palgrave communications
HUMANITIES | SOCIAL SCIENCES | BUSINESS

ARTICLE

Received 23 May 2016 | Accepted 15 Nov 2016 | Published 13 Dec 2016

DOI: 10.1057/palcomms.2016.92 OPEN

# Scientific assessments to facilitate deliberative policy learning

Martin Kowarsch[1], Jennifer Garard[1,2,5], Pauline Riousset[3,5], Dominic Lenzi[1,5], Marcel J. Dorsch[1,2,5], Brigitte Knopf[1,5], Jan-Albrecht Harrs[1,5] and Ottmar Edenhofer[1,2,4,6]

**ABSTRACT** Putting the recently adopted global Sustainable Development Goals or the Paris Agreement on international climate policy into action will require careful policy choices. Appropriately informing decision-makers about longer-term, wicked policy issues remains a considerable challenge for the scientific community. Typically, these vital policy issues are highly uncertain, value-laden and disputed, and affect multiple temporal and spatial scales, governance levels, policy fields, and socioeconomic contexts simultaneously. In light of this, science-policy interfaces should help facilitate learning processes and open deliberation among all actors involved about potentially acceptable policy pathways. For this purpose, science-policy interfaces must strive to foster some enabling conditions: (1) "representation" in terms of engaging with diverse stakeholders (including experts) and acknowledging divergent viewpoints; (2) "empowerment" of underrepresented societal groups by co-developing and integrating policy scenarios that reflect their specific knowledge systems and worldviews; (3) "capacity building" regarding methods and skills for integration and synthesis, as well as through the provision of knowledge synthesis about the policy solution space; and (4) "spaces for deliberation", facilitating direct interaction between different stakeholders, including governments and scientists. We argue that integrated, multi-stakeholder, scientific assessment processes—particularly the collaborative assessments of policy alternatives and their various implications—offer potential advantages in this regard, compared with alternatives for bridging scientific expertise and public policy. This article is part of a collection on scientific advice to governments.

[1] Mercator Research Institute on Global Commons and Climate Change, Berlin, Germany [2] Technical University Berlin, Berlin, Germany [3] Institute for Ecological Economy Research, Berlin, Germany [4] Potsdam Institute for Climate Impact Research, Potsdam, Germany Correspondence: (email: kowarsch@mcc-berlin.net)

[5]These authors contributed equally to this work.

[6]These authors jointly supervised this work.

## Introduction: science-policy interfaces for longer-term, wicked policy problems

The global environmental challenges of the twenty-first century affect the fundamental interests of billions of people and other living beings, both now and into the future. International climate policy in light of the Paris Agreement, as well as the recently adopted global Sustainable Development Goals (SDGs), are clear examples of policy goals and processes focused on longer-term, "wicked", yet vital policy issues. Wicked policy problems are mainly characterized by complexity[1] and related interdependencies, high uncertainty, divergence of viewpoints and values, and fluid problem definition.[2] This is also true for other grand challenges beyond the environmental realm, such as global health, demographic change and social welfare, and economic and financial crises. (ICSU, 2010).

Wicked policy problems are complex because they typically affect and involve multiple temporal and spatial scales, governance levels, policy fields, and socioeconomic contexts simultaneously. This is well illustrated by recent work on the relationship between SDGs and climate policy goals (e.g., Jakob and Steckel, 2016; von Stechow *et al.*, 2016). Sustainable development seems to be among the wickedest current policy problems (Haas, 2004). Given the many potential synergies and tradeoffs involved, carefully considered policy choices are required. It remains unclear what the most appropriate policy options are in these cases, partly due to a lack of knowledge about the many direct and indirect effects of policy options; conventional methods of scientific inquiry usually do not suffice to address wicked issues effectively (Funtowicz and Ravetz, 1991; Turnpenny *et al.*, 2009). Three examples illustrate the importance of revealing and addressing indirect effects of, for instance, climate policies:

- How to ensure climate protection without undermining the right to development, given that economic development has been linked to the exploitation of fossil fuels in the past?
- How to use bioenergy for climate change mitigation without negatively affecting biodiversity, forest protection, water availability and food security?
- How to tackle the distributional questions related to policy instruments such as carbon pricing for climate change mitigation?

There is thus an obvious need for scientific expertise informing policy on these (and many other) wicked policy issues to better understand, among other things, the available policy options and their practical implications.

However, what are the best types of bridges between various forms of scientific expertise and public policy processes—in short: science-policy interfaces (SPIs)—to effectively and appropriately inform decision-makers regarding longer-term, wicked policy problems? This is the central question underlying this paper. The longer-term, wicked policy issues *amplify and specify* some of the old profound challenges for SPIs (Pielke, 2007; Hulme, 2009; Aitsi-Selmi *et al.*., 2016; Kowarsch, 2016a; von Stechow *et al.*., 2016). First, the complexity of these policy issues makes it difficult for scientific studies to address the high number of policy-relevant aspects of the issues at stake in a truly integrated manner, across disciplines and approaches. Second, another challenge is high uncertainty resulting from the need to go beyond traditional areas of research to address large-scale, long-term and non-linear risks. However, increasing bodies of literature do not necessarily reduce uncertainty and disagreement, particularly in the social sciences where research is often not aggregated (Hunter and Schmidt, 1996; van Slyke *et al.*., 2010). This also endangers scientific credibility. Third, the contestable yet unavoidable normative assumptions involved—such as those related to prioritizing policy goals or means, policy evaluation criteria or any evaluation of uncertainty (Putnam, 2004; Douglas, 2009; Dietz, 2013)—raise questions of legitimacy, particularly if these assumptions are related to power asymmetries, or the large number actors with many diverging worldviews and interests. Furthermore, related to these three challenges, there are also tradeoffs between salience, credibility and legitimacy (Cash *et al.*, 2003; Mitchell *et al.*, 2006), particularly when it comes to the highly value-laden and often uncertain social-science evaluation of controversial policy response options. If these challenges and tradeoffs for SPIs on longer-term, wicked policy problems are not appropriately addressed, these SPIs will be largely ineffective because their effectiveness presupposes salience, credibility and legitimacy from the perspective of various actors involved at SPIs (Cash *et al.*, 2003; Sarewitz, 2004; Pielke, 2007; Hulme, 2009; Kowarsch, 2016b, Chapter. 3).

Large-scale, integrated scientific assessment processes are examples of SPIs that are often used to address the longer-term, wicked policy problems. There is currently a proliferation of, for instance, global environmental assessments. More than 130 of them have been initiated over the past four decades, mostly at the behest of governmental bodies. Examples include the United Nations Environment Programme's (UNEP) Global Environment Outlook (GEO) series and the assessment reports provided by the Intergovernmental Panel on Climate Change (IPCC). However, there are many calls for far-reaching reform of scientific assessment processes, and some lament that they are overly laborious, time-consuming, and institutionally or politically constrained. Yet, while agreeing with the need for continuous assessment reform, are the available alternatives—such as standing expert committees, or scientific reports produced without multi-stakeholder processes—better than large-scale assessments for addressing longer-term, wicked policy problems? We will argue that assessments, particularly when following the Pragmatic-Enlightened Model (Edenhofer and Kowarsch, 2015), are comparatively promising tools in this regard, because these assessment processes have higher potential for facilitating deliberative policy learning among all actors involved.

There is a large literature on assessments, their characteristics, failures and reform options (e.g., Cash *et al.*, 2003; Mitchell *et al.*, 2006; Norgaard, 2008a; Carraro *et al.*, 2015; Victor, 2015; Aitsi-Selmi *et al.*, 2016). There are also many case studies (mostly on national or sub-national scale) on various other SPIs (see literature provided in the following section), as well as normative models for scientific policy advice in general (e.g., Habermas, 1971; Pielke, 2007; Brown, 2009; Kowarsch, 2016b). Typically, these studies move beyond the still predominant technocratic model for scientific policy advice. Instead, ideals and goals of deliberative democracy, and the potential role of scientific expertise for policy learning and policy change (e.g., Hajer, 1993; Schmidt and Radaelli, 2004; Sabatier, 2007) are widely discussed and acknowledged in the literature. However, these concepts must be better translated into criteria (i.e. metrics of success) for SPIs which should then be applied to evaluate the ability of different SPIs to respond to longer-term, wicked policy problems.

Based on a brief characterization of selected SPIs and some criteria for SPIs mainly distilled from major "building blocks" of deliberative policy learning, we discuss the extent to which assessments, as SPIs, measure up against alternative SPIs on longer-term, wicked policy problems—in terms of their theoretical relative potential to realize these building blocks. We then empirically analyse the extent to which assessments actually contribute, as overall outcome, to deliberative policy learning.

### Integrated scientific assessments, and other science-policy interfaces

Let us briefly introduce the core characteristics of some selected SPIs. We mainly focus on large-scale, integrated scientific assessment processes here, because they turned out to be the most promising SPIs for longer-term, wicked policy problems (see further below).

The international community has mandated and supported a number of prominent, large-scale assessments of environmental issues in recent years (see IPBES, 2013 for an overview). Conducting large-scale assessments requires hundreds of researchers from different disciplines, experts from non-academic institutions, and several years of collaborative knowledge synthesis and significant financial resources. At present, a diversity of projects at both global and sub-global scales can be called "scientific assessments." However, there is some confusion about what this term means. Below, we propose a conceptualization of contemporary integrated scientific assessments in the context of public policy-making processes. However, we acknowledge that there is considerable diversity among assessments, which provide the different types of scientific outputs, and whose characteristics can change over time. Nonetheless, three central characteristics of integrated scientific assessments are:

- *Assembling the available scientific knowledge* (and identifying research gaps) in order to provide a rich, interdisciplinary and highly integrated image of the policy-relevant considerations. Additionally, and to a greater extent than in literature reviews, peer-reviewed *synthesis* of the available publications and information is required to identify the confidence level that can be associated with the scientific findings in assessments, and to put the available scientific knowledge into decision-making contexts by pointing out the potential implications for policy debates. Synthesis necessarily involves "assessment" itself and informed judgment, as well as a high level of integration and coherence.
- Striving to *provide policy-relevant scientific knowledge in a publicly accessible manner* to support public policy-making processes and deliberation. This means formulating scientific insights that may (1) help frame and define the societal problem at stake, including the policy goals and objectives, (2) shed more light on available policy means (such as policy instruments, institutions, measures), and/or (3) reveal potential or actual (ex post or ex ante) implications of these means in terms of direct effects, adverse side effects (costs, risks, etc.) and synergies (co-benefits).
- Taking into account *different viewpoints* in terms of controversial scientific statements and approaches, uncertainty, and disputed societal values and conflicting interests. Besides making areas of disagreement transparent in the assessment outputs, (1) engaging with policy-makers and other stakeholders, as well as (2) involving a number of authors with various backgrounds, approaches and viewpoints, are likely assessment design elements to realize this. As such, assessments can be regarded as formal social processes to scientifically discuss policy-relevant issues, which usually facilitate learning among the participants. Assessments usually are not advocacy pieces.

To summarize, integrated scientific assessments are multi-stakeholder processes for distilling and synthesizing knowledge in particular fields to inform policy, involving (regionally and intellectually) diverse experts and stakeholders (see Mitchell *et al.*, 2006, 3, for a more comprehensive definition). Depending on the degree to which these characteristics are realized, one can distinguish between smaller-scale and larger-scale assessments. Many assessments are formally mandated by policy-makers, which indicates demand and may facilitate their impact. An intermediate scientific activity between standard research and assessments is doing pre-assessments (or "pilot assessments"), for instance meta-studies that aggregate knowledge to fill research gaps identified in previous assessment processes.

The IPCC assessments will serve as one of our key examples of assessments.[3] The IPCC was created in 1988 by several international organizations as an intergovernmental panel for knowledge synthesis to inform climate policy, involving both governments and scientists. The IPCC has become the leading international body for assessing and synthesizing knowledge in climate change and its potential environmental and socio-economic impacts. Its core products are its lengthy periodical Assessment Reports mainly produced by a decentralized, worldwide network of thousands of scholars and experts in order to assess the current state of scientific knowledge in a scientifically sound and policy-relevant but not policy-prescriptive manner. The IPCC's scientific rigor, hybrid structure and impact at the science-policy interface inspired a number of other assessment processes following the IPCC model.

Having described some general common characteristics of the diverse existing integrated scientific assessment processes as SPIs, what should be more specific features of an assessment from a *normative* perspective? The Pragmatic-Enlightened Model (PEM), developed by Edenhofer and Kowarsch (2015) based on John Dewey's pragmatist philosophy of ends-means-interdependency, provides guidance on assessment design, at least for large-scale assessments of wicked policy issues that face high uncertainty and disputed value-laden viewpoints. For policy evaluations, the PEM assumes the interdependency of policy objectives, means and their implications. For example, the extensive use of bioenergy for ambitious climate change mitigation can negatively affect biodiversity and food security which may require a revaluation of the means and even the initial policy goals. Key claims of the PEM thus are (1) to thoroughly explore the various practical implications of policy means in quantitative and qualitative terms, making uncertainty transparent; (2) to explore and present alternative, disputed policy pathways in the assessment, related to different policy objectives and values; and (3) to engage diverse stakeholders at different stages of the assessment process, enabling the co-production of reliable knowledge based on scientific methods. The PEM envisages the role of scientific experts as mapmakers of alternative policy pathways and their implications, while policy-makers bear the role of navigators. In this way, assessments (though not value-free) may avoid policy-prescription, while still allowing for learning about policy pathways. Although assessment always implies value judgments and uncertainty, reliable and objective scientific knowledge in assessments remains possible and desirable (Kowarsch, 2016b). Assessments should strictly be based on rigorous (and wherever possible, peer-reviewed) scientific research.

Based on that, this paper will discuss two different types of integrated scientific assessments: first, conventional integrated scientific assessments; and second, medium-scale and large-scale integrated scientific assessments that are more or less conducted in the "spirit" of the ideals claimed by the PEM model. A prominent example of such PEM-inspired assessments, as we call them in this paper, is the recent contribution of the IPCC Working Group III on climate change mitigation options to the IPCC's Fifth Assessment Report (IPCC, 2014a; Edenhofer and Kowarsch, 2015). However, for the purpose of this article, the group of "PEM-inspired assessments" also include current or past —but also potential future—assessment processes that do *not explicitly* follow the recently developed PEM model, but at least

are more or less in line with some core claims of the PEM, i.e. the collaborative exploration of policy alternatives. We thus use the term "PEM-inspired assessments" rather broadly in this paper.

Predominant alternative SPIs—which are used in the subsequent sections to illustrate the relatively high potential of integrated scientific assessments, particularly the PEM-inspired ones—include:

1. Conventional peer-reviewed research papers and studies addressing policy-relevant issues, including in particular meta-studies (i.e., meta-analysis, model intercomparisons, etc.) and pre-assessments;
2. Topic-specific, individual or small-group advice and consultancy (relatively homogeneous intellectually, politically and geographically; including, e.g., think tanks, lobby groups, foundations, chief scientific advisors, learned societies, and national research organizations), often one-off; mainly through the provision of scientific reports to inform policy based on self-determined processes (but also through direct dialogue with decision-makers);
3. Permanent expert committees (or councils, or panels, or other formalized structures) with formalized procedures which provide direct (e.g., oral) advice to individual policy-makers or parliaments, written policy briefs, or smaller reports, but which do not provide more comprehensive, systematic assessments;
4. Standardized impact assessment reports which are usually more national or sub-national;

The core characteristics of these selected alternative SPIs are summarized in Tables 1 and 2, together with major characteristics of conventional and PEM-inspired integrated scientific assessments.

These six key SPIs are by no means comprehensive. As described by Kohler *et al.* (2012), a huge variety of SPIs have been developed over the past 25 years as the need for deliberate scientific input to decision-making has been increasingly recognized. This paper explores a selection of key interfaces employed regularly, which have been chosen to reflect the variety of options, scales and actors involved. Within the six SPIs listed in Tables 1 and 2, there is a high degree of variability. The different SPIs were grouped in order to provide some stylized clustering of the different interfaces which are not perfectly distinct. There is sometimes a high degree of variability of the performance of a particular format for scientific policy advice with regards to different characteristics.

Many of these interfaces are interlinked; for example, large-scale assessments often build on conventional research papers, pre-assessments, smaller-scale assessments as well as the work done at other SPIs. Conversely, for instance, research gaps identified in large-scale assessments can often inspire new directions for future research papers, pre-assessments and other initiatives, and longer-term policy issues can also involve the need for short-term scientific policy advice, etc.

## Deliberative democratic theory provides criteria for science-policy interfaces

We ground our discussion of the selected SPIs in a popular branch of political theory—namely deliberative democratic theory. While the reason for this might not be obvious at a first glance, it is compelling upon further reflection. A central claim of deliberative democratic theory (see also Box 1) is "that legitimacy requires the right, opportunity, and capacity of those subject to a collective decision to participate in consequential deliberation about the decision" (Stevenson and Dryzek, 2012: 2). Unlike traditional theories of representative democracy that primarily link legitimacy and public consent solely to voting-mechanisms or policy output, deliberative democratic theorists claim that political decisions must be based on and linked to inclusive and deliberative public debates—especially of those subject to these decisions. In these deliberation processes, different arguments, viewpoints and interests on a specific matter are brought up and are exchanged to allow for mutual learning and reasoning among all actors involved. Any decision made subsequently has to be justified on the basis of this deliberation—it has to be made plausible in light of the arguments heard before (Habermas, 1996b). This deliberation process can shape individual preferences by arguments and reasoning, going beyond blatant self-interest. Legitimacy is thus conceptually linked with accountability, involving public reason-giving discourses aimed at both informing and justifying public policy (Chambers, 2003: 308f). Due to their rigorous methodologies and systematic analyses of various policy aspects, the sciences as one societal subsystem can and should provide knowledge relevant for policy deliberation.

Deliberative democratic theory is an appropriate normative stance especially for discussing alternative SPIs responding to longer-term, wicked policy problems. This is due to the particular characteristics of these wicked issues described in the introduction, in light of which the strengths of deliberative democratic theory as a normative grounding for SPIs become particularly apparent. Given (1) the considerable multi-dimensionality and complexity, (2) deep uncertainty, and (3) divergent normative viewpoints and stakes involved in such policy issues (both regarding potential solutions and the underlying problem definition), no straightforward (scientific) method for identifying appropriate policy pathways exists. Some experts might advocate strong opinions on particular policy issues, but doing so relies upon contestable normative assumptions. Rather, for effective

**Deliberative democratic theory**

Deliberative democracy is properly described as a family of theories, including Dryzek's "discursive" theory (Dryzek, 1990, 2010), Pettit's "republican-contestation" theory (Pettit, 2000), Chambers' "reasonable" democratic theory (Chambers, 1996), Young's "communicative" theory (Young, 1993), and, perhaps most famously, Habermas' "deliberative politics" (Habermas, 1996a, 1996b,), and earlier works such as Dewey (1927). Although these theories differ in important respects, Wiklund has pointed out that the concept of a participatory "voice" is common to all of them (Wiklund, 2005: 283). Dryzek helpfully defines the key concept of deliberation itself as an actualization of this voice within pluralistic public dialogue, which hopefully "induces reflection upon preferences in a non-coercive fashion" (Dryzek, 2000: 2). Chambers adds that deliberation is "debate and discussion aimed at producing reasonable, well-informed opinions in which participants are willing to revise preferences in light of discussion, new information, and claims made by fellow participants" (Chambers, 2003: 309). The general process of deliberation itself can take many forms. It can be very open and unbound (take the Internet as an example of global scale) but also very narrow and bounded (think about the members of a local floriculture club debating a new statute).

Deliberative processes can also be convincing for those who are otherwise unaffiliated with deliberative democratic theory (Chambers, 2003: 308). In a report for the US National Academy of Sciences, Dietz and Stern conclude that "substantial evidence shows that effective public participation can help agencies do a better job in achieving public purposes for the environment by ensuring better decisions and increasing the likelihood that they will be implemented effectively" (Dietz and Stern, 2008: 226). Well-designed public participation not only "improves the quality and legitimacy of a decision and builds the capacity of all involved to engage in the policy process", but can also "lead to better results in terms of environmental quality and other social objectives" (Dietz and Stern, 2008: 226).

**Table 1 | Organizational and procedural setup of a selected few science-policy interfaces**

| | *Research studies (including meta-studies, pre-assessments and so on)* | *Individual or small groups of scientific advisors or consultants* | *Permanent scientific advisory bodies (not providing larger-scale assessments)* | *Standardized impact assessments* | *Conventional medium-scale to large-scale integrated scientific assessments* | *Integrated scientific assessments in the spirit of the PEM* |
|---|---|---|---|---|---|---|
| *Examples* | Transport and ETS (See https://ideas.repec.org/p/ecc/wpaper/2.html, accessed 15 May 2016); ADAM (See https://www.pik-potsdam.de/research/sustainable-solutions/flagshipspld/MitigationScenarios/adam/adam-project, accessed 15 May 2016); papers commissioned by, for example, WHO, CBD and so on (See, e.g., https://www.gov.uk/government/collections/reducing-risk-of-future-disasters#supporting-evidence, http://www.who.int/hia/health_indicators/en/, and http://www.cbd.int/doc/health/guide-biodiversity-health-en.pdf, accessed 15 May 2016)). | Reports by science academies, think tanks, etc. (see examples & discussions in OECD, 2015; McGann *et al.*, 2014; Mentzel, 1999); Chief Scientific Advisor (CSA) | To national governments (that is, US SAB and multiple sub-committees), India SAC-PM and SAC-C) or to International Conventions (for example, UNFCCC SBSTA, UNESCO SAB, CBD SBSTTA; EU SAM) | EU Impact Assessment (See http://ec.europa.eu/smart-regulation/impact/index_en.htm, accessed 15 May 2016) | Stern Review; UK Foresight Reports ( See https://www.gov.uk/government/collections/foresight-projects, accessed 15 May 2016); UN Emissions Gap Report (See, for example, http://www.unep.org/publications/ebooks/emissionsgapreport/, accessed 15 May 2016); HDR (See, for example, http://hdr.undp.org/sites/default/files/hdr14-report-en-1.pdf, accessed 15 May 2016); UNEP's GEO series; GBO; IAASTD. See IPBES (2013) for details and acronyms | IPCC WG III AR5; MA; envisioned IPBES assessments (see IPBES, 2013 for details and acronyms) |
| **Organizational and procedural setup** | | | | | | |
| *Number of experts and disciplines* | Few (sometimes interdisciplinary) | 1 to a few (sometimes >10 authors, sometimes inter-disciplinary) | Many (often >10, often inter-disciplinary) | Often >15 authors, often inter-disciplinary | Very high number (usually >100 authors, mostly interdisciplinary) | Very high number (usually >100 authors, always inter-disciplinary) |
| *Author selection process* | Self-selected; rarely nominated, e.g. if papers have been commissioned | Self-selected, or sometimes nominated, e.g. CSA, or if reports have been commissioned | Often through formalized, criteria-based nomination process envisaging balanced representation | Usually nominated | Formalized process, often many divergent perspectives, and regional representation | Highly formalized process, many divergent perspectives, and regional representation |
| *Duration of process* | Between a few months and 1–2 years | Often <1 year (sometimes more) | Often for a set term, for example, 4 years | Often <2 years | Usually several years | Often 4–7 years |
| *Governance structure (incl. formalization of processes)* | informal | Mostly informal | Often formalized, for example, under guidance of a chairperson | Usually, parts of the process are formalized | Highly formalized re. committees, procedures, conflict of interest policies etc. | Highly formalized re. committees, procedures, conflict of interest policies and so on. |

**Table 2 | Focus and methodology as well as relationship to stakeholders and political context of a selected few science-policy interfaces**

| | *Research studies (including meta-studies, pre-assessments, etc.)* | *Individual or small groups of scientific advisors or consultants* | *Permanent scientific advisory bodies (not providing larger-scale assessments)* | *Standardized impact assessments* | *Conventional medium-scale to large-scale integrated scientific assessments* | *Integrated scientific assessments in the spirit of the PEM* |
|---|---|---|---|---|---|---|
| **Focus and methodology** | | | | | | |
| *Scope of literature* | Various, depending on type of study | Most relevant literature; sometimes synthesis of additional literature. | Most relevant literature; sometimes synthesis of additional literature. | Synthesis of most relevant literature | Relatively comprehensive synthesis of literature | Relatively comprehensive synthesis of literature |
| *Policy options / alternatives explored* | One or a few key options (if policy analysis is done at all) | One or a few key options (policy analysis is often done) | High variation, sometimes several key options explored | Exploring a few (pre-determined) options | Exploring several options (if policy analysis is done at all) | Many alternative pathways, options, and diverse practical implications |
| *Governance levels & other complexities considered* | rather few | High variation, often only a few; sometimes analysis of interdependencies | High variation, often a few, sometimes with analysis of interdependencies | Only a few, sometimes with analysis of interdependencies | High variation, often multiple, and with analysis of interdependencies | Several, and with extensive analysis of interdependencies |
| *Transparency of levels of confidence and agreement* | Some | High variation, sometimes only little | High variation, sometimes high | Moderate to high | High variation; moderate to very high | Very high |
| *Review process* | Double-blind formalized peer review | High variation, often no external peer-review | Usually extensive internal review processes | Usually formalized internal review processes | High variation, often a larger group of external reviewers sometimes including policy makers and other stakeholders | Large-scale, formalized, multi-stage review process including policy makers and other stakeholders |
| **Relationship to stakeholders and political context** | | | | | | |
| *Mandate from governing bodies* | Very rarely | Rarely | Nearly always | Usually | Mostly | Mostly |
| *Interaction with stakeholders (beyond scientists)* | If at all, usually only little (and informally, i.e. bilateral conversations) | High variation, sometimes little (informally, or surveys/ interviews) | Usually formalized interaction with particular governmental bodies | Usually formalized interaction with particular governmental bodies | Very often, and increasingly observed through multiple formats and with many groups | Extensive, and through multiple formats; many diverse groups |
| *Inclusion of divergent viewpoints* | If at all, overview of major divergent viewpoints; sometimes a few scenarios | Sometimes overview of major divergent viewpoints; sometimes a few scenarios | Often the committee itself represents divergent views; sometimes scenario exploration | Exploration of a few scenarios related to pre-selected policy alternatives | Overview and exploration of several divergent viewpoints, sometimes through co-produced scenarios | Exploration of many relevant divergent viewpoints, mostly through co-produced scenarios |
| *Outreach and communication* | Little (beyond scientific community) | Often extensive efforts | Mostly extensive efforts | Moderate | Often extensive efforts | Mostly extensive efforts |

and legitimate decision-making on these wicked policy problems, an inclusive, open and integrated deliberation process about policy problems and the policy solution space is essential, involving diverse perspectives and actor groups as well as scientific disciplines. Deliberative democratic theory rejects technocratic approaches, according to which scientific experts and engineers alone can comprehensively determine the most appropriate policy options in an allegedly objective and reliable manner. Drawing clear boundaries between allegedly "neutral" scientific expertise on the one hand and value-laden policy-making on the other is impossible. Technocratic approaches are thus inappropriate here, and largely ineffective in terms of impact on policy processes, inter alia due to the implicit, entangled and often disputed value judgments and uncertainties (e.g. Jasanoff, 1990; Sarewitz, 2004; Pielke, 2007; Kowarsch, 2016b).

While aggregating and weighing individual preferences and viewpoints in one form or another towards any kind of social welfare is indispensable for collective policy-making and policy evaluation, Arrow showed the infeasibility of traditional economic approaches to preference aggregation (Arrow, 1970). The aggregation of citizens' individual interests through mere voting mechanism, polls, or revealed preferences is either not possible or not sufficient to determine social welfare and "the public good". However, well-designed deliberation processes within SPIs can provide a valuable approximation in this regard; they serve as valuable fora for exchanging and justifying competing preferences, values and beliefs. This could lead to learning processes where one's own preferences might also change over time, and it could facilitate an inclusive and transparent process of *deliberative* preference aggregation that can feed into political decision-making processes. Deliberative design of SPIs aims at convergence on policy solutions where possible—not as a result of individual bargaining power, but as a result of public justification, which means providing and judging on reasons. Hence, the purpose of deliberation also in SPIs is simultaneously "epistemic and practical", aiming "to uncover facts about interests and equality and how best to pursue them for the purpose of making good collective decisions" (Christiano, 2012: 27).

Additionally, deliberative and ongoing learning procedures generate further benefits for dealing with wicked policy problems. Actively involved agents in well-designed deliberation processes may better understand the complexity and uncertainty connected to policy decisions and have the opportunity to co-develop appropriate policy options. Deliberative processes within SPIs can thus increase the legitimacy and social acceptance of subsequent policy decisions, making them more compelling, resilient and sustainable than decisions based on mere bargaining between interest groups. Wicked problems like climate change affect people around the world in different regions and jurisdictions. They are increasingly demanding an input into policy debates on wicked problems according to their different vulnerabilities and policy preferences. Some even argue that democratizing the diverse forms of international science-policy discourse by means of deliberative practices may well prove to be a "shorter road" than democratizing the constitution of international institutions (Dryzek, 1999: 35).

To conclude, although successful deliberation itself is demanding (e.g. Ryfe, 2005), deliberative design of SPIs is indispensable especially in response to wicked problems where simple all-embracing and uncontroversial solution pathways do not exist (Dietz, 2013). Precisely because of the multi-faceted complexity and interdependencies of these problems, ongoing and transparent deliberation on available knowledge and response options is essential—closely echoing John Dewey's early call for open-ended deliberative public experimentation of different means and different goals, not least based on scientific methodology (Dewey, 1927).

However, newly introducing extensive deliberation processes might not be a top priority for *all* contexts of policy-making and related SPIs on various governance levels. Urgent policy issues might require rapid policy responses, such as to health crises (e.g. the Ebola epidemic beginning in 2013–14), or natural disasters and other catastrophes (e.g. the earthquake in L'Aquila 2009, or the Fukushima nuclear disaster in 2011). Nevertheless, initiating long-term deliberative learning processes might still be very valuable, for instance to help deal with long-term impacts or to increase preparation for future disasters.

We now expand on the basic deliberative ideals to distill more specific criteria for discussing SPIs for longer-term, wicked problems.

**Distilling criteria from building blocks of deliberative policy learning**. SPIs should be designed to allow for deliberative and inclusive, participatory learning and justification processes concerning the wicked policy problems and the related policy solution spaces—in short: *deliberative policy learning*. Contributing to deliberative policy learning should be a primary goal, and key promise, of SPIs. While "deliberative" refers to the inclusive and argumentative way of designing the process, "policy learning" can be understood as an updating of beliefs about policies (i.e., their rationale, performance or required institutions) resulting from a combination of social interaction, personal experiences, value change and scientific policy analysis (Dunlop and Radaelli, 2013).

Although SPIs can hardly ensure the comprehensive realization of deliberative policy learning in political processes alone (see section below on actual achievements of assessments for other relevant factors), SPIs could at least realize some major *building blocks* (i.e., *enabling conditions*) of deliberative policy learning that are highlighted in the literature on deliberative democratic theory. Table 3 provides an overview of four such major building blocks and subsequent criteria for SPIs. We distill these criteria from the building blocks in light of what was said in the introduction about the policy context and the general roles and guidelines for scientific expertise in these policy contexts. Being consistent with, but going beyond the seminal SPI criteria of salience, credibility, and legitimacy (see introduction), our criteria thus envision the realization of four major building blocks of deliberative policy learning within SPIs. This is similar to Miller and Erickson's approach (2006).

***Representation***, the first building block of deliberative policy learning, aims at facilitating the inclusion of those subject to a decision, and especially of competing voices. The encouragement of greater public participation in policy debate—including participation in the framing of policy goals, greater public awareness of the need for cooperative solutions, and mutual respect of differing viewpoints—can potentially lead to better collective decision-making (Chambers, 2003: 316; see also Gutmann and Thompson, 1996). Everyone who is affected by a decision-making process should also have a chance to be involved in the deliberative process leading to this decision. In democratic societies, all citizens are part of the general public of their state and are—at least in principle—able to participate, for instance by raising their voice in the media, at public demonstrations, or to organize into interest groups, including as members of political parties. To claim effectiveness and legitimacy it is not necessary for every viewpoint to find its way directly into a parliamentary decision. Different institutions and actors can serve as intermediate suppliers of a political core area to screen and structure discourses, being an amplifier (as well as filter) for particular messages and interests (Habermas, 1996b). The fairer, more equal

**Table 3 | Overview of criteria for science-policy interfaces responding to wicked policy issues, distilled from building blocks of deliberative policy learning**

| *Building block* | *Criteria for SPIs* | |
|---|---|---|
| Representation | Incorporating wide variety of viewpoints and stakeholders | Input side |
| | Incorporating wide variety of scientific insights and disciplines | Input side |
| Empowerment | Critically scrutinizing access barriers; supporting development of policy options and scenarios reflecting marginalized viewpoints | Input side |
| | Critically scrutinizing requirements to adequately participate; organizational support throughout the process | During the process |
| Capacity building | Building internal capacity of participants: knowledge integration and synthesis | During the process |
| | Realizing external capacity building: providing knowledge about implications of alternative policy pathways, disclosing key uncertainties and normative assumptions | Output side |
| Spaces for deliberation | Realizing continuous and iterative face-to-face deliberation | During the process |
| | Realizing vertical and horizontal linkage of spaces for deliberation | Input and output side |

and inclusive a deliberative procedure is, the more *legitimate* the decision based on such a procedure would be. The better the link between the multifaceted perception of a specific problem and a decision-making body (or SPI) dealing with this problem, the more *effective* it should be. To realize this goal, appropriate representation is also of key importance for SPIs responding to highly value-laden and disputed policy problems. This also is easily observable, for example, in the IPCC where representation is frequently a major issue discussed by the country delegations.

To distill criteria for SPIs from this building block, realizing proper representation is fostered by incorporating both (1) a high diversity of viewpoints and stakeholders, and (2) a wide variety of scientific insights (Pregernig, 2006; Whitfield *et al.*, 2011). The latter should also include the representation of different disciplinary knowledge as well as of knowledge systems that transcend traditional scientific domains including, for example, traditional ecological or local knowledge. Facing the global magnitude of many wicked problems, engaging with all affected is clearly infeasible. Nonetheless, the challenge for large-scale policy advisory processes is to determine which stakes are relevant or important enough to be incorporated into specific SPIs. Indeed, this is often a point of contention. One potentially promising method of addressing this is through discursive representation, i.e. by identifying the most prominent societal discourses relevant to a particular debate and ensuring they are represented—as for instance in the case of discourses within United Nations Framework Convention on Climate Change (UNFCCC) negotiations (Stevenson and Dryzek, 2012). Discursive representation could help to minimize the extent to which groups rely on power in deliberations, and increasing the importance of higher-quality arguments and rationalization (Deitelhoff, 2012).

Following this recurring tension between receptive representation and powerful interests, deliberative processes should also enable some form of ***empowerment*** as the second building block. This is relevant especially for previously marginalized groups or overlooked viewpoints that have too little (institutional) power to organize themselves appropriately e.g. across all policy levels, but are meaningful enough to have a say in the specific policy process. Many proposals for actual methods of empowering marginalized viewpoints within a deliberative democratic framework might have been vague at best (Chambers, 2003), leaving much room for improvement. Both sensitivity at the input side, which involves scrutinizing access barriers for marginalized viewpoints, and sensitivity regarding their requirements to adequately participate during the process are essential aspects of empowerment. This also applies to SPIs designed in response to wicked problems. Regarding global environmental assessments, for example, attempts to create a unified scientific framework to inform policy also fostered the exclusion of many voices from having a meaningful input to global decision-making (Miller, 2003,2004). Prominent examples for exclusion include indigenous groups, local knowledge-holders, or academic disciplines traditionally excluded from the provision of scientific policy advice. Indigenous and local groups often don't have the financial or institutional means to organize visibility for their perspectives, even though those might clarify important trade-offs of policy solutions on the ground. Avoiding such exclusion is an ongoing challenge for SPIs with deliberative aspirations (Berkes *et al.*, 2006). Empowerment therefore aims to diminish unjustified barriers on the input side, and to enable meaningful contribution of marginalized viewpoints to scientific policy advice.

More specific criteria for SPIs thus target the inclusion of marginalized viewpoints, different disciplinary insights, local knowledge systems or institutional knowledge both (1) by actively targeting underrepresented viewpoints at SPIs and (2) by active organizational support for marginalized groups throughout science-policy processes. To realize the former, experts and marginalized stakeholders could, for instance, co-produce elaborate policy options or broader policy scenarios that are in line with the knowledge systems, values and educational backgrounds of the marginalized stakeholders, and finally integrate these policy options and scenarios back into the scientific policy advisory process (using deliberation and tools appropriate for integrating knowledge systems). While Wiklund notes that the tools, such as public hearings and submissions, usually used to foster participation in science-policy interactions are limited, there is "an increasing experimentation with more inclusive, dialogue-based tools" (Wiklund, 2005: 289). One example of a highly promising tool to integrate knowledge systems and empower marginalized viewpoints is participatory scenario development (Patel *et al.*, 2007). A very functional tool for providing organizational support throughout the process is, for instance, to install secretariat-like bodies responsive to the individual needs of participating stakeholders.

Effective deliberative policy learning also presupposes ***capacity building*** as another important building block. Deliberative democratic theory often claims that all participating individuals *ideally* have appropriate capacity to constructively contribute to and influence decision-making (Cohen, 1997). This is especially important to avoid perpetuating extant power structures which tend to favour more highly educated and already empowered individuals or groups. While achieving equal capacity among

participants is likely impossible, two things can still be done: first, realizing internal capacity building of the participants to allow for meaningful iteration and intellectual progress in deliberation processes, and second, realizing external capacity building, for instance by transparent deliberations and insights, leading to an accessible outcome for subsequent decision-making. Effective capacity building is also decisive for SPIs, especially if they aim to contribute significantly to the resolution of longer-term, wicked policy problems (Miller and Erickson, 2006).

Two more specific criteria can be distilled from the aim of capacity building: (1) the diverse participating stakeholders should have the chance to learn not only how to approach different information and knowledge properly, but also how to integrate and synthesize different kinds of knowledge in a scientifically sound, collaborative and policy-relevant manner (internal capacity building). While working collaboratively with multiple disciplines and knowledge systems is no easy task, the iterative learning process among various actors in this regard may also create long-term benefits for many SPI participants. (2) Even more importantly, SPIs should provide the public with capacity in terms of the required *knowledge about different policy pathways and their practical implications* (see introduction). This external capacity building presupposes an integration (but not necessarily reconciling) and synthesis of knowledge across different disciplines and knowledge systems, and requires transparency of key uncertainties and normative assumptions. Such explorations of the policy solution space can become a valuable point of reference for decision-making processes, potentially even across governance levels.

A final key building block aims at the provision of actual ***spaces for deliberation***. Without direct interactions between different stakeholders, deliberative policy learning can hardly be realized (Wiklund, 2005). Deliberation processes among multiple stakeholders can stimulate mutual understanding and learning; these communicative activities can also increase mutual trust and enhance cooperation (Cole, 2015). For SPIs, the dynamics of continued and inclusive face-to-face deliberation can facilitate reasonable policy learning and even convergence regarding the multifaceted viewpoints and different disciplinary accounts, and may change individual and collective preferences. Engaging in interdisciplinary collaboration seriously, and challenging traditional problem framings or other assumptions in new ways, are examples of how spaces for deliberation can facilitate new insights and increase the evidence base particularly regarding the policy solution space (Reid *et al.*, 2006). In contrast, technocratic approaches might at best advocate limited input from non-scientific actors to e.g. co-determine the initial scope of an individual SPI process or by holding a session where the results are in essence "taught" to policy-makers, following a linear model of SPIs. Moreover, linkage of different (already existing or novel) spaces for deliberation is crucial to increase the effectiveness of deliberation processes and their influence on various policy processes on different governance levels, and to allow for mutual inputs to and exchange between different spaces for deliberation (including exchange with other scales and government levels) at the input side or output side of SPI processes. Linking with other deliberation processes can also reduce the burden of expectations for an individual deliberation space (e.g. within SPIs).

Two specific criteria for SPIs directly from this building block thus are: (1) the provision of effective opportunities for continuous, open and iterative face-to-face deliberation between different stakeholders involved at SPIs (including governmental officials and scientists); and (2) a high degree of vertical and horizontal linkage of spaces for deliberation. The former can considerably be fostered by having an appropriate and clear mandate from governing bodies in this regard. In some cases, SPIs may benefit from technical tools such as online platforms to decrease the various costs of face-to-face meetings involving large groups. Regarding the second criterion, inputs to global SPI processes from lower-scale spaces for deliberation could be realized, for instance, by integrating the findings of local or regional mini-publics (Niemeyer, 2014), by organizing regional consultation processes or by affiliating to existing subsidiary SPI bodies on the national or local level. Linkage at the output side of SPIs can help ensure proper utilization and dissemination of the SPI products and results (Cash *et al.*, 2006). This linking can increase the efficiency of resource-intensive deliberation processes by taking advantage of what is already being done and expanding the network of involved and interested actors beyond what the original process might have achieved. Linking to policy processes, for example by directly engaging with actors who play a role in policy debates, is crucial for SPIs to have an influence on decisions (Fouilleux, 2004).

All these criteria will guide our discussions of alternative SPIs for longer-term, wicked policy issues. However, these criteria are limited in several respects: (1) they do not cover all possible aspects of deliberative policy learning, but only focus on a few crucial aspects as a "proxy" to deliberative policy learning; and (2) these criteria would be insufficient to comprehensively guide the design of all varieties of SPIs, which would require more specific and differentiated criteria (also with regards to potential trade-offs between the four building blocks or with other values). The primary purpose of our criteria is to help reveal the *relative* prospects of alternative SPIs responding to wicked issues in the following section. Despite these limitations, our criteria are potentially valuable means for drawing attention to the most normatively important aspects (based on widely accepted ideals from the diverse literature on deliberative democracy), rather than highlighting only some more "technical" requirements for SPIs or overly narrow metrics of SPI successes.

## Realizing the building blocks of deliberative policy learning

Employing these criteria for what all SPIs should envisage when addressing longer-term, wicked policy problems, we now discuss the relative extent to which different SPIs may contribute to deliberative policy learning in terms of potentially realizing the major building blocks. Due to space constraints, we cannot comprehensively discuss all SPI types (six) in light of all criteria (eight). Rather, for each of the building blocks and related criteria, we selectively elaborate in this section on those SPIs that show relatively high potential—which will mostly be large-scale assessments, particularly the PEM-inspired ones. Aside from considering the relative prospects of particular SPIs, we also point out major challenges, conditions and limitations of SPIs in meeting the criteria distilled above.

The main method used for this discussion is to conceptually combine the characteristics of different SPIs identified above (see particularly Tabs. 1 and 2) with the criteria (see Table 3 for an overview). We hypothesize about the implications of different SPI designs for the potential of these SPIs to meet the criteria. As such, our method is mainly theoretical, but will be complemented by some empirical underpinnings and illustrated by various examples. While this section cannot provide a full-fledged evaluation and empirical comparison of the alternative SPIs, it may offer an interesting set of hypotheses on the performance of alternative SPIs that could be further examined in future, specific comparative case studies.

**Realizing representation**. As discussed above, wicked policy problems involve a wide array of disputed issues and competing stakes. This complexity actually presents SPIs with an

opportunity to realize the criteria associated with representation and to be regarded as relatively legitimate (Mitchell *et al.*, 2006; Norgaard, 2008a; Reed, 2008; Kowarsch, 2016b). One way to achieve this is by designing SPIs to be collaborative and inclusive processes, striving for fair and equal representation of different viewpoints and stakeholders as well as incorporating insights from a wide variety of scientific disciplines. SPIs must be transparent in their activities in order to remain accountable to the diversity of individuals and groups represented.

Large-scale integrated scientific assessments seem to have great potential in this regard. Some prominent examples of highly inclusive and collaborative SPI processes include the Millenium Ecosystem Assessment (MA), the International Assessment of Agricultural Knowledge, Science and Technology for Development (IAASTD), and the more recently established Intergovernmental Platform for Biodiversity and Ecosystem Services (IPBES). Both the MA and the IAASTD assessment processes explicitly engaged with diverse groups who had multiple roles and participated in a diversity of activities over the course of entire assessment processes (Norgaard, 2008a; Feldman and Biggs, 2012). The latter was expressly designed as a social experiment in terms of bringing such a diversity of viewpoints and stakeholders together (Watson, 2009). The IPBES process, which has built upon knowledge gained inter alia by the MA, IAASTD and IPCC, can be seen as the most ambitious assessment to date regarding inclusiveness. While some have criticized IPBES for not yet living up to these ambitions (Hotes and Opgenoorth, 2014; Montana and Borie, 2016), IPBES has nonetheless made significant strides in terms of encouraging collaboration between a broad diversity of disciplinary expertise, explicitly striving to achieve a balance of participants according to gender, geographic location and affiliation, and openly acknowledging and integrating different knowledge systems and viewpoints in both its design and assessment practices (Díaz *et al.*, 2015). In addition, all PEM-inspired assessments[4] can, in theory, strongly contribute to realizing representation also by emphasizing the explicit, transparent and direct inclusion of divergent viewpoints and different policy narratives in the exploration of alternative policy pathway scenarios (Edenhofer and Kowarsch, 2015; Kowarsch, 2016b). The explicit discussion of some divergent viewpoints and normative assumptions in the recent IPCC Working Group III assessment of climate change mitigation options is a good example in this regard.

Compared with nearly all other SPIs, large-scale assessments are long-term processes with a high profile and often an intergovernmental structure. Engaging with diverse stakeholder groups and their viewpoints is usually part of their practice. Moreover, they can often be more multi- and interdisciplinary than other SPIs simply because of the higher number of experts involved that may represent different disciplines and competing approaches, also due to the more elaborate author nomination processes (see Tabs. 1 and 2 for details). Large-scale integrated scientific assessments, particularly the PEM-inspired ones, thus have a higher potential for achieving the criteria associated with representation as compared to, for example, conventional research papers, policy reports, expert panels, or impact assessments, which usually do not engage so extensively with stakeholders, or with diverse viewpoints.

However, there are also numerous challenges and limitations to realizing representation in assessments and other SPIs (Sénit *et al.*, 2016), of which we will highlight three. Firstly, there are inevitable and significant legitimacy issues involved in selecting, interpreting and evaluating the societally most relevant viewpoints. This challenge might actually be less pronounced at lower levels, for example in smaller-scale impact assessments where identifying the most relevant stakes and selecting representatives for these groups might involve less controversy. Questioning the legitimacy of who exactly is included in deliberations in turn leads to questioning the deliberative learning process itself as well as decisions the process contributes to.

Secondly, there are numerous methodological barriers which must be overcome in order for the social sciences and other traditionally underrepresented disciplines to more meaningfully participate in SPIs. For example, there is still a long way to go in order for social sciences to be able to rigorously translate divergent worldviews into consistent policy scenarios and their evaluation, and to respond to the most pressing societal questions, such as political economy features (e.g. winners and losers) of policies (Norgaard, 2008a; Carraro *et al.*, 2015; Victor, 2015). In addition, existing social science methods that support the integration of various viewpoints like participative scenario development tend to be based on contradicting epistemological principals, as they come from different schools of thought which makes their integration difficult (Schultz, 2006; van Asselt *et al.*, 2010).

Thirdly, the potential for legitimate representation of divergent viewpoints when assessing policy pathway scenarios is limited by the fact that a single assessment process can realistically only explore a few such scenarios in-depth (in terms of multi-criteria policy evaluation) since this is typically resource-intensive. This means that some potential scenarios, which might be preferable to some viewpoints or better reflect some knowledge systems, must be excluded from the analysis in order for it to remain feasible. Moreover, there can be resistance in some assessments from particular governments or other interest groups against an open exploration of alternative policy viewpoints and pathways (Edenhofer and Minx, 2014).

**Realizing empowerment**. As several case studies suggest, inclusive SPI processes can help empower local communities and marginalized social groups (O'Faircheallaigh, 2010), and integrate indigenous and other knowledge systems with prevalent scientific discourses (e.g. Berkes *et al.*, 2006 on MA; Thaman *et al.*, 2013 on IPBES). But which SPI is most promising in this regard?

In terms of supporting the development of policy options and scenarios that reflect marginalized viewpoints (as the first criterion associated with empowerment), there are significant opportunities for empowerment at smaller SPI scales (see Tabs 1 and 2 for details), and in fact working at a lower scale can actually improve the specificity of approaches to certain groups and enable a more comprehensive understanding of power asymmetries. For example, as acknowledgement of the power dynamics at play in environmental impact assessments has increased so too has research seeking to improve the level of equality between different participants (e.g. Cashmore and Richardson, 2013). In addition, scenario exercises have been undertaken in smaller scale SPI activities, for example in coastal planning (Tomkins *et al.*, 2008) and land use assessments in the United Kingdom (Foresight, 2010) and in multi-scale assessments seeking to integrate knowledge accumulated at local scales into progressively larger-scale processes (Kok *et al.*, 2007).

However, these examples also show that emphasizing empowerment through participatory scenario-building requires longer time frames, access to the expertise necessary to organize such exercises, as well as mechanisms to motivate participation amongst marginalized groups. If scenarios are not co-developed but rather provided in a top-down fashion, or if only "advocacy pieces" for powerful groups are provided, they lack the sensitivity to marginalized groups' reality required for empowerment by forgoing deliberation, lowering their potential for realizing this building block.

The larger resource requirements (particularly in terms of time) for empowerment, which may be out of reach for some SPIs, point towards a higher potential for empowerment—in terms of both criteria associated with this building block—through large-scale scientific assessments which last for several years, even though these may be poorly funded in some cases. One example of how large-scale assessments can support the development of particular policy pathways can be found in UNEP's Global Environment Outlook (GEO) Small Island Developing States (SIDS) Outlook, which provided a valuable source of information as well as a cohesive narrative for an often-marginalized group of countries in international negotiations on environmental and development issues. Another example is the IAASTD, which went to great lengths to ensure that minority voices were heard and taken seriously, thus empowering, for example, small-scale farmers, and especially women (Stokstad, 2008; Labbouz and Treyer, 2010). The MA employed a participatory scenario-building exercise which expressly aimed at integrating different epistemologies including across qualitative and quantitative information sources, across major disciplinary divides, as well as across scientific knowledge and other systems (including local and traditional ecological knowledge). This scenario-building exercise served to empower different groups by identifying and then striving to restore balance to power asymmetries observed in previous assessments (Bennett and Zurek, 2006). Moreover, one could argue that any assessment seriously following the PEM science-policy model, given its idea of a cartography of alternative policy pathways, also would put high emphasis on the collaborative development of policy options and scenarios that reflect relevant, but marginalized viewpoints. PEM-inspired assessments thus seem to provide a clear theoretical avenue for the integration of minority viewpoints and marginalized groups.

Also in terms of critically scrutinizing requirements to adequately participate (e.g. organizational support throughout the process), many large-scale assessment processes might be more promising than other SPIs (for example, the rather supply-driven one-off studies, see Tabs. 1 and 2), given their formalized, longer-term procedures and institutions in this regard. This may be particularly true for intergovernmental assessment bodies related to UN processes, such as GEO or the IPCC assessments, for instance.

Besides the costs (in terms of time, funds and so on), another major challenge—for all SPIs—with realizing the two criteria associated with empowerment is that altering pre-existing power structures during the course of SPI activities could have unknown and potentially adverse side effects. Since it "appears unlikely that those who hold power will yield gracefully to groups pushing for a share of it," altering the balance of power through SPIs runs the risk of actually entrenching power asymmetries if the groups in power push back (O'Faircheallaigh, 2010: 23). Furthermore, participation in SPI processes presupposes at least a certain degree of empowerment already, at least in the sense that there is some form of organization and the potential for representation which would enable participation. For the most marginalized societal groups, this is rarely the case and some form of empowerment must actually occur before participation in SPI processes, and potentially further empowerment through these processes, can take place (Esteves and Vanclay, 2009).

**Realizing capacity building**. While participants in smaller-scale SPI activities (see Tabs. 1 and 2) likely also learn about synthesizing knowledge through their participation, the specific characteristics of *large-scale* assessments (see above) make this type of capacity-building not only a precursor to successfully producing an assessment, but also require learning to be more diverse in terms of the sheer number of disciplines, stakeholders and different viewpoints involved. Interdisciplinary scientific assessments for environmental governance, like the IPCC, MA or UNEP assessments, prompt a general shift in scientists' perspective from local to global and disciplinary to interdisciplinary, broaden their perspectives and help them to develop deeper and richer analyzes of complex systems (Norgaard, 2008b). Recurring assessments, which often have at least some returning scientific experts from one cycle to the next, can benefit from the experience that these returning experts gained in previous assessment cycles in terms of interdisciplinary collaboration and knowledge integration and synthesis. This is important since these tasks go beyond common academic endeavors and require learning together.

Building capacity inter alia for working across disciplines and knowledge synthesis is, however, no easy task. Since participants of large-scale assessments are nearly always volunteers, a major challenge with this is ensuring that actors still have motivation and incentives to participate given the significant demands their participation entails (Carraro *et al.*, 2015). Norgaard reflected that in the context of the MA "[m]ultiscientist deliberation requires considerable humility, hard work, respect, and patience" (2008b: 6). One potential way forward could be programs such as the IPCC "Chapter Scientists" program, which decreases the burden on authors while also enables learning amongst early career researchers (Schulte-Uebbing *et al.*, 2015).

Moreover, as the "external capacity building" criterion claims, public debates on longer-term, wicked policy problems require a credible, rigorous synthesis of the available knowledge (and uncertainties) across different disciplines, including insights from different worldviews and working with different knowledge systems in a policy-relevant manner—which primarily means the exploration of the direct and indirect effects of available alternative policy options. This in turn requires the capacity to not only acknowledge and understand knowledge from diverse perspectives but also to work collaboratively to integrate it.

All of the SPIs presented in Tabs. 1 and 2 can build external capacity in terms of enabling policy learning and influencing decision making. This is particularly the case when a SPI has a direct path of influence on decisions, as is often the case for example with impact assessments (O'Faircheallaigh, 2010) or with other SPI activities which were expressly requested by policy makers.

However, large-scale integrated scientific assessments seem to have particularly high potential here (Mitchell *et al.*, 2006; Norgaard, 2008a; Kowarsch, 2016a); in fact this is their major strength when measured up against alternative SPIs. Ideally, these assessments provide a systematic processes for knowledge synthesis and integration involving a large number of scientific experts from different disciplines, elaborate synthesis methodologies as well as extensive, multi-stage peer-review processes to ensure credible results. When designed and conducted this way, such assessments can integrate knowledge across different disciplines, competing approaches and various policy fields in a more rigorous, credible and comprehensive manner than, for example, policy advice or consulting provided by smaller, and more homogeneous groups of experts (particularly if they lack peer-review). Again, PEM-inspired assessments may have particularly high potential regarding the provision of policy-relevant knowledge maps (that is, external capacity building), in the sense that the focus on exploring policy alternatives and their implications emphasizes a highly policy-relevant knowledge synthesis, which explicitly includes of a vast diversity of viewpoints, disciplines and approaches, and could directly address disputed normative issues as well.[5] Such an emphasis

can increase the quality, comprehensiveness and credibility of knowledge synthesis to inform policy processes.

Furthermore, the potential scale of influence of large-scale assessments (see below for examples) give them a higher relative potential in terms of external building capacity among much larger groups, compared with other SPIs. In addition, intergovernmental large-scale assessment processes can provide the necessary methodologies for national actors to take environmental policy-making forward *domestically*, thus building external capacity which has been acknowledged by numerous governments in the process of reforming the IPCC (2013). Large-scale assessments can also motivate smaller-scale SPI activities. For example, the GEO-5 assessment process was a catalyst and provided the methodological foundations for regional assessments all over the world (Ivanova, 2009).

Table 4 provides an overview of some theoretical advantages of PEM-inspired and other assessments regarding the realization of both internal and external capacity building.

However, given the methodological, organizational and other limitations and challenges that large-scale assessments face (for example, Carraro *et al.*, 2015; Victor, 2015), a fully comprehensive assessment of all relevant aspects of the complex policy solution space as ideally envisaged by the PEM is impossible. A further limitation to this approach is that many scientists and decision-makers continue to underestimate the need for, and significant challenges of, seriously and rigorously exploring policy pathways and their implications.

Another major challenge regarding external (but also internal) capacity building is the exploding literature in some fields. Many large-scale assessments, such as the IPCC assessments, envisage the provision of comprehensive assessments of the available literature. This is a noble idea. However, as the example of exploding literatures on climate change clearly shows (Grieneseisen and Zhang, 2011), the high number of existing publications in this field implies huge challenges for current assessment-making, and perhaps requires more sophisticated bibliometric methods in future assessment processes.

A further challenge related to external capacity building is that it is hard to prove that policy actors have indeed learned from the outputs of an SPI process. A truthful evaluation of the overall influence of one single SPI is hard, and can at best be conducted after one or two decades (Sabatier, 1988). Yet, multiple analyses illustrated that assessments have contributed to problem recognition, alternative narratives and conceptualizations of environmental changes and new framing of problems (Miller, 2000; Farrell and Keating, 2006; van Deveer, 2006; see below for more details). The influence of science on policy via discourses and frames result from a multitude of activities in which multiple political actors are involved and which they pursue simultaneously. Policy change results from the cumulative learning and deliberative processes.

**Realizing spaces for deliberation**. Spaces for deliberation can be provided by several SPIs if they facilitate the direct (or at least indirect, but effective) interaction between different stakeholders, including governments and scientists. While direct face-to-face interactions in particular can engender trust, in-person meetings are costly, and this cost only increases with the scale of the SPI and the related in-person meetings. Large-scale assessments which strive to bring a large and diverse group of individuals together, for example as in the IAASTD or GEO-5 regional consultations, must invest large amounts of resources (both time and funds) inter alia in designing the meetings, selecting the stakeholders and actually convening participants for a long enough time to enable deliberation. However, this seems

**Table 4 | Capacity building for policy deliberation: overview of the potential advantages of integrated scientific assessments, in particular the PEM-inspired ones, measured up against other science-policy interfaces**

| *Sub-criteria for science-policy interfaces to realize capacity building* | *Conventional integrated scientific assessments* | *PEM-inspired integrated scientific assessments* | *Other science-policy interfaces* |
|---|---|---|---|
| Provision of policy-relevant knowledge on direct and indirect effects of different policy options (see introduction) | Engagement with stakeholders and exploration of policy scenarios helps ensure policy-relevance | Extensive engagement with stakeholders ensures policy-relevance, together with the highly relevant exploration of policy alternatives and their implications | If any, the engagement with stakeholders helps ensure policy-relevance; impact assessments are policy-relevant, but only few options and effects considered; conventional scientific research often does not address policy-relevant issues |
| Incorporating exploding bodies of literature in some fields (for example, Grieneseisen and Zhang, 2011; McKinnon *et al.*, 2015) | Facilitated by involving a high number of researchers, and by employing meta-analysis methodology | Facilitated by involving a high number of researchers, and by employing meta-analysis methodology | Only effective when involving a high number of researchers and when employing meta-analysis methodology |
| Rigorous synthesis across different disciplines, approaches, policy fields, scales (making research gaps transparent) (see introduction; Norgaard, 2008a) | Credibility through disciplinary diversity in author teams (and diversity of viewpoints and approaches), and through elaborate synthesis and integration methodologies | Focus on pathway exploration facilitates such synthesis; credibility through diversity of disciplines & viewpoints involved and elaborate synthesis and integration methodologies | Due to the complexity and magnitude, conventional research papers cannot deliver such synthesis; also small-group policy advice typically does not deliver such synthesis |
| Evaluation of uncertainty, disagreement and inconclusive (non-aggregated) results, particularly in social science research (see introduction) | Credibility through disciplinary diversity in author teams (and diversity of viewpoints and approaches), and through elaborate synthesis and integration methodologies | Credibility through disciplinary diversity in author teams (and diversity of viewpoints and approaches), and through elaborate synthesis and integration methodologies; focus on pathway exploration fosters such synthesis; | High variation; requires appropriate methodologies and the involvement of researchers with diverse viewpoints |
| Capacity building in terms of methods and skills for knowledge integration and synthesis | High potential given that these assessments are essentially synthesis processes | Particularly high potential given the focus on collaborative exploration of alternative policy pathways | High variation; success inter alia depending on the degree to which synthesis is done |

worthwhile. The MA, which also aimed to foster spaces for deliberation between very diverse actors seems to have been successful in the sense that "MA participants formed a learning community to connect their disparate disciplinary perspectives on socioecological systems" (Norgaard, 2008b: 6). Similarly, all PEM-inspired assessments may have a high potential to provide spaces for deliberation through their direct and collaborative exploration of highly relevant alternative policy pathways and their implications.

Other SPIs, such as conventional research, meta-studies with particular policy-relevance, or small-group policy advice (for example, the various policy recommendations provided by different competing reports), usually do not extensively provide such an opportunity for face-to-face interactions with diverse individuals, at least not to the extent observed in large-scale assessments. Impact assessments also often do not involve such extensive multi-stakeholder deliberation processes, often facilitating instead a one-way flow of information from the public to the assessment, identifying major potential impacts and making judgments regarding their perceived severity through public consultation meetings (Wiklund, 2005). This could in theory be done in a deliberative manner, for example having different actors explicitly explain arguments and values underlying their prioritizations of different impacts, which is augmented by the explicit mandate for most impact assessments to convene affected individuals (Bartlett and Baber, 1999; Wiklund, 2005). However, this is rarely achieved in reality, often because of time constraints posed by the pending project which prompted the impact assessment in the first place. Permanent expert committees and different types of integrated scientific assessments can better facilitate the provision of spaces for face-to-face deliberation, and can carry the authority to convene diverse groups which might not otherwise sit down together. That being said, multi-stakeholder, collaborative assessment processes theoretically have even higher potential in this regard due to the (ideally) extensive and formalized, large-scale involvement of so many different stakeholders over a very long period of time. The length of engagement is important as, in combination with face-to-face meetings, ongoing engagement can help to build trust over time which is an important precondition for truly open deliberation (Stevenson and Dryzek, 2012).

An interesting example here is the innovative mechanism that was introduced by UNEP to facilitate the production of GEO-SIDS and to provide a space for deliberation via an online platform organizing participants into communities of practice. Organizing meetings in person for such a geographically disparate group of participants given the short timeline and limited resources available would have been practically infeasible. Providing an online forum through UNEP Live, and organizing work via well-thought-out Communities of Practice, seems to have overcome this challenge and contributed strongly to a deliberative democratic learning process in GEO-SIDS. However, many participants had already engaged together in face-to-face deliberations in the past and so were already well-acquainted and had built up some degree of trust. This highlights the fact that perhaps online mechanisms can add value by building off of face-to-face meetings, but it is unlikely that remote deliberation can completely replace coming together in person.

However, a number of internal and external obstacles can limit ability to provide effective spaces for deliberation and are relevant to numerous SPIs (for example, Reed, 2008). Participatory mechanisms can be misused, for instance when prioritizing scientific expert participation in face-to-face meetings and only including non-scientific viewpoints through voting mechanisms, in case of an excessive focus on participation as a stand-alone method rather than conceptualizing it as part of a broader process, or by providing cursory spaces for deliberation in which participants' inputs are not taken seriously or are used only to provide post-hoc justifications for pre-determined decisions. Another challenge of providing spaces for deliberation within SPIs is to still come to clear and highly policy-relevant conclusions, even if they are undesirable for some of the political parties involved (for example, Edenhofer and Minx, 2014). In any case, effective deliberation presupposes the realization of the other SPI criteria discussed above.

Concerning the second criterion associated with spaces for deliberation, forging vertical and horizontal linkages, both to other SPIs and to decision-making processes, is a crucial part of realizing spaces for deliberation to adequately frame problems and their solutions. Forging these linkages presupposes knowledge of the complex and ever-changing landscape of different SPIs and decision-making processes, which might perhaps be more efficiently accomplished by a larger-scale process with a better overview of such a landscape as compared to a smaller-scale process. Often, the institutions mandating or involved in overseeing large-scale assessments have pre-existing linkages to or at least a strong knowledge of other SPI processes. However, it is important to note that this does not necessarily mean that the assessment process will capitalize on these linkages. One challenge, for example, is that there could be some sense of competition between different SPIs which leave organizers hesitant to share information or the fruits of intensive deliberative processes.

The Regional Consultations organized for the GEO-5 assessment, for instance, have been an attempt to convene representatives of the public at smaller scales and to coordinate their contribution to larger-scale decision-making via the large-scale assessments process. There are also vertical linkages between regional-scale GEO assessments and the global process which capitalize on deliberative processes contributing to these smaller-scale SPIs in the global assessment. Such linkages also exist in other SPIs. For example, the multi-scale assessments mentioned earlier which in one case involved deliberation via scenario-building exercises on the topic of land degradation and potential mitigation measures at local scales, the results of which were then incorporated into further participatory scenario exercises as part of a larger Mediterranean regional-scale assessment (Kok *et al.*, 2007). However, given the scale of large-scale assessments, their scope for identifying relevant deliberative SPI processes, as well as their relatively higher profile and ability to provide motivation for other SPIs to engage with them, large-scale assessments would fare comparatively well in terms of realizing linkages between spaces for deliberation.

One of the most prominent examples of engaging with policy actors in SPIs is through the approval processes which often come at the end of a large-scale assessments process. Here, policy makers representing governments from around the world come together to approve the final documents and in some case negotiate the text of a summary for policymakers document. While some argue that bringing policy makers into deliberations about content results in watering down of findings based on political (and unscientific) reasoning, this process nonetheless creates higher governmental buy-in and mutual learning than scientific reports or expert panels without stakeholder engagement (Mitchell *et al.*, 2006; Norgaard, 2008a; Reed, 2008; Kowarsch, 2016a). Such buy-in can enable the effective use of assessments as a deliberative learning platform. Impact assessments and the public deliberations they employ also have a direct linkage to policy debates, since they are formal requirements in many project proposals (for example construction projects) and must legally be taken into account before a final decision on a project is made (Bartlett and Baber, 1999; Wiklund, 2005).

However, these linkages are project-specific, limiting the breadth of influence, and also do not encourage governmental buy-in since governmental actors do not usually take an active role in deliberations. Once again, large-scale integrated assessments have a higher potential than smaller-scale SPIs, this time in terms of linking externally to policy debates.

To conclude, large-scale scientific assessments, and especially the PEM-inspired assessments, seem to have relatively high potential to realize many of the deliberative criteria. Within their lifespan, they can provide constant structures for inclusive and extensive deliberation and transparent justification. Scientific assessments constitute a valuable means not only of providing an inclusive space for argumentation and mutual learning but also of structurally linking the potentially unbounded realm of diversified public opinion and knowledge production with rule-making bodies at different policy levels. In taking up the available scientific knowledge and divergent viewpoints, scientific assessments must be highly receptive on their input-side, assuring a high degree of representativeness. The assessment process itself can produce consensus, convergence (through the iterative exploration of concrete future scenarios, areas of overlap or convergence between competing ethical claims or policy narratives would likely be revealed), or at least transparency over divergent viewpoints and underlying assumptions.[6] The assessment process also has the potential to empower participating actors and actors groups and to build capacity among and beyond participants. However, much depends on specific contexts and the more specific design of SPIs. This section only discussed the relative, basic theoretical potential of stylized SPIs.

## Assessments can actually contribute to deliberative policy learning

We have argued that assessments, and especially those informed by the PEM, may have at least as high or even higher theoretical potential as other SPIs for realizing major building blocks of deliberative policy learning. We now take a broader (ex-post) view of large-scale integrated scientific assessments as SPIs, in order to examine their *actual outcomes* regarding deliberative policy learning overall, drawing on empirical evidence from recent large-scale integrated assessment processes.

The effectiveness of assessments can be limited and some assessments have considerable weaknesses and shortcomings (see introduction). Many different factors, preconditions and obstacles (and sometimes tradeoffs) influence and possibly limit the success of assessments (Mitchell *et al.*, 2006; Sabatier, 2007), despite their sometimes explicit *ambition* to contribute to deliberative policy learning (for example, Edenhofer and Kowarsch, 2015). For instance, extensive stakeholder engagement and public participation techniques in general do not automatically lead to higher legitimacy and successful deliberation. Rather, they can sometimes even be counter-productive if not well-designed (Ryfe, 2005; Reed, 2008; Edenhofer and Minx, 2014), and become "just another layer of technocracy" (Rayner, 2003). Consequently, there are grounds for skepticism (for example, Gluckman, 2016), especially regarding the real potential of multi-stakeholder, transnational assessments (Rayner, 2003).

Realizing the criteria from above is thus a necessary but not sufficient condition for facilitating deliberative policy learning within SPIs. Hence, the question remains in what sense and to what extent existing assessments can be shown empirically to have substantially contributed to the facilitation of deliberative policy learning. While we cannot discuss the various factors influencing the impacts and influences of assessment processes on policy here, we provide some empirical evidence that diverse assessment processes have actually helped facilitate deliberative policy learning in the past. We draw on both extensive expert interviews, which point to policy learning outcomes from selected global environmental assessments, as well as an empirical document analysis of the IPCC contributions to UNFCCC debates. Our analysis confirms what a number of scholars have noted regarding the relatively high potential of increased deliberative SPI formats in general (for example, NRC, 1996; Dunlop and Radaelli, 2013; see also the Aarhus Convention) and for environmental assessments in particular to facilitate better environmental policy discourses (for example, Baber and Bartlett, 2001; Norgaard, 2008a; for a review, see Wiklund, 2005).

More specifically, we draw on empirical evidence gathered from 99 semi-structured interviews.[7] Seventy-six of these interviews were conducted with authors, producers, or other scientific experts involved in the production of either GEO-5 or IPCC WGIII AR5, 13 interviews were conducted with governmental officials who had participated in one of these two assessments, and 10 interviews were conducted with non-involved intended target audience of these and other global environmental assessments. While a large number of individuals were contacted from each group for interviews in an attempt to capture as many diverse perspectives as possible, actual interviews conducted depended largely on the response rate. For authors, 31.5% of requests resulted in an actual interview, 19.4% for governmental officials, and only 12.7% for intended target audiences. The interview duration extended from 20 min to 2 h (55 min on average). The interviews were mainly conducted via Skype and telephone between August 2013 and February 2015. All of the interviews were recorded and transcribed with the participants' prior consent. The interviews were analysed using Grounded Theory Analysis methodology in Max QDA software, with iterative coding employed in order to organize the vast amount of information collected following the guidance of Strauss and Corbin (1998). From within the general theme of GEA impacts and influence, we identified four strands of relevant information regarding the contribution of GEAs to creating conditions for policy learning, the perceived value of learning processes both within and external to GEAs, and the extension of learning beyond governments and scientific authors to include also civil society actors. In general, interviewees with long histories of participating in global environmental assessments placed particular value on their function as engagement platforms and the fact that they help integrate numerous divergent viewpoints. In order to provide a more complete picture, we complement the results of our empirical analysis by referring to other examples analysed in the past.

First, the interviews revealed that assessments actually contribute to the emergence of policy learning. They confirmed that assessments bring people together and create or solidify the "sustained collegial interactions" that are central to reflexive learning about policy issues and the achievement of both policy and scientific goals (Clark *et al.*, 2011). For instance, the empirical material we collected points to the fact that the IAASTD assessment brought people together to discuss issues which some key actors had not contributed much in the past. This was also the case of the assessment on Long-Range Transboundary Air Pollution (LRTAP), which mobilized a wide range of actors beyond the initial core of scientists from developed countries on the issue of long-range transboundary air pollution (Selin, 2006). LRTAP assessment meetings helped various actors to get to know each other, which facilitated the official negotiations later on (Selin, 2006). Thus, assessments bridged the cultural gap between scientists and policy-makers so that policymakers better understood the practical implications for policy of the different scientific and normative perspectives, and scientist got sensitized for the political risk and value of different pathways.

Second, participants find particular value in the learning processes that occur during assessment processes. The breadth and the richness of the information synthesized in large-scale assessments allow researchers, stakeholders and government representatives to forge connections between disciplinary silos, policy fields and national experiences. Such processes help policy actors reframe and refine their understanding of the interconnectedness of the environmental, societal and economic dimensions of policy issues, and to link them in innovative ways. For instance, GEO-5 and GEO-6 have created opportunities for policy actors to negotiate the mainstreaming of environmental concern across the new global development goals and enhance the integration of this better understanding of interconnected environmental and economic aspects, both in the formulation of the goals and the monitoring of progress towards their achievement. The GEO series thus influenced the Post-2015 Development Agenda (Dodds *et al.*, 2014). Assessments can also, according to a government representative involved in GEO-5, provide participants with a unique opportunity to learn about experiences in other countries. The IPCC focal point of a developed country explains that the UNFCCC Structured Expert Dialogue on international climate policy 2014–2015 provided an opportunity to learn about the concerns of different governments and their perspectives, and to exchange experiences about how governments have responded to similar issues in the past.

Finally, assessments not only generate learning among researchers and policymakers but they also support policy learning among civil society organizations, which collaborate with and provide scientific policy advice to local governments, for instance. One civil society interviewee stated that "we can point towards experiences in other jurisdictions across the world and highlight which experiences have been made there and which lessons have been learned and what things might be better to avoid when designing [a policy instrument]" highlighting learning that can be translated directly into scientific policy advice. Civil society actors use the lessons they draw from global environmental assessments to raise awareness for environmental issues and confront running conservative governments with the lessons learned in assessments processes: "[global environmental assessments] are mainly useful … as tools to communicate [messages] to media and also to publics". For some, the best use of an assessment is via civil society groups who design arguments relative to policy instruments and economic arguments about the creation of jobs and green growth, in order to challenge the status quo. As one interviewee stated "the real policy-making process is with civil society".

**The IPCC as an historic, unprecedented deliberation platform.** There is some evidence that the IPCC assessments, particularly the recent PEM-inspired assessment provided by the Working Group III on climate change mitigation (IPCC, 2014a), strongly influenced the negotiations leading to the Paris Agreement on climate change (adopted in December 2015). The Paris Agreement has to be regarded as a milestone in international environmental governance, but also a huge achievement of multilateralism in general. Again, the decisive question is not to what extent the IPCC directly influenced policy decisions, but rather to what extent the IPCC helped facilitate deliberative policy learning in these cases.

In her capacity as Executive Secretary of the UNFCCC, Christiana Figueres stated that "the ambitious agreement reached in Paris would not have been possible without the IPCC".[8] The Paris Agreement actually makes strong reference to the scientific information provided by the IPCC, for instance:

> In order to achieve the long-term temperature goal set out in Article 2, Parties aim to reach global peaking of greenhouse gas emissions as soon as possible, recognizing that peaking will take longer for developing country Parties, and to undertake rapid reductions thereafter *in accordance with best available science*, so as to achieve a balance between anthropogenic emissions by sources and removals by sinks of greenhouse gases *in the second half of this century*, (Article 4.1, emphasis added) (See http://unfccc.int/resource/docs/2015/cop21/eng/l09.pdf, accessed 15 May 2016).

Moreover, in the run-up of Paris, the G7 leaders published a declaration with an explicit reference to concrete numbers taken from the IPCC assessment, explicitly highlighting the concept of "decarbonization", a key concept in the IPCC Working Group III assessment:

> "Mindful of this goal [to hold the increase in global average temperature below 2 °C] and considering the latest IPCC results, we emphasize that deep cuts in global greenhouse gas emissions are required with a decarbonisation of the global economy over the course of this century. Accordingly, as a common vision for a global goal of greenhouse gas emissions reductions we support sharing with all parties to the UNFCCC the upper end of the latest IPCC recommendation of 40 to 70% reductions by 2050 compared to 2010 recognizing that this challenge can only be met by a global response" (G-7 Leaders' Declaration, June 2015) (See https://www.whitehouse.gov/the-press-office/2015/06/08/g-7-leaders-declaration, accessed 15 May 2016).

To illustrate the similarities, see the IPCC summary for policymakers:

> "Scenarios […] consistent with a likely chance to keep temperature change below 2 °C […] are characterized by lower global GHG emissions in 2050 than in 2010, 40 to 70% lower globally, and emissions levels near zero GtCO2eq or below in 2100" (IPCC, 2014a: 10–12).

The numbers mentioned in the G7 declaration, and the long-term ambition to achieve zero emissions in the course of the century, have been an important building block of the options discussed in the UNFCCC negotiations in Paris. In the end, these numbers were dropped, but the requirements of achieving net zero emissions in the long-term has been maintained in the final text of the Paris Agreement.

As some of the co-authors of this paper were involved in the IPCC assessment processes and attended some IPCC plenaries as well as the very effective Structured Expert Dialogue on international climate policy held in 2014 and 2015 under the UN Framework Convention on Climate Change (UNFCCC), we had the opportunity to witness, in our interpretation, the remarkable effectiveness of these IPCC-related processes in terms of enabling deliberation and learning about policy problems and the policy solution space (that is, policy alternatives and their implications). In this context, all stakeholders involved (including the scientists) especially learned about the costs, technologies, policies, institutional requirements, synergies and tradeoffs of ambitious climate change mitigation efforts to keep global mean temperature below 2 °C as agreed by the international community.

Multiple countries consider the IPCC to be the main source of information to the UNFCCC (e.g. IPCC, 2013) and a "reference point for policy debates at global level" (EU, a view also shared by Denmark, see IPCC, 2014b). Finland (IPCC, 2013) perceives the

IPCC to be "the most significant science-policy interface institution, which has a global impact on political decision making" and points at "the ability [of the IPCC] to respond efficiently to questions and needs that arise in the UNFCCC process". A high-level UNFCCC interview respondent points that: "Of course, we are the main client of IPCC and as such, we interact with IPCC in various ways". Interfaces such as the UNFCCC Structured Expert Dialogue have evolved towards providing platforms by which information is exchanged between UNFCCC and IPCC in an iterative manner and in a more informal manner than international processes, which often emphasize negotiation rather than actual deliberation, usually allow for, as our interviews reveal. One government official interviewed explained that the Structured Expert Dialogue gives "you an opportunity of a real give-and-take, (…), a real-time Q&A (…)". A UNFCCC high level interviewee explains that parties who had participated in a previous expert structured dialogue would come back to ask the IPCC experts more specific questions, pointing at potential contradictions that had developed over time in the scientific discourse of IPCC authors. For instance, the Structured Expert Dialogue in Lima 2014 helped clarify the positions of the countries regarding short-lived climate pollutants, which accelerated the UNFCCC negotiations as the negotiators could start to discuss the synergetic effects of these pollutants right away—for instance synergies between mitigation and adaptation short-term health effects and longer-term climate effects.

Although representation in the IPCC is still focused on governments and scientists, it is in our opinion a unique, historic deliberative learning platform. International climate policy negotiations have been influenced by the IPCC assessments over decades already, leading inter alia to the Paris Agreement. The IPCC furthermore managed to evolve over time and to adapt to the changing political contexts, as well as to improve its structure, procedures and methodologies. As a large-scale deliberative learning platform, the IPCC is virtually unprecedented. In terms of the large scale and societal relevance of the deliberation process, it can perhaps be compared with the process leading to the adoption of the Human Rights Declaration after World War II, or with Catholic Ecumenical Councils such as the Second Vatican Council (1962–1965). Notwithstanding the need for continuous IPCC reform, some of the critics of the IPCC in the past may have overlooked its remarkable value for contributing to the facilitation of international deliberative policy learning to some extent, and for linking the IPCC to deliberation among publics around the world.

## Conclusion: the promise of assessment-making

This study has provided: (1) a normative idea for what the ultimate goal of SPIs for longer-term, wicked policy issues could be: contributing to the facilitation of deliberative policy learning; (2) a subsequent criteria-based discussion of alternative SPIs regarding their theoretical potential to achieve this ultimate goal, finding that PEM-inspired assessments seem to have relatively high potential; and (3) some anecdotal evidence that large-scale integrated scientific assessment processes are actually contributing to deliberative policy learning overall. Although the focus in this article is on environment-related global policy issues and assessments, the promises of assessment-making as a key tool at the SPI are certainly also relevant for policy fields and scientific advice beyond the environmental realm (see also Kowarsch, 2016a).

These results, however, do not show that assessments are *always and necessarily* better than other SPIs responding to longer-term, wicked policy issues. Moreover, these other SPIs are sometimes definitely needed as well in these contexts. One reason is that, despite the long time horizon of these wicked issues, there can be short-term windows of opportunity for political action, more specific pressing issues requiring rapid response from scientific experts and so on, for which time-consuming assessments might not be the first choice. Because they can enable later larger-scale assessments of alternative policy pathways, there is furthermore an important role for individual studies that explore only a particular policy pathway in depth.

However, when measured against alternative SPIs, well-designed large-scale integrated scientific assessments are highly promising and perhaps indispensable SPIs to bridge scientific expertise and policy-making on longer-term, wicked policy problems such as the difficult achievement of the SDGs. Such assessments are potentially more comprehensive, more integrated, and—through the inclusion of divergent viewpoints and diverse stakeholders—more politically legitimate. In many policy cases, it's very likely that there would actually be sufficient time to organize a broader deliberation process despite what some might have said. For example, climate policy has already been debated for three decades.

How do assessments (and other SPIs) actually manage to achieve considerable influence on policy? Even if assessments realize the major building blocks of deliberative policy learning, other factors may still be highly relevant as well, including in particular political constraints (for example, described by Cairney *et al.*, 2016). In fact, some governments and other stakeholders do not always have much sympathy for an open, public and critical scientific exploration of policy alternatives, since they want to primarily protect their particular interests (e.g. Edenhofer and Minx, 2014). However, large-scale assessments can perhaps better cope with this challenge than other SPIs. If a critical mass of (diverse) actors becomes involved in a large-scale, mandated assessment process, it may become hard for particular interest groups to avoid any engagement in an open deliberation process, that is, in the open exchange of scientifically informed arguments (Sabatier, 2007). This helps to unfold and scrutinize established political beliefs and "information shortcuts" (Cairney *et al.*, 2016). Many large-scale assessment bodies are intergovernmental organizations with mandates from governing bodies which further helps to facilitate buy-in from governments regarding assessment results. Finally, if the publicly accessible output of an assessment process is a reasonable cartography of alternative policy pathways and their implications, this can influence broader public discourses and, with it, put some pressure on governments to also accept inconvenient insights. Future research should examine the more precise conditions under which policymakers and particular interest groups accept an open exploration of policy alternatives particularly in intergovernmental, "hybrid" SPIs, and the extent to which particular policy narratives of particular stakeholder groups change during the course of an assessment. While this paper mainly focused on the science-policy nexus, future research should also more thoroughly address the various roles of and multiple interactions with societies and different publics regarding different SPIs. We have argued above that deliberative design of large-scale assessments provides the opportunity to bring scientific expertise, policy-making processes and the broader society better together.

Another considerable political challenge for those large-scale assessments that address *global* policy issues (such as climate change or the SDGs) is that politically, to varying degrees depending on the specific policy issue, these issues are also, or even primarily, addressed at national and local levels in diverse contexts. Large-scale assessments usually have not much to say about these. While this multi-scale issue cannot be

satisfactorily discussed in a few lines here (for a deeper discussion see, for example, Berkes *et al.*, 2006 and Cash *et al.*, 2006), we believe that this challenge does not make global assessments meaningless for the national and local level. To mention just one example, global assessments may facilitate the global dissemination of particularly interesting national and local policy lessons that are also highly relevant to some other contexts.

Related to this is another important issue, namely the various necessary but sometimes ineffective links between national structures of scientific expertise and international assessment bodies (or other international SPIs). Again, this cannot be discussed here, but would ideally be part of a future research agenda.

Also in light of these various limitations and open questions, evaluating science-policy interfaces, including large-scale assessments, in a rigorous, consistent and collaborative manner is absolutely crucial to improve processes of evidence-informed policy making. This requires effective, perhaps formalized processes to facilitate the uptake of lessons learned regarding past processes of evidence-informed policy making to inform the design of future science-policy interactions. It also requires processes to facilitate more explicit and systematic, multi-stakeholder critical discussions about the different potential methods of evaluation (and about the underlying different normative understandings of both policy processes and scientific knowledge production). A number of alternative evaluation criteria and guidelines for SPIs exists (for example, OECD, 2015; Cash *et al.*, 2003). Many of them focus on more specific, practical issues and might thus be complementary to our approach. The "metrics of success" of SPIs, that is, the measurements of actual or envisaged impacts of scientific advisory bodies on policy processes etc., or their enabling conditions, must, however, go beyond the overly narrow and unrealistic (yet still predominant) assumption of SPIs directly influencing policy decisions (see also Cash *et al.*, 2003; Rayner, 2003; Mitchell *et al.*, 2006). Such criteria should rather emphasize the diverse, indirect yet valuable influences that SPIs can have on broader policy discourses including, for instance, agenda-setting and the diffusion of programmatic policy ideas (such as, for example, Emissions Trading Schemes).

A feasible and desirable outcome for SPIs in light of our deliberative criteria may thus be confronting and enriching policy discourses, rather than seeking to directly determine policy decisions. Assessments can substantially contribute to such deliberative policy learning and thus indirectly help facilitate "learning democracy" (Ansell, 2011). Against a backdrop of current isolationist nationalism in some countries and the wicked policy challenges and crises of the 21st century, there are compelling reasons to bolster global cooperation based on inclusive, transnational deliberative policy learning (including participation of those disappointed by the political establishment). The impressive milestones recently achieved in international governance, such as the Paris Agreement and the signing of the SDGs, along with bottom-up democracy movements around the world, should encourage greater faith in multilateral cooperation and collaborative deliberation. International solidarity and societal coherence can possibly be strengthened by anchoring policy proposals within deliberatively democratic processes. Thus, the laborious deliberatively democratic science-policy approaches may be the worst responses to such overlapping and interconnected challenges, apart from all the others. Actively "to participate in the making of knowledge is the highest prerogative of man and the only warrant of his freedom" (Dewey, quoted in Brown, 2009: 135).

## Notes

1 Although there is no consensus on the more precise meaning of complexity in public policy, this term basically refers to policy phenomena emerging from (but being greater than the sum of) many interacting elements, while being highly sensitive to initial conditions and typically evolving in a non-linear manner, adapting to changing environments (OECD, 2009; Cairney, 2012). Furthermore, while complex problems are not necessarily complicated, wicked policy problems certainly are complicated.

2 For a more detailed characterization see Rittel and Webber (1973); Carley and Christie (2000: 156); Head (2008).

3 For more information about the IPCC's organizational structure, processes and history see http://www.ipcc.ch/organization/organization_structure.shtml (accessed 20 October 2016).

4 Also the MA, IAASTD and IPBES, for instance, are more or less in line with the particular PEM claim to highlight and explore divergent viewpoints in assessments (although these assessments are obviously are not explicitly following the PEM model which was developed more recently). See also the section on the different science-policy interfaces above.

5 Following the PEM idea, conflicts could perhaps be more resolvable (and less divisive or ideological) when assessing and comparing concrete future scenarios based on alternative disputed normative ideas, rather than potentially endless disputes about the abstract normative ideas themselves (Kowarsch and Edenhofer, 2016). A key tool of policy debates and policy learning alike are policy narratives to provide a simplified orientation of complex value-laden policy issues (Shanahan *et al.*, 2011; Urhammer and Røpke, 2013). Assessments of alternative policy pathways and their implications could provide the necessary evidence base and capacity to facilitate learning, that is, changes, regarding the policy narratives held by different stakeholder groups.

6 Through deliberation, principles and values which are often implicitly used to judge policy options can be made apparent in assessments. This in turn makes it possible to translate them into more specific and concrete metrics or indicators for explicitly and critically evaluating the acceptability of different policy pathways or policy options (Kowarsch and Edenhofer, 2016). Earlier IPCC assessments, for example, strove to avoid normative-ethical discussions entirely. However, disputed normative assumptions are often "hidden" in policy-relevant scientific findings anyway (Biewald *et al.*, 2015). In its recent assessment, however, the IPCC Working Group III added a chapter explicitly discussing ethical issues (IPCC, 2014a, Chapter 3), which is a remarkable deliberative achievement in our view.

7 The interviews were mainly conducted to inform the research initiative on "The Future of Global Environmental Assessment Making" (http://www.mcc-berlin.net/en/research/cooperation/unep.html, accessed 11 November 2016). In addition, our involvement in recent IPCC and UNEP GEO processes as well as in several multi-stakeholder workshops on assessment-making provided us with opportunities to collect information on how governments and scientific experts perceive the deliberative value of assessment-making. Given the rigorous interview methodology applied according to the state of the art, our involvement in recent assessment processes does not make our results considerably biased. The claims about the overall value of the IPCC assessments further below are, however, more opinionated (as made transparent there).

8 See her video briefing to the IPCC Plenary in April 2016 (min 0:56), available at https://www.youtube.com/watch?v=lZpTr5kOpGU (accessed 15 May 2016).

## References

Aitsi-Selmi A, Blanchard K and Murray V (2016) Ensuring science is useful, usable and used in global disaster risk reduction and sustainable development: a view through the Sendai framework lens. *Palgrave Communications*; **2** (16016), doi: 10.1057/palcomms.2016.16.

Ansell C K (2011) *Pragmatist Democracy: Evolutionary Learning as Public Philosophy*. Oxford University Press: Oxford.

Arrow K J (1970) *Social Choice and Individual Values. 2nd edition*. Yale Univ Press: New Haven, CT.

Bartlett R V and Baber W F (1999) From rationality to reasonableness in environmental administration: moving beyond proverbs. *Journal of Management History*; **5** (2): 55–67.

Baber W F and Bartlett R V (2001) Toward environmental democracy. *The Kansas Journal of Law and Public Policy*; **11** (1): 35–64.

Bennett E, Zurek E (2006) Integrating epistemologies through scenarios In: Reid W, Berkes F, Wilbanks T and Capistrano D (eds.) *Bridging Scales and Epistemologies: Linking Local Knowledge and Global Science in Environmental Assessments*. Island Press: Washington DC, pp 275–293.

Berkes F, Reid W V, Wilbanks T J, Capistrano D (2006) Conclusions In: Reid W V, Berkes F, Wilbanks T J and Capistrano D (eds) *Bridging Scales and Knowledge Systems: Concepts and Applications in Ecosystem Assessment*. Island Press: Washington DC, pp 315–331.

Biewald A, Kowarsch M, Lotze-Campen H and Gerten D (2015) Ethical aspects in the economic modeling of water policy options. *Global Environmental Change*; **30** (1): 80–91.

Brown M B (2009) *Science in democracy: Expertise, institutions, and representation*. MIT Press: Cambridge, MA.

Cairney P (2012) Complexity theory in political science and public policy. *Political Studies Review*; **10** (3): 346–358.
Cairney P, Oliver K and Wellstead A (2016) To bridge the divide between evidence and policy: Reduce ambiguity as much as uncertainty. *Public Administration Review*; **76** (3): 399–402.
Carley M and Christie I (2000) *Managing Sustainable Development*. Earthscan: London.
Carraro C, Edenhofer O, Flachsland C, Kolstad C, Stavins R and Stowe R (2015) The IPCC at a crossroads: Opportunities for reform. *Science*; **350**, 34f.
Cash D W *et al* (2003) Knowledge systems for sustainable development. *PNAS*; **100** (14): 8086–8091.
Cash D W *et al* (2006) Scale and cross-scale dynamics: Governance and information in a multilevel world. *Ecology and Society*; **11** (2): 8.
Cashmore M and Richardson T (2013) Power and environmental assessment: Introduction to the special issue. *Environmental Impact Assessment Review*; **39**, 1–4.
Chambers S (1996) *Reasonable Democracy: Jürgen Habermas and the Politics of Discourse*. Cornell University Press: Ithaca, NY.
Chambers S (2003) Deliberative democratic theory. *Annual Review of Political Science*; **6** (1): 307–326.
Christiano T (2012) Rational deliberation among experts and citizens In: Parkinson J and Mansbridge J (eds) *Deliberative Systems: Deliberative Democracy at the Large Scale*. Cambridge University Press: Cambridge, UK, pp 27–51.
Clark W C, Tomich T P, van Noordwijk M, Guston D, Catacutan D, Dickson N M and McNie E (2011) Boundary work for sustainable development: Natural resource management at the Consultative Group on International Agricultural Research (CGIAR) Inaugural Article: Knowledge Systems for Sustainable Development, Special Feature Sackler Colloquium. PNAS 200900231.
Cohen J (1997) Deliberation and democratic legitimacy In: Bohman J a and ndRehg W. (eds.) *Deliberative Democracy: Essays on Reason and Politics*. MIT Press: Cambridge, MA, pp 67–91.
Cole D H (2015) Advantages of a polycentric approach to climate change policy. *Nature Climate Change*; **5** (2): 114–118.
Deitelhoff N (2012) (Is) Fair enough?—Legitimation internationalen Regierens durch deliberative Verfahren In: Niesen P (ed.) *Transnationale Gerechtigkeit und Demokratie*. Frankfurt aM: Campus, pp 103–130.
Dewey J (1927) *The Public and its Problems*. Henry Hold: New York.
Díaz S *et al* (2015) The IPBES conceptual framework—Connecting nature and people. *Current Opinion in Environmental Sustainability*; **14**, 1–16.
Dietz T (2013) Bringing values and deliberation to science communication. *PNAS*; **110** (3): 14081–14087.
Dietz T and Stern P C(NRC) (eds) (2008) *Public Participation in Environmental Assessment and Decision Making*.: National Academy Press: Washington DC.
Dodds F, Laguna-Celis J and Thompson L (2014) *From Rio+20 to a New Development Agenda. Building a bridge to a sustainable future*. Routledge: New York City.
Douglas H E (2009) *Science, Policy, and the Value-Free Ideal*. University of Pittsburgh Press: Pittsburgh, PA.
Dryzek J (1990) *Discursive Democracy: Politics, Policy, and Political Science*. Cambridge University Press: Cambridge, UK.
Dryzek J (1999) Transnational democracy. *Journal of Political Philosophy*; **7** (1): 30–51.
Dryzek J (2000) *Deliberative Democracy and Beyond: liberals, critics, contestations*. Oxford Univ Press: Oxford.
Dryzek J (2010) *Foundations and Frontiers of Deliberative Governance*. Oxford Univ Press: Oxford.
Dunlop C and Radaelli C (2013) Systematising policy learning: from monolith to dimensions. *Political Studies*; **61** (3): 599–619.
Edenhofer O and Minx J C (2014) Mapmakers and navigators, facts and values. *Science*; **345** (6192): 37–38.
Edenhofer O and Kowarsch M (2015) Cartography of pathways: A new model for environmental policy assessments. *Environmental Science & Policy*; **51**, 56–64.
Esteves A M and Vanclay F (2009) Social Development Needs Analysis as a tool for SIA to guide corporate-community investment: Applications in the minerals industry. *Environmental Impact Assessment Review*; **29** (2): 137–145.
Farrell A E, Keating T J (2006) Dissent and Trust in Multilateral Environmental Assessments In: Farrell A E and Jäger J (eds) *Assessments of Regional and Global Environmental Risks. Designing Processes for the Effective Use of Science in Decisionmaking*. RFF Press Book: Danvers, MA, pp 65–83.
Feldman S and Biggs S (2012) The politics of international assessments: The IAASTD process, reception and significance. *Journal of Agrarian Change*; **12** (1): 144–169.
Fouilleux E (2004) CAP reforms and multilateral trade negotiations: Another view on discourse efficiency. *West European Politics*; **27** (2): 235–255.
Foresight. (2010) Land Use Futures Project. Final Project Report. The Government Office for Science, London, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/288843/10-631-land-use-futures.pdf. accessed 27 January 2016.
Funtowicz S O, Ravetz J R (1991) A new scientific methodology for global environmental issues In: Costanza R (ed) *Ecological Economics: The Science and Management of Sustainability*. Columbia University Press: New York, pp 137–152.
Gluckman P (2016) Understanding the challenges and opportunities at the Science-Policy Interface: domestic and global perspectives. Talk at United Nations Environmental Assembly, Science-Policy Forum, Nairobi, Kenya, May 2016, http://www.ingsa.org/wp-content/uploads/2016/05/Gluckman-UNEA-keynote.pdf. accessed 19 May 2016.
Gutmann A and Thompson D (1996) *Democracy and Disagreement*. Harvard University Press: Cambridge, MA.
Grieneseisen M and Zhang M (2011) The current status of climate change research. *Nature Climate Change*; **1** (2): 72–730.
Haas P M (2004) When does power listen to truth? A constructivist approach to the policy process. *Journal of European Public Policy*; **11** (4): 569–592.
Habermas J (1996a) *Die Einbeziehung des Anderen. Studien zur politischen Theorie*. Suhrkamp: Frankfurt am Main, Germany.
Habermas J (1996b) *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*: Polity Press: Cambridge, UK.
Habermas J (1971) *Toward a Rational Society*. Beacon Press: Boston, MA.
Hajer M A (1993) Discourse coalitions and the institutionalisation of practice: The case of acid rain in Britain In: Fischer F and Forester J (eds) *The Argumentative Turn in Policy Analysis and Planning*. UCL Press Limited: London, pp 43–76.
Head B W (2008) Wicked problems in public policy. *Public Policy*; **3** (2): 101–118.
Hotes S and Opgenoorth L. (2014) Trust and control at the science-policy interface in IPBES. *Bioscience*; **64**, 277–278.
Hulme M (2009) *Why we disagree about climate change: Understanding controversy, inaction and opportunity*. Cambridge Univ Press: Cambridge, MA.
Hunter J E and Schmidt F L (1996) Cumulative research knowledge and social policy formulation: The critical role of meta-analysis. *Psychology, Public Policy, and Law*; **2** (2): 324–347.
IPBES (Intergovernmental Platform on Biodiversity and Ecosystem Services). (2013) Catalogue of Assessments on Biodiversity and Ecosystem Services, http://catalog.ipbes.net/ accessed 15 May 2016.
ICSU. (2010) Earth System Science for Global Sustainability: The Grand Challenges. International Council for Science, Paris, http://www.icsu.org/publications/reports-and-reviews/grand-challenges/GrandChallenges_Oct2010.pdf. accessed 15 May 2016.
IPCC. (2013) Collated comments from Governments for the thirty-seventh session of the IPCC on the Future of the IPCC (Batumi, Georgia, 14–18 October 2013). Submitted by the IPCC Secretariat, http://www.ipcc.ch/apps/eventmanager/documents/5/030920131000-INF_1_p37.pdf. accessed 15 May 2016.
IPCC, Edenhofer O *et al* (Eds.) (2014a) *Climate Change 2014: Mitigation of Climate Change: Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press: Cambridge, UK.
IPCC. (2014b) Collated comments from Governments for the thirty-ninth session of the IPCC on the Future of the IPCC (Berlin, Germany, 7–12 April 2014). Submitted by the IPCC Secretariat, https://www.ipcc.ch/apps/eventmanager/documents/11/280220141142-inf1_future_of_ipcc_govt_comments.pdf. accessed 15 May 2016.
Ivanova L (2009) UNEP as Anchor organization InBiermann F, Siebenhüner B and Schreyögg A (eds) *International Organizations in Global Environmental Governance*. Routledge: London.
Jakob M and Steckel J C (2016) Implications of climate change mitigation for sustainable development. *Environmental Research Letters*; **11** (10): 104010.
Jasanoff S (1990) *The Fifth Branch: Science Advisers as Policymakers*. Harvard University Press: Cambridge, MA.
Kohler P, Conliffe A, Gutierrez M, Jungcurt S, Yamineva Y (2012) Informing policy: The role of science and knowledge in global environmental agreements In: Chasek P S and Wagner L M (eds) *The Roads from Rio: Lessons Learned from Twenty Years of Multilateral Environmental Negotiations*. RFF Press, Taylor and Francis: New York, pp 59–81.
Kok K, Biggs R and Zurek M (2007) Methods for developing multiscale participatory scenarios: Insights from Southern Africa and Europe. *Ecology and Society*; **12** (1): 8–23.
Kowarsch M (2016a) Policy assessments to enhance EU scientific advice. *Nature Climate Change*; **6** (1): 15–17.
Kowarsch M (2016b) *A Pragmatist Orientation for the Social Sciences in Climate Policy. How to Make Integrated Economic Assessments Serve Society. Boston Studies in the Philosophy and History of Science 323*. Springer: Switzerland.
Kowarsch M, Edenhofer O (2016) Principles or pathways? Improving the contribution of philosophical ethics to climate policy In: Roser D and Heyward C (eds.) *Climate Justice in a Non-Ideal World*. Oxford Univ Press: Oxford, pp 296–318.
Labbouz B and Treyer S (2010) The mobilization of international expertise for global governance in IAASTD: a failure in consensus or a successful advocacy strategy? Paper presented at the Social Dimensions of Environmental Change and Governance: 2010 Berlin Conference on the Human Dimensions of Global

Environmental Change, http://edocs.fu-berlin.de/docs/servlets/MCRFileNodeServlet/FUDOCS_derivate_000000001300/Labbouz-The_mobilization_of_international_expertise-220.pdf?hosts=. accessed 15 May 2016.
McGann J G, Viden A and Rafferty J (2014) *How think tanks shape social development policies*. University of Pennsylvania Press: Philadelphia, PA.
McKinnon M C, Cheng S H, Garside R, Masuda Y J and Miller D C (2015) Map the evidence. *Nature*; **528** (7581): 185–187.
Mentzel M (1999) Think tanks, policy-making, and a Dutch advisory council. *Science and Public Policy*; **26** (3): 171–178.
Miller CA (2000) The Dynamics of Framing Environmental Values and Policy: Four Models of Societal Processes. *Environmental Values*; **9** (2): 211–233.
Miller C (2003) Knowledge and accountability in global governance In: Tetreault M and Teske R (eds) *Partial truths about the politics of community*. University of South Carolina: Charleston, SC, pp 315–41.
Miller C (2004) Resisting empire: Globalization, relocalization, and the politics of knowledge In: Jasanoff S*Martello M*; (eds) *Earthly politics: Local and global in environmental governance*. MIT Press: Cambridge, MA, pp 81–102.
Miller C, Erickson P (2006) The politics of bridging scales and epistemologies In: Reid W V, Berkes F, Wilbanks T J and Capistrano D (eds) *Bridging Scales and Knowledge Systems: Concepts and Applications in Ecosystem Assessment*. Island Press: Washington DC, pp 297–314.
Mitchell R B, Clark W C, Cash D and Dickson N. (eds.) (2006) *Global Environmental Assessments: Information and Influence*. MIT Press: Cambridge, MA.
Montana J and Borie M (2016) IPBES and biodiversity expertise: Regional, gender, and disciplinary balance in the composition of the interim and 2015 multidisciplinary expert panel. *Conservation Letters*; **9** (2): 138–142.
Niemeyer S (2014) Scaling up deliberation to mass publics: Harnessing mini-publics in a deliberative system InGrönlund K, Bächtiger A and Setälä M (eds.) *Deliberative Mini-Publics: Involving Citizens in the Democratic Process*. ECPR Press: Colchester, UK, pp 177–202.
Norgaard R (2008a) Finding hope in the millennium ecosystem assessment. *Conservation Biology*; **22** (4): 862–869.
Norgaard R (2008b) The implication of interdisciplinary scientific assessments for environmental governance In: Ranganatha J, Munasinghe M and Irwin F (eds) *Policies for Sustainable Governance of of Global Ecosystem Services*. Edward Elgar: Cheltenham, UK.
NRC. (1996) *Understanding risk: Informing decisions in a democratic society*. National Academy Press: Washington DC.
OECD. (2009) Applications of Complexity Science for Public Policy, Paris: OECD Publishing, https://www.oecd.org/science/sci-tech/43891980.pdf. accessed 1 October 2016.
OECD. (2015) Scientific Advice for Policy Making: The Role and Responsibility of Expert Bodies and Individual Scientists. OECD Science, Technology and Industry Policy Papers, No. 21, Paris: OECD Publishing, http://dx.doi.org/10.1787/5js33l1jcpwb-en. accessed 15 May 2016.
O'Faircheallaigh C (2010) Public participation and environmental impact assessment: purposes, implications, and lessons for public policy making. *Environmental Impact Assessment Review*; **30** (1): 19–27.
Patel M, Kok K and Rothman S S (2007) Participatory scenario construction in land use analysis: an insight into the experiences created by stakeholder involvement in the Northern Mediterranean. *Land Use Policy*; **24** (3): 546–561.
Pettit P (2000) Democracy, electoral and contestatory In: Shapiro I and Macedo S (eds) *Designing Democratic Institutions*. New York Univ Press: New York, pp 105–144.
Pielke R A (2007) *The Honest Broker: Making Sense of Science in Policy and Politics*. Cambridge Univ Press: Cambridge, MA.
Pregernig M (2006) Transdisciplinarity viewed from afar: Science-policy assessments as forums for the creation of transdisciplinary knowledge. *Science & Public Policy*; **33** (6): 445–455.
Putnam H (2004) *The Collapse of the Fact/Value Dichotomy and Other Essays*. Harvard University Press: Cambridge MA.
Rayner S (2003) Democracy in the age of assessment: Reflections on the roles of expertise and democracy in public-sector decision making. *Science and Public Policy*; **30** (3): 163–170.
Reed M S (2008) Stakeholder participation for environmental management: a literature review. *Biological Conservation*; **141** (10): 2417–2431.
Reid W V, Berkes F, Wilbanks T J, Capistrano D (2006) Introduction InReid W V, Berkes F, Wilbanks T J a and ndCapistrano D (eds) *Bridging Scales and Knowledge Systems: Concepts and Applications in Ecosystem Assessment*. Island Press: Washington DC, pp 1–17.
Rittel H and Webber M (1973) Dilemmas in a general theory of planning. *Policy Sciences*; **4** (2): 155–169.
Ryfe D M (2005) Does deliberative democracy work? *Annual Review of Political Science*; **8** (1): 49–71.
Sabatier P A (1988) An advocacy coalition framework of policy change and the role of policy-oriented learning therein. *Policy Sciences*; **21** (2–3): 129–168.
Sabatier P A (ed) (2007) *Theories of the Policy Process*, 2nd edn., Westview Press: Boulder, CO.
Sarewitz D (2004) How science makes environmental controversies worse. *Environmental Science & Policy*; **7** (5): 385–403.
Schmidt V A and Radaelli C M (2004) Policy change and discourse in europe: conceptual and methodological issues. *West European Politics*; **27** (2): 183–210.
Schulte-Uebbing L, Hansen G, Hernández A M and Winter M (2015) Chapter scientists in the IPCC AR5—Experiences and lessons learned. *Current Opinion in Environmental Sustainability*; **14**, 250–256.
Schultz W L (2006) The cultural contradictions of managing change: using horizon scanning in an evidence-based policy context. *Foresight*; **8** (4): 3–12.
Selin N E (2006) Applying assessment lessons to new challenges InFarrell A E and Jäger J (eds) *Assessments of Regional and global Environmental Risks. Designing Processes for the Effective Use of Science in Decisionmaking*. RFF Press Book: Danvers, MA, pp 84–100.
Sénit C, Biermann F and Kalfagianni A (2016) The representativeness of global deliberation: A critical assessment of civil society consultations for sustainable development. *Glob Policy*; doi:10.1111/1758-5899.12371.
Shanahan E A, Jones M D and McBeth M K (2011) Policy narratives and policy processes. *The Policy Studies Journal*; **39** (3): 535–561.
Stevenson H and Dryzek J (2012) The legitimacy of multilateral climate governance: A deliberative democratic approach. *Critical Policy Studies*; **6** (1): 1–18.
Stokstad E (2008) Dueling visions for hungry world. *Science*; **319** (5869): 1474–1476.
Strauss A and Corbin J (1998) *The Basics of Qualitative Research*. Sage Publications: Thousand Oaks.
Thaman R, Lyver P, Mpande R, Perez E, Cariño J and Takeuchi K (eds) (2013) The Contribution of Indigenous and Local Knowledge Systems to IPBES: Building Synergies with Science. IPBES Expert Meeting Report, UNESCO/UNU. Paris: UNESCO, http://unesdoc.unesco.org/images/0022/002252/225242e.pdf accessed 15 May 2016.
Tomkins E L, Few R and Brown K (2008) Scenario-based stakeholder engagement: Incorporating stakeholders preferences into coastal planning for climate change. *Journal of Environmental Management*; **88** (4): 1580–1592.
Turnpenny J, Lorenzoni I and Jones M (2009) Noisy and definitely not normal: responding to wicked issues in the environment, energy and health. *Environmental Science & Policy*; **12** (3): 347–358.
Urhammer E and Røpke I (2013) Macroeconomic narratives in a world of crises: An analysis of stories about solving the system crisis. *Ecological Economics*; **96** (6): 62–70.
van Asselt M B A, van't Klooster S A, van Notten P W F and Smits L A (2010) *Foresight in action. Developing policy-oriented scenarios*. Earthscan: London.
van Deveer S D (2006) European politics with a scientific face In: Farrell A E and Jäger J (eds) *Assessments of Regional and global Environmental Risks. Designing Processes for the Effective Use of Science in Decisionmaking*. RFF Press Book: Danvers, MA, pp 84–100.
van Slyke D M, O'Leary R, Kim S (2010) Conclusion: Challenges and opportunities, crosscutting themes, and thoughts on the future of public administration InO'Leary R, Van Slyke D M and Kim S (eds) *The Future of Public Administration Around the World. The Minnowbrook Perspective*. Georgetown Univ Press: Washington DC, pp 281–294.
Victor D (2015) Climate change: Embed the social sciences in climate policy. *Nature*; **520** (7545): 27–29.
von Stechow C *et al* (2016) 2°C and SDGs: United they stand, divided they fall? *Environmental Research Letters*; **11** (3): 034022.
Watson R T (August, 2009) Letter by Robert T. Watson, director of the IAASTD, to Bureau Members and Donors regarding the independent evaluation of the IAASTD commissioned by the World Bank to Howard Elliot, http://www.agassessment-watch.org/docs/Watson%20Response%20to%20final%20IAASTD%20evaluation%20august%202009.pdf.
Whitfield S, Geist H J and Ioris A A R (2011) Deliberative assessment in complex socioecological systems: recommendations for environmental assessment in drylands. *Environmental Monitoring and Assessment*; **183** (1): 465–483.
Wiklund H (2005) In search of arenas for democratic deliberation: a Habermasian review of environmental assessment. *Impact Assessment and Project Appraisal*; **23** (4): 281–292.
Young I M (1993) Justice and communicative democracy InGottlieb R (ed) *Radical Philosophy: Tradition, Counter Tradition, Politics*. Temple Univ Press: Philadelphia, PA, pp 123–143.

## Data availability

The datasets analysed during the current study are not publicly available to protect the anonymity of the interviewees but are available (in anonymized form, as MAX QDA file) from the corresponding author on reasonable request.

## Additional information

**Competing interests:** The Authors declare no competing financial interests.

**Reprints and permission** information is available at http://www.palgrave-journals.com/pal/authors/rights_and_permissions.html

**How to cite this article**: Kowarsch M *et al* (2016) Scientific assessments to facilitate deliberative policy learning. *Palgrave Communications*. 2:16092 doi: 10.1057/palcomms.2016.92.

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/