Mercator Research Institute on
Global Commons and Climate Change
(MCC) gemeinnützige GmbH

What are scientific assessments?

Dr. Martin Kowarsch

MCC Working paper 4/2014

MCC was founded jointly by Stiftung
Mercator and the Potsdam Institute for
Climate Impact Research.

What are scientific assessments?

Martin Kowarsch[1] – MCC working paper, draft version 24 Nov 2014

The international community has mandated and supported a number of widely known, large-scale assessments of complex environmental issues in recent years (see IPBES, 2013, for an overview), such as the assessments by the Intergovernmental Panel on Climate Change (IPCC) and the United Nations Environment Programme's (UNEP) Global Environment Outlook (GEO) series. Conducting such large-scale assessments requires hundreds of researchers from different disciplines, experts from non-academic institutions, and several years of research and significant financial resources. At present, a diversity of projects at the global as well as at sub-global scales are called "scientific assessments." However, there are divergent understandings of this term, sometimes leading to confusion. In the following, a ***conceptualization of contemporary "scientific assessments" in the context of public policy-making processes*** is suggested to facilitate a consistent, differentiated and constructive discussion about such assessments, and their relationship to other scientific efforts – although there is in fact a broad spectrum of scientific efforts, a thin line between the types of scientific efforts discussed below, and a change in key characteristics of assessments over time. Given these considerations, the three central characteristics of assessment are:

- Intending the provision of somehow **policy-relevant** scientific knowledge in a publicly accessible manner to support public policy-making processes and deliberation. This means to envisage 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), acknowledging the existence and interdependencies of multiple policy objectives and multi-functional policy instruments.
- **Assembling the available scientific knowledge** (and identifying research gaps) in order to provide a rich, interdisciplinary and highly integrated image of the policy-relevant aspects considered. Additionally, and to a greater extent than in literature reviews, peer-reviewed *synthesis* of the available publications and information is required (i) to identify the confidence level that can be associated with the scientific findings in assessments, and (ii) 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.
- Taking into account **different viewpoints** in terms of controversial scientific statements and approaches, uncertainty, but also in terms of disputed societal values and conflicting interests. Besides making areas of disagreement transparent in the assessment outputs, (i) engaging with policy-makers and other stakeholders, as well as (ii) involving a number of authors with various backgrounds, approaches and viewpoints, are typical assessment design elements to realize this. As such, assessments can be regarded as social processes to scientifically discuss policy-relevant issues, which usually facilitate learning among the participants. Assessments are not advocacy pieces.

[1] MCC working group "Assessment and Scientific Policy Advice." Contact: kowarsch@mcc-berlin.net.

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**. Pre-assessments are scientific studies that (i) address crucial research gaps relevant to specific assessment projects, for example socio-economic research papers on public policy analysis for assessments, and (ii) that offer a better understanding and overview of the related literature and political or societal controversies to some extent, in order to provide some of the well-structured research material and overview needed for successful scientific assessments. Pre-assessments can alternatively consist of research studies and literature reviews on assessment methodology and conceptual frameworks, assessment design options, and impact strategies. Though *pre*-assessments should not be regarded as full assessments themselves, they can considerably support and facilitate full assessments.

The **added value** of scientific assessments – for instance compared to standard scientific policy analyses – is that they are potentially more comprehensive, more integrated and interdisciplinary, and – through the inclusion of divergent viewpoints and diverse stakeholders – relatively more legitimate. In this sense, scientific assessments ideally bridge the gap from assembling scientific knowledge to managing the complexity of societal challenges by facilitating learning about these policy-relevant issues. This ambition, however, can only be realized if assessments are appropriately designed with regard to the considerable challenges and trade-offs they face (Cash et al., 2003).

For this reason, going beyond the above descriptive conceptualization, the **pragmatic-enlightened model** (PEM; see Kowarsch and Edenhofer, 2014, for more detail) provides some **normative guidance** on assessment design, at least for large-scale assessments of complex policy fields that face high uncertainty and disputed value-laden viewpoints. For policy evaluations, the PEM assumes the interdependency of policy objectives, means and their implications. Key claims of the PEM 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 may avoid policy-prescription, while still allowing for learning about policy options. Though assessment always implies value judgments, and though co-production of knowledge with stakeholders is envisaged, reliable and objective scientific knowledge in assessments remains possible and desirable according to the PEM. Assessments should therefore be strictly based on rigorous (and wherever possible, peer-reviewed) scientific research.

A **new culture in academia** is needed that leads to the provision of the scientific material necessary for a successful assessment of policy options, including pre-assessments. Moreover, the onerous and intellectually challenging synthesis effort in assessments should be accepted as a fully respectable, rigorous and societally very useful scientific task in and of itself. Contributing to both assessments and to rigorous scientific research need not be mutually exclusive, because assessments are dependent on the availability of high-quality, peer-reviewed scientific research, and because the synthesis exercise as such – i.e. the creation of maps of knowledge for and jointly with policymakers – must be increasingly seen as highly challenging and genuinely scientific work as well.

	Research paper	Pre-assessment	Smaller-scale assessment	Larger-scale assessment
Examples	Peer-reviewed journal articles, e.g., on climate economics	Transport & ETS;[2] ADAM;[3] commissioned papers;[4] by WHO;[5] by CBD[6]	Stern Review (2007); UK Foresight;[7] UN Emissions Gap Report;[8] HDR[9]	IPCC ARs; UNEP's GEO series; GBO; IAASTD. See IPBES (2013) for details and acronyms
Political mandate	Very rarely	Rarely	Very often	Always
Scope of literature used	Only directly relevant literature	Relevant literature (and overview of existing lit.)	Synthesis of much of the relevant literature	Almost comprehensive literature synthesis
Policy-relevant synthesis and integration	Usually not	To some extent	High degree of synthesis and integration	High degree of synthesis and integration
Inclusion of divergent viewpoints (uncertainty, values, etc.)	Usually only in order to clarify own viewpoint	Overview of divergent viewpoints; in addition, possibly a few scenarios	Overview, and exploration of scenarios based on some crucial divergent viewpoints	Exploration of all highly relevant divergent viewpoints (e.g., through scenarios)
Geographical scales and/or governance levels considered	Usually 1 or 2	1 to few	A few; analysis of their interdependency	Several; analysis of their interdependency
Systemic effects and conditions[10] considered	1 to few	Few	Several key ones	All of the most relevant ones
Basic policy objectives and evaluation criteria considered	Usually 1 to 2 (at least implicitly; e.g., avoiding climate change)	Usually 1 to few (transparent)	Several key ones (transparent)	All of the most relevant ones (transparent)
Policy options (i.e., policy means and sets of objectives) explored	Usually 1 to 2 (if policy analysis is done)	A few key ones (if policy analysis is done)	Some key ones (if policy analysis is done)	Several (if policy analysis is done)

[2] See, e.g., https://ideas.repec.org/p/ecc/wpaper/2.html.
[3] See https://www.pik-potsdam.de/research/sustainable-solutions/flagshipspld/MitigationScenarios/adam/adam-project.
[4] E.g., https://www.gov.uk/government/collections/reducing-risk-of-future-disasters#supporting-evidence.
[5] See http://www.who.int/hia/health_indicators/en/.
[6] See http://www.cbd.int/doc/health/guide-biodiversity-health-en.pdf.
[7] See https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/286476/12-1289-reducing-risks-of-future-disasters-report.pdf.
[8] See http://www.unep.org/publications/ebooks/emissionsgapreport/.
[9] See, e.g., http://hdr.undp.org/sites/default/files/hdr14-report-en-1.pdf.
[10] For instance, different assumptions about the future availability of low-carbon energy technologies.

Analyzed practical implications of policy means	1 to few (if policy analysis is done)	A few (if policy analysis is done)	Some (if policy analysis is done)	Several (if policy analysis is done)
Transparency of confidence level and agreement	Some	Some	High	Very high
Number of scientific experts and disciplines	1 to few (often only one discipline)	A few (sometimes interdisciplinary)	Many (often >15 authors and >2 institutions; often interdisciplinary)	Very high number (usually >100 authors, often interdisciplinary)
Process of author selection	No specific process	Self-selected, or nominated (if commissioned papers, for example)	Often formal criteria for author selection to better represent divergent viewpoints and approaches	Highly formalized process; often many divergent perspectives and regional representation envisaged
Governance structure	Informal	Mostly informal	Often formalized	Highly formalized: committees, procedures; confl. of interest policy
Stakeholder engagement (beyond scientists)	Usually not necessary	Often only little (e.g. bilateral conversations)	Some (e.g., workshops with different stakeholder groups)	Extensive (multiple formats, many groups – at least more recently)
Review process	Double blind-peer review	Double blind-peer review; often additional expert reviews	Bigger group of external reviewers (incl. policy-makers & stakeholders)	Large-scale, formalized, multi-stage review process (incl. policy-makers & stakeholders)
Typical duration of production process	Often <1 yr.	Often 1 to 2 yrs.	Often >2 yrs.	Often 4 to 7 yrs.
Typical output	Article in peer-reviewed scientific journal	Research article(s), or scientific report	Journal special issue, ed. Volume, or scientific report	Assessment report, including targeted summaries (SPM)
Outreach activities and communication	Usually none	Little	High	Very high

Table 1: Differences between assessments and other scientific efforts*. This table focuses on the key differences between contemporary (i) typical research papers, e.g. on policy analysis, (ii) pre-assessments, (iii) smaller-scale assessments, (iv) and larger-scale assessments. The table lists the minimum requirements for the respective type of scientific work. These types are interlinked, because assessments necessarily have to build on research papers, and pre-assessments – which are only rarely produced thus far – can help a lot to facilitate assessments, given the increasing complexity and amount of literature assessments have to deal with. The identification of research gaps in assessment processes again can result in new research papers. To simplify matters, the table does not comprehensively cover the full range of scientific activities and products.*[11] *Note that there is often a thin line between the types of scientific efforts, rather than a strict and clear divide.*

[11] For instance, it does not include research on assessment methodology which can also be regarded as "pre-assessment."

ANNEX A: Selected conceptualizations of "scientific assessment" in the literature

According to Mitchell et al. (2006, p. 3; slightly adapted), scientific assessments in public policy contexts can be understood as *formal efforts by a number of experts and stakeholders to assemble and synthesize the knowledge in a particular research field with a view toward making it publicly available in a form that is intended to be useful for policy-making*.

The full original quotation reads as follows:

> "We define 'assessments' as formal efforts to assemble selected knowledge with a view toward making it publicly available in a form intended to be useful for decision making. By 'formal,' we mean that an assessment is sufficiently organized that such aspects as products, participants and issuing authority can be identified relatively easy. By 'efforts to assemble selected knowledge,' we seek to recognize that assessments vary both with respect to how comprehensive they are and whether they involve conducting new, or summarizing and evaluating existing, research. We interpret 'knowledge' broadly, treating the question of which kinds of information or expertise a specific assessment chooses to incorporate as an empirical rather than definitional one. We emphasize 'publicly available' to distinguish assessments from technical advice prepared for the private use of decision makers. Finally, we use 'decision makers' to encompass actors in government, private corporations, research laboratories, nongovernmental organizations (NGOs), and civil society more generally."

The following table is taken from Jasanoff (1990, p. 80, adapted):

	Regulatory science	Research science
Goals	"Truths" relevant to policy	"Truths" of originality and significance
Institutions	Government, industry	Universities
Products	Studies and data analyses, often unpublished	Published papers
Incentives	Compliance with legal requirements	Professional recognition and advancement
Time-frame	Statutory timetables, political pressure	Open-ended
Options	Acceptance of evidence; rejection of evidence	Acceptance of evidence; rejection of evidence; waiting for more data
Accountability: Institutions	Congress; courts; media	Professional peers
Accountability: Procedures	Audits and site visits; regulatory peer review; judicial review; legislative oversight	Peer review, formal and informal
Accountability: Standards	Absence of fraud or misrepresentation; conformity to approved protocols and agency guidelines; legal tests of sufficiency (e.g., substantial evidence, preponderance of the evidence)	Absence of fraud or misrepresentation; conformity to methods accepted by peer scientists; statistical significance

From the UNEP-hosted International Assessment of Agricultural Knowledge, Science and Technology for Development website (http://www.unep.org/dewa/agassessment/docs/assessmentdef.doc):

What is an Assessment?

- An assessment is a critical evaluation of information for purposes of guiding decisions on a complex, public issue. The stakeholders, who are typically decision-makers, define the topic.
- Assessments are policy relevant, but not prescriptive.
- Assessments are conducted by a credible group of experts with a broad range of disciplinary and geographical experience, in a balanced transparent way.
- Assessments *reduce* complexity by summarization, synthesis and sorting what is known and widely accepted from what is not known (or not agreed).
- Assessments relate to the situation at a particular time and in a geographical domain.

An assessment is not…

A research project
- Most data should already be collected, peer-reviewed and in the public domain
- Gap filling, research on assessment, new runs of old modules and synthesis are permissible

A review paper
- Focused on policy
- Judgment, clearly labeled as such, is required

An advocacy piece
- Must be balanced and evidence-based

It is not an opportunity to promote your pet topics or own work
It is not a vehicle to develop a research agenda

Assessment Characteristics

- Involve all stakeholders in the preparation and peer-review – ownership of the process and results is essential
- Conducted according to an open, transparent, representative and legitimate process
- Findings to be policy relevant, not policy prescriptive; use "if…then" approach
- Technically accurate
- Incorporate different views
- Take a local, regional and global perspective
- Risk assessment, management and communication

Assessment vs Review

	Review	**Assessment**
Audience	Scientists	Decision-makers
Done by	One or a few	Large and varied group
Topic	Simple and narrow	Broad and complex
Identifies gaps in	Research: curiosity-driven	Knowledge for implementation: problem-driven
(Un) Certainty statements	Not required	Essential
Judgement	Hidden	Required but clearly flagged
Coverage	Exhaustive, historical	Sufficient to deal with main range of uncertainty
Synthesis	Not required	Essential to reduce complexity

From UNEP Division of Early Warning and Assessment website (http://www.unep.org/dewa/assessments):

Assessments bring together diverse strands of knowledge in a way that is useful for decision making. Assessments are key mechanisms through which science informs decision making.

Assessments are fundamentally communication processes, not simply reports, which share many similar features, regardless of their scope. The following characteristics define an environmental assessment

- It is a critical, peer-reviewed evaluation of information, for purposes of guiding decisions on a complex public issue, following a well-defined process.
- The scope (topic) is defined by multiple stakeholders, typically decision makers. Findings are policy-relevant but not prescriptive.
- Conducted by a credible group of experts with a broad range of disciplinary and geographical expertise, in a balanced and transparent manner.

DEWA also undertakes extensive review of the latest science on specific issues to identify consensus by sorting out what is known and widely accepted from what is known and not agreed.

ANNEX B: References and acknowledgements

References

Cash, David W., William C. Clark, Frank Alcock, Nancy M. Dickson, Noelle Eckley, David H. Guston, Jill Jäger, and Ronald B. Mitchell. 2003. Knowledge systems for sustainable development. *PNAS* 100(14): 8086–8091.

Edenhofer, O., Kowarsch, M. (equal contribution), 2014. Cartography of policy paths: A model for solution-oriented environmental assessments. MCC Working Paper. Online: http://www.mcc-berlin.net/fileadmin/data/pdf/Publikationen/Edenhofer-Kowarsch_PEM_article.pdf.

IPBES (Intergovernmental Platform on Biodiversity and Ecosystem Services), 2013. Catalogue of Assessments on Biodiversity and Ecosystem Services. Online: http://catalog.ipbes.net/.

Jasanoff, S. 1990. *The Fifth Branch: Science Advisers as Policymakers*. Cambridge, Mass.: Harvard University Press.

Mitchell, R.B., Clark, W.C., Cash, D.W., Dickson, N.M. (Eds.), 2006. Global Environmental Assessments: Information and Influence. MIT Press, Cambridge, MA.

Stern, N.. 2007. *The economics of climate change. The Stern review*. Cambridge: Cambridge University Press.

Acknowledgements

This paper builds upon the many useful comments I graciously received from colleagues at MCC and PIK, in particular those from Jennifer Garard and Christian Flachsland.

ANNEX C: MCC contribution to scientific assessments

- The Mercator Research Institute on Global Commons and Climate Change (MCC) is planning to initiate own scientific assessments, such as the research project on emission trading schemes (ETS) as a climate policy instrument.[12]
- Moreover, MCC staff are involved in the production of a number of assessment reports together with other institutions. For example, Ottmar Edenhofer, in his capacity as co-chair of IPCC WG III, coordinated the development of the WG III contribution to the Fifth Assessment Report (AR5). MCC researchers contributing as authors and experts to the IPCC WG III AR5 are Felix Creutzig, Michael Jakob, Jan Steckel, Christoph von Stechow and Christian Flachsland. Blanca Fernandez and Felix Creutzig will moreover be co-authoring the Second Assessment Report of the Urban Climate Change Research Network. They examine how climate protection in cities impacts equality and how an eco-friendly urban design may affect land rents. Within the Euro-CASE Energy Platform MCC researchers in cooperation with the Potsdam Institute for Climate Impact Research (PIK) seek to identify new paths to reforms of the European carbon trading system EU ETS. They examine lessons-learned, its main current and future challenges, and how the system influences, and is influenced by, other government policies, such as the promotion of renewable energies. Another recent example of MCC involvement in the production of assessment reports is the Institute's cooperation with The New Climate Economy. The Commission, for which MCC Director Ottmar Edenhofer participates in the Economic Advisory Panel, has the aim to identify and communicate the economic benefits and costs of reducing greenhouse gas emissions.
- Because MCC is so much involved in assessment making, MCC also conducts research on how to improve assessment making, for instance in a collaborative research initiative with UNEP.[13]

[12] For an intermediate output of the ongoing research, see http://www.euro-case.org/images/stories/pdf/position-paper/Euro-CASE-policy-paper-ETS-reform.pdf.
[13] For more information, see, for instance, http://www.mcc-berlin.net/en/research/working-groups/assessments-and-scientific-policy-advice.html and http://www.mcc-berlin.net/en/research/cooperation/unep.html.