through social media analysis).

4.4.5 Other case-specific methods of analysis

Numerous methodological proposals have been made in the literature for the scientific analysis of individual effects at project level. To name but a few, these are, for example, the methods of "Process tracing"[33], "Most significant change analysis"[34] procedure and the "contribution analysis"[35] technique, which have been tested in development cooperation and political science.

The use of these methods of analysis should be decided on a case-by-case basis depending on the focus of interest (which particular effect/effects?) and the availability of relevant data (are data readily available or easily accessible?). Here too, a careful consideration of the cost-benefit ratio as a whole and for each individual case should be carried out.

5 Questions for the Advisory Board

- ➔ Are there any other examples of impact assessment at the institute or state levels that could be instructive for the future IASS framework?
- ➔ Is the theoretical framing coherent and conclusive?
- ➔ Are you aware of other indicators of societal effects that could complement the proposed framework?
- ➔ Which methods of documentation, measurement and analysis (noted here or otherwise) are particularly appropriate for the proposed indicators? Could they be implemented without requiring disproportionate efforts?

[33] This allows causal hypotheses to be tested by systematically collecting and discussing evidence both in favor of and against the assumption (cf. e.g. Beach, D. (2016). "It's all about mechanisms – what process-tracing case studies should be tracing." New Political Economy **21**(5): 463-472.

[34] Here, participant interviews are used to collect a sample of significant-change-stories and to illuminate the underlying causation models. In conjunction with other data, these interviews are used to investigate causal relationships or, at minimum, contribution chains. (cf. for example Davies, R. and J. Dart (2005). The 'Most Significant Change' (MSC) Technique. A Guide to Its Use.

[35] This method facilitates a systematic review in six steps of the causation models associated with particular effects. In the process, a list of the premises behind the Theory of Change is developed and relevant evidence categorized accordingly. (cf. for example Mayne, J. (2012). "Contribution analysis: Coming of age?" Evaluation **18**(3): 270-280.

6 References

Beach, D. (2016). "It's all about mechanisms – what process-tracing case studies should be tracing." New Political Economy **21**(5): 463-472.

Belcher, B. M., K. E. Rasmussen, M. R. Kemshaw and D. A. Zornes (2016). "Defining and assessing research quality in a transdisciplinary context." Research Evaluation **25**(1): 1-17.

Bergmann, M., M. Schäfer and T. Jahn (2017). Wirkungen verstehen und feststellen. Arbeitspapier aus dem BMBF-Verbundprojekt TransImpact.

Bornmann, L. and W. Marx (2014). "How should the societal impact of research be generated and measured? A proposal for a simple and practicable approach to allow interdisciplinary comparisons." Scientometrics **98**(1): 211-219.

Carew, A. L. and F. Wickson (2010). "The TD Wheel: A heuristic to shape, support and evaluate transdisciplinary research." Futures **42**(10): 1146-1155.

Chilvers, J. and M. Kearnes, Eds. (2016). Remaking Participation: Science, Environment and Emergent Publics, Routledge.

Davies, R. and J. Dart (2005). The 'Most Significant Change' (MSC) Technique. A Guide to Its Use.

Eric (2010). Evaluating the societal relevance of academic research: A guide.

Geels, F. W. (2002). "Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study." Research Policy **31**(8): 1257-1274.

Gemeinschaft, H. and B. Hansjürgens (2016). Wissenstransfer in der Helmholtz Gemeinschaft. Konzept zur strategischen Weiterentwicklung und Stärkung. , Arbeitsgruppe Wissenstransfer der Helmholtz-Gemeinschaft.

Godin, B. and C. Doré (2007). Measuring the Impacts of Science: Beyond the Economic Dimension HIST Lecture. Helsinki Institute for Science and Technology Studies.

Helmholtz-Gemeinschaft (2015). "ECKPUNKTEPAPIER. Strategische Weiterentwicklung des Wissenstransfers in der Helmholtz-Gemeinschaft."

IASS (2017). "Kriterien für die Evaluation der IASS."

Kaufmann-Hayoz, R., R. Defila, A. Di Giulio and M. Winkelmann (2016). Was man sich erhoffen darf – Zur gesellschaftlichen Wirkung transdisziplinärer Forschung. Transdisziplinär forschen – zwischen Ideal und gelebter Praxis. Hotspots, Geschichten, Wirkungen. . R. Defila and A. Di Giulio. Frankfurt a.M., Campus**:** 289-327.

Kaufmann, A. and A. Kasztler (2009). "Differences in publication and dissemination practices between disciplinary and transdisciplinary science and the consequences for research evaluation." Science and Public Policy **36**(3): 215-227.

Krainer, L. and V. Winiwarter (2016). "Die Universität als Akteurin der transformativen Wissenschaft: Konsequenzen für die Messung der Qualität transdisziplinärer Forschung." GAIA - Ecological Perspectives for Science and Society **25**(2): 110-116.

Leibniz-Gemeinschaft (2014). Grundsätze des Evaluierungsverfahrens des Senats der Leibniz-Gemeinschaft in der Fassung vom 17. Juli 2014 Leibniz-Gemeinschaft.

Leibniz-Gemeinschaft (2016). Wissenstransfer in der Leibniz-Gemeinschaft. Definitionsvorschlag aus dem AK Wissenstransfer als Arbeitsgrundlage für die Mitgliedseinrichtungen.

Leibniz-Gemeinschaft, D. S. d. (2017). Stellungnahme zum Deutschen Schiffahrtsmuseum – Leibniz-Institut für deutsche Schifffahrtsgeschichte, Bremerhaven (DSM)

Mayne, J. (2012). "Contribution analysis: Coming of age?" Evaluation **18**(3): 270-280.

Meagher, L., C. Lyall and S. Nutley (2008). Flows of knowledge, expertise and influence: a method for assessing policy and practice impacts from social science research. Research Evaluation. **17:** 163-173.

Molas, J., A. Salter, P. Patel, A. Scott and X. Duran (2002). Measuring Third Stream Activities.

Penfield, T., M. J. Baker, R. Scoble and M. C. Wykes (2014). "Assessment, evaluations, and definitions of research impact: A review." Research Evaluation **23**(1): 21-32.

REF (2011). REF 2014: Assessment framework and guidance on submissions.

REF (2012). REF 2014: Panel criteria and working methods.

Schuck-Zöller, S., J. Cortekar and D. Jacob (2017). "Evaluating co-creation of knowledge: from quality criteria and indicators to methods." Adv. Sci. Res. **14**: 305-312.

SEP (2014). Standard Evaluation Protocol 2015 – 2021. Protocol for Research Assessments in the Netherlands.

Spaapen, J. and L. van Drooge (2011). "Introducing 'productive interactions' in social impact assessment." Research Evaluation **20**(3): 211-218.

Spaapen, J., L. van Drooge, T. Propp, B. van der Meulen, T. Shinn, A. Marcovich, P. van den Besselaar, S. de Jong, K. Barker, D. Cox, K. Morrison, T. Sveinsdottir, D. Pearson and B. D'Ippolito (2011). Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society (SIAMPI) final report.

Stern, N. (2016). Building on success and learning from experience : an independent review of the Research Excellence Framework

Walter, A. I., S. Helgenberger, A. Wiek and R. W. Scholz (2007). "Measuring societal effects of transdisciplinary research projects: Design and application of an evaluation method." Evaluation and Program Planning **30**(4): 325-338.

Wiek, A., S. Talwar, M. O'Shea and J. Robinson (2014). "Toward a methodological scheme for capturing societal effects of participatory sustainability research." Research Evaluation **23**(2): 117-132.

Williams, S. (2017). Evaluating transition experiments in times of rapid change. . International Sustainability Transitions 2017. Gothenburg, Sweden.

Wissenschaftsrat (2007). Kriterien des Ausschusses Ressortforschung für die Begutachtung von Bundeseinrichtungen mit FuE-Aufgaben.

Wissenschaftsrat (2012). Stellungnahme zum Wuppertal Institut für Klima, Umwelt, Energie GmbH, Wuppertal. Bremen

Wissenschaftsrat (2014). Aufgaben, Kriterien und Verfahren des Evaluationsausschusses des Wissenschaftsrates. Greifswald, Geschäftsstelle.

Wissenschaftsrat (2015). Stellungnahme zum Umweltbundesamt (UBA), Dessau-Roßlau. Saarbrücken.

Wissenschaftsrat (2016). Empfehlungen zur Spezifikation des Kerndatensatz Forschung.

Wissenschaftsrat (2016). Stellungnahme zum Institut für sozial-ökologische Forschung (ISOE), Frankfurt a. M. Kiel

Wissenschaftsrat (2016). Wissens- und Technologietransfer als Gegenstand institutioneller Strategien. Positionspapier. Weimar.

Wolf, B., T. Lindenthal, M. Szerencsits, J. B. Holbrook and J. Heß (2013). "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 **22**(2): 104-114.

Annex 1: Methods and necessary modifications

	Category	Indicators	Methods	Necessary modifications
Process	Nature of the participatory research process	Participative events (number, type, phase, sequence)	- Output reporting in RDB (dialog-focused events) - Project story boards	
		· Stakeholders' motivation in participating	- Event evaluation	- New process
		· Adequacy of the role of stakeholders in events	*- process-focussed quality assurancene (methodology) (indirect, proxy)*	
		· Perceived importance of events	- Event evaluation	- New process
	Quality of the participatory research process	Proper representation of the range of opinions and perspectives	*- process-focussed quality assurancene (methodology) (indirect, proxy)*	
		· Fulfilment of the most important roles for participation	*- process-focussed quality assurancene (methodology) (indirect, proxy)*	
		· Sufficient level of interaction	*- process-focussed quality assurancene (methodology) (indirect, proxy)*	
		· Adequate appreciation and processing of stakeholder input	*- process-focussed quality assurancene (methodology) (indirect, proxy)*	
		Successful representation and resolution of dissent and conflict	*- process-focussed quality assurancene (methodology) (indirect, proxy)*	
		Diversity of participation activities	*- process-focussed quality assurancene (methodology) (indirect, proxy)*	
Prod. interaction	Direct productive interaction	- One-on-one conversations - Participation in advisory boards - Publications/talks for non-scientific audiences - Events	- Effect stories category: "Direct productive interaction" - Output reporting in RDB (memberships, talks and panels)	- Effect Categories **modify** - **Add** impact stories template to effect in the RDB - these **attributes** must adopted and **relevant data added retrospectively** to facilitate the evaluation of talks and panels + memberships with respect to target groups
	Indirect productive interaction	· "Contextual response"	- Effect stories category: "Direct productive interaction" - Secondary assessment of media analysis (Reach) - possible contextual response analysis (Spaapen/Drooge 2011, *op. cit.* , p. 28), or context monitoring (Williams 2017, *op. cit.* , p. 16) at project/policy area level	- Effect Categories **modify** Requires an **addition** to the RDB effect category for the entry of these results. (supplements)
	Financially mediated productive interaction	Contracts, licenses, level/source of project funding	- Output reporting in RDB (memberships, talks and panels) - Effect stories category: "Financial productive interaction"	- Effect Categories **modify**
First order effects	Utilisable scientific insights (Technologies & societal innovations, products/services, instrumental knowledge, publications)	Scientific advice (based on own research and/or evaluation of currently available knowledge; ability to conduct advisory activities independently; transparent procedures, relevance of advisory activities to target groups)	- Output reporting in RDB (Number of consultations, committee memberships) - possibly effect stories category: "Advice" *- Process-focussed quality assurance (consultancy concept/strategy) (indirect, proxy)* *- Articles of association, GO, Guidelines for good policy advice (safeguarding inst. independence) (indirect, proxy)*	- Effect Categories **modify** - in order to show that policy advice is grounded in research, the field "corresponding research" should be **added** to the RDB category "Advice" and **relevant data collected retrospectively**. - the current distinction in the RDB between "Effect" and "Advice" may need to be abandoned and data relevant to "Advice" (even if this is not analytically correct) should collected by means of a structured effect story
		Publications for non-scientific audiences	- Output reporting in RDB (non-scientific publication formats (brochures, magazine articles, blogs, Twitter, homepages))	- these **attributes** must adopted and **relevant data added retrospectively** to facilitate the evaluation of publications with respect to target groups and the further development of the publication strategy
		Events for non-scientific audiences	- Output reporting in RDB (Events for non-scientific audiences)	
		Products/materials for non-scientific audiences	- Output reporting in RDB	
	Improved capacities (Increased cognitive and anticipative capacity in relevant non-scientific partners)	improved understanding among non-scientific partners	- Surveys in selected cases, targeted queries/interviews at project level - Effect stories category: "Capacities - Understanding" *- Output reporting in RDB (Number non-scientific participants at events (dialogue), (indirect/proxy)* *- Reporting P&C (download statistics for publications)(indirect/proxy)*	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of these survey results. (supplements) - Effect stories template **must be added** to the effect category in the RDB
		new knowledge among relevant non-scientific partners	- Surveys in selected cases, targeted queries/interviews at project level - Effect stories category: "Capacities - Knowledge" *- Output reporting in RDB (Number non-scientific participants at events (dialogue), (indirect/proxy)* *- Reporting P&C (download statistics for publications)(indirect/proxy)*	- Requires an **addition** to the RDB effect category for the entry of these survey results. (supplements) - Effect stories template **must be added** to the effect category in the RDB
		improved communication capacities	- Surveys in selected cases, targeted queries/interviews at project level - Effect stories category: "Capacities - Knowledge" *- Output reporting in RDB (Number non-scientific participants at events (dialogue), (indirect/proxy)* *- Reporting P&C (download statistics for publications)(indirect/proxy)*	- Requires an **addition** to the RDB effect category for the entry of these survey results. (supplements) - Effect stories template **must be added** to the effect category in the RDB
	Network effects (improved networks and communities)	Number of networks created or extended	- Output reporting in RDB (number of research activities in cooperation with NGOs, public agencies, cultural institutions, businesses) '- Effect stories category: "Networks - Expansion"	- Effect Categories **modify** - **Add** effect stories template to the effect category in the RDB
		Cross-system collaborations and partnerships	- Output reporting in RDB (number of research activities in cooperation with NGOs, public agencies, cultural institutions, businesses) '- Effect stories category: "Networks - Cross-system Collaboration"	- Effect Categories **modify** - **Add** effect stories template to the effect category in the RDB

Annex 1: Methods and necessary modifications

		· New Contacts	- Output reporting in RDB (number of research activities in cooperation with NGOs, public agencies, cultural institutions, businesses) '- Effect stories category: "Networks - Contacts"	- Effect Categories **modify** - **Add** effect stories template to the effect category in the RDB
		Enhanced trust within networks	- Surveys in selected cases, targeted queries/interviews at project level '- Effect stories category: "Networks - Trust"	- Effect Categories **modify** - **Add** effect stories template to the effect category in the RDB
		Strengthened identity within communities	- Surveys in selected cases, targeted queries/interviews at project level '- Effect stories category: "Networks - Identity"	- Effect Categories **modify** - **Add** effect stories template to the effect category in the RDB
		Willingness to share with other networks	- Surveys in selected cases, targeted queries/interviews at project level '- Effect stories category: "Networks - Sharing"	- Effect Categories **modify** - **Add** effect stories template to the effect category in the RDB
	New discourses and narratives (New/changed public discourses, new/changed public narratives)	Number /quality of new narrative elements in media discourse	- Bellwether interviews in selected cases - Context analyses in selected cases - Interviews with process participants in selected cases - Effect stories category: "Discourses - Narrative elements"	- Effect Categories **modify** - **Add** effect stories template to the effect category in the RDB
		Quality of changes to discourse	- Bellwether interviews in selected cases - Context analyses in selected cases - Interviews with process participants in selected cases - Effect stories category: "Discourses - Discursive shifts"	- Effect Categories **modify** - **Add** effect stories template to the effect category in the RDB
Second order effects	Policy effects (knowledge-based policies, knowledge-based decisions, implemented solutions, instrumental or conceptual use of knowledge, economic benefits)	Number /quality of knowledge-based policies & laws adopted	- Bellwether interviews in selected cases - Interviews with process participants in selected cases - Effect stories category: "Policy effect - Laws/Regulation" - in-depth analyses of selected cases using a suitable method	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of these interview results. (supplements) - Effect stories template **must be added** to the effect category in the RDB
		Number of knowledge-based decisions	- Bellwether interviews in selected cases - Interviews with process participants in selected cases - Effect stories category: "Policy effects - Decisions" - in-depth analyses of selected cases using a suitable method	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of these interview results. (supplements) - Effect stories template must be added to the effect category in the RDB
		Change in policy-related discourse	- Bellwether interviews in selected cases - Interviews with process participants in selected cases - Impact stories category: "Policy effects - discourse" - in-depth analyses of selected cases using a suitable method (outcome harvesting, process tracing, most significant change, contribution analysis)	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of these interview results and analyses of individual cases. (supplements) - Effect stories template **must be added** to the effect category in the RDB
	Organizational change (Changed organizational context for current and future work, new organizations, new standards, new partnerships, new business models)	Changes in responsibilities and roles	- Bellwether interviews - Interviews with process participants - Impact stories category: "Organizational change - Roles/Responsibilities" - in-depth analyses of selected cases using a suitable method (outcome harvesting, process tracing, most significant change, contribution analysis)	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of these interview results and analyses of individual cases. (supplements) - Effect stories template must be added to the effect category in the RDB
		Changes in the rules of participation / organization of processes	- Bellwether interviews in selected cases - Interviews with process participants in selected cases - Impact stories category: "Organizational change - Participation rules/processes" - in-depth analyses of selected cases using a suitable method (outcome harvesting, process tracing, most significant change, contribution analysis)	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of these interview results and analyses of individual cases. (supplements) - Effect stories template **must be added** to the effect category in the RDB
		New institutional framework	- Bellwether interviews in selected cases - Interviews with process participants in selected cases - Impact stories category: "Organizational change - Institutional framework" - in-depth analyses of selected cases using a suitable method (outcome harvesting, process tracing, most significant change, contribution analysis)	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of these interview results and analyses of individual cases. (supplements) - Effect stories template **must be added** to the effect category in the RDB
		Changes in investment strategies	- Bellwether interviews in selected cases - Interviews with process participants in selected cases - Impact stories category: "Organizational change - Investment strategy" - in-depth analyses of selected cases using a suitable method (outcome harvesting, process tracing, most significant change, contribution analysis)	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of these interview results and analyses of individual cases. (supplements) - Effect stories template **must be added** to the effect category in the RDB

Annex 1: Methods and necessary modifications

Third order effects	Alternative Visions and Social Imaginaries (changes in collective goals and visions, greater social cohesion across groups and beyond)	Quality/resilience of new social imaginaries	- Bellwether interviews in selected cases - Context analyses in selected cases - Interviews with process participants in selected cases - Impact stories category: "Imaginaries - Discourses" - in-depth analyses of selected cases using a suitable method (outcome harvesting, process tracing, most significant change, contribution analysis)	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of interview results, context analysis, and analyses of individual cases. (supplements) - Effect stories template **must be added** to the effect category in the RDB
		Quality/resilience of new visions	- Bellwether interviews in selected cases - Context analyses in selected cases - Interviews with process participants in selected cases - Impact stories category: "Imaginaries - Discourses" - in-depth analyses of selected cases using a suitable method (outcome harvesting, process tracing, most significant change, contribution analysis)	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of interview results, context analysis, and analyses of individual cases. (supplements) - Effect stories template **must be added** to the effect category in the RDB
	Transformed social practice (Changes in the socio-cultural landscape, adjustment of or change in standards, inclusion of new actors or themes in public spaces and discourses, changes in practices of participation, new spaces for innovation and experimentation)	Changes in the behaviour of collective actors	- Bellwether interviews in selected cases - Interviews with process participants in selected cases - Impact stories category: "Social practice - Behavior" - in-depth analyses of selected cases using a suitable method	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of these interview results and analyses of individual cases. (supplements) - Effect stories template **must be added** to the effect category in the RDB
		Increased capacity for collective action	- Bellwether interviews in selected cases - Interviews with process participants in selected cases - Impact stories category: "Social practice - Capacity" - in-depth analyses of selected cases using a suitable method	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of these interview results and analyses of individual cases. (supplements) - Effect stories template **must be added** to the effect category in the RDB
		New forms of participation	- Bellwether interviews in selected cases - Interviews with process participants in selected cases - Impact stories category: "Social practice - Behavior" - in-depth analyses of selected cases using a suitable method	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of these interview results and analyses of individual cases. (supplements) - Effect stories template **must be added** to the effect category in the RDB
		Deeper networks and solidarity	- Bellwether interviews - Interviews with process participants - Impact stories category: "Social practice - Behavior" - in-depth analyses of selected cases using a suitable method	- Effect Categories **modify** - Requires an **addition** to the RDB effect category for the entry of these interview results and analyses of individual cases. (supplements) - Effect stories template **must be added** to the effect category in the RDB

Annex 2: Template for effect stories

Effect stories are used to collect data for the purpose of ex-post quality assurance. Here, they will be used to collect data relating to individual indicators/evidence as part of the quality model for the institution as a whole and will inform the selection of individual cases for further analysis.

The stories should not exceed 4 pages in length and should refer to one or more effect categories in order to ensure the relevance of their contents to the evaluation framework. The stories should be linked to at least one research output in order to demonstrate the research-basis of effects and thus meet a central requirement of the Science Council. In addition to this, (objective) evidence is to be provided for the effects described (REF 2011, REF 2012).

Title of the effect story:
Effect category: [Selection from Vocabulary 1. Productive interaction: a) Direct, b) Indirect, c) Financial), 2. Capacities: a) Understanding, b) Knowledge, c) Communication, 3. Networks: a) Expansion, b) Cross-system collaboration, c) Contacts, d) Trust, e) Identity, f) Sharing, 4. New discourses/Narratives: a) Narrative element, b) Qualitative change in discourse, 5. Policy effect: a) Law/Regulation, b) Decision, c) Discourse, 6. Organizational change: a) Roles/Responsibilities, b) Process/Rules of participation, c) Institutional framework, d) Investment strategy, 7. Alternative visions and imaginaries: a) Imaginaries, b) Visions, 8. Social practice: a) Behavior, b) Capacity. (multiple capacities may be identified]
Corresponding project: [Link to project – multiple projects may be identified]
1. Summary of effect (max. 100 words):
2. Relevant research results (max. 200 words):
3. Reference to research (max. 6 references): [Selection of publications listed in RDB]
4. Detailed description of effect (max. 750 words): This section should offer a narrative that covers the following: - How do the research results relate to the effect? In what way do they make a vital and material contribution to the effect? - What is the nature and extent of the effect? The following details should be included here: - A clear explanation of the process and/or means by which the research led to, underpinned or supported the effect (e.g. how the results were disseminated and how this influenced some users in their actions or attitudes). - The contributing research should be placed within a larger research context (incl. mention of other key results in the field) (e.g. cooperative activities) and its specific contribution to the effect noted in each case. - Beneficiaries - who or which group or organization was affected positively?

- Descriptions of the effect (including reach and significance) on beneficiaries - how have they benefited, or been otherwise influenced or affected?
- Data on when effects impacted.

5. Evidence of effect (max. 10 references)

Evidence for the effects described above (incl. their occurrence, significance, and reach) should be detailed in this section.

Depending on the specific instance, this may include:

- reports, reviews, web links or other documented sources accessible in the public sphere,
- confidential reports or documents to which access can be provided on request (under the condition that confidentiality be maintained),
- naming of individual beneficiaries who are willing to provide testimonies
- factual statements by beneficiaries[1]

REF (2011). REF 2014: Assessment framework and guidance on submissions.
REF (2012). REF 2014: Panel criteria and working methods.

[1] Unlike an "individual testimonial", a "factual statement" should provide only verifiable details that do not relate to specific persons. Cf. REF 2012: