



A review of *Discerning Experts*

Gary W. Yohe¹

Received: 4 April 2019 / Accepted: 23 May 2019 / Published online: 12 June 2019
© The Author(s) 2019

Discerning Experts (Oppenheimer et al. 2019) is a meticulously researched study of the evolution of decision-makers' commissioning assessments of various forms to inform their considerations of what they should or could do. It relies on carefully sculpted interviews as well as thorough historical analyses of three diverse case studies through 2014. This reliance on three 'points of data' may appear limiting, but the conclusions described here ring true to my ears—ears that have participated in many climate and environmental assessments over the past three decades.

Put another way, Oppenheimer et al. (2019) have produced what is essentially an assessment of assessments, and so its syntheses of insights across the case-study chapters add new knowledge even while they validate old knowledge that had formerly been largely anecdotal for most of us. This volume is therefore essential reading for participants of any large environmental assessment. Past participants looking for a refresher course will see themselves in the pages, just like I did. Young scholars invited to participate in their first assessment will come to understand what will be expected of them. And readers of *Climatic Change* will gain a glimpse of 'how the sausage' that they read about in these pages may have been made.

My review will refer to other assessments and my own experiences. But, here is my lead.

I agree with Chris Field (his comments from the cover of Oppenheimer et al. (2019)) on the basis of my experiences before and after the 2014 threshold date of the volume. Professor Field wrote: 'As a "first" study of the internal workings of large environmental assessments, this book reveals their strengths and weakness, and explains what assessments can – and cannot – be expected to contribute to public policy and the common good (my emphasis of "first").'

Discerning Experts is not exactly the 'first' comparative study of multiple assessments, but it is an important one because of the extraordinary quality of its documentation and analysis as well as its clever creation of critical and instructive diversity across its three case studies.

The first section responds to the introduction in Chapter 1 and its roots from many centuries ago. The second section offers some commentary on the case studies that

✉ Gary W. Yohe
gyohe@wesleyan.edu; <https://www.gyohe.faculty.wesleyan.edu>

¹ Huffington Foundation Professor of Economics, and Environmental Studies, Wesleyan University, Middletown, CT 06459, USA

occupy Chapters 2, 3 and 4—acid rain, the ozone hole and the West Antarctic Ice Sheet (the WAIS), respectively. It should be noted that only Chapter 4 (WAIS) covers an assessment process that was active in this century. It is joined by Chapter 3 on ozone as a topic that could be directly related to what might commonly be thought to be climate change. However, Chapter 2, on acidic precipitation, does not. It covers an environmental issue of a fundamentally different character. It focuses on a relatively short-term externality for which impacts are local or at most regional and for which solutions can be effectively implemented in relatively short order. The ozone hole discussion in Chapter 3 extends the time frame a bit, but a solution does exist; banning fluorocarbons has made the hole shrink—slowly but steadily. As a result, the associated risks will, eventually, become negligible. The WAIS chapter is an example of fundamental long-term climate change for which a ‘fix’ cannot be anticipated and for which damages and risks are both global and irreversible.

The third section considers the synthetic nature of Chapters 5 and 6—the ‘policy boundary’ and what do ‘assessments do’ topics. In making my case, here, I will offer some personal insight into how well the content and analysis in the text applies more broadly beyond the three case studies from which they are drawn. It is here where many expanded experiences come into play to illustrate the incredible scope of the volume. Finally, a last section recasts my lead and expands the suggested readership.

1 Section 1: the introduction

In the introduction (Chapter 1), the authors do some historical excavation. They go back to Northern Europe in the fourteenth century where “divinely inspired prophets merited veneration, but false saints and demoniacs demanded condemnation (pg2)”. It became essential, back then, to be able to tell one from the other; and so Brigit of Sweden became a test case for the value of an orderly assessment by “independent” and preeminent ecclesiastical ‘experts’. She was ultimately canonized. Catherine of Siena was examined similarly in Southern Europe, but she did not survive; she died of ‘self-inflicted’ starvation in 1830 before her canonization. From the very beginning, therefore, it was clear that the stakes for assessments could be enormous. It follows that an examination of historic and best practices is more than a worthwhile exercise.

As fascinating as these historical antecedents are, the real value of Chapter 1 is its analysis of the evolution of independent decision support processes all the way through the release of the 5th Assessment (AR5) of the Intergovernmental Panel on Climate Change Intergovernmental Panel on Climate Change (IPCC) 2014a, b, c, d). Readers should also know that the data that support this analysis extend beyond 2014 because the primary authors have been fixtures in the IPCC for decades and continue to contribute their time and expertise (for free) to IPCC projects to this day.

The historical coverage in this volume leads the reader to understand modern-day institutions’ reliance on a series of assessments to inform their actions and intentions of their ‘clients’. For example:

1. IPCC assessments every 6 years or so with ad hoc special reports in the interim
2. US National Climate Assessment (NCA) every 4 years by an act of Congress with a commitment for intermediate and ongoing assessments

3. Up to 200 mostly smaller assessments every year from the National Research Council (NRC) in direct response to demand from public and/or private clients for something authoritative
4. Three iterations over the past 10 years of the New York (City) Panel on Climate Change (Blake et al. 2019)

While the US experience is the focus of this book, parallel assessments are a way of life across the globe. For example, National Climate Action Plans communicated to the UN Framework on Climate Change numbered in excess of 125 leading up to the Paris Accord negotiations (WRI 2015). Each was the product of a domestic assessment of climate risk; most highlighted mitigation and transparency plans, but many also added some detail to their adaptation plans.

The authors' historical perspective informs understanding of more recent assessments; that is to say, it informs understanding of not only historical assessments but also more recent assessments (at least in climate change—my particular venue). And, it also informs what can be expected in the future. These are, of course, the fundamental points of the introduction; and they will weave their way through the text of my review.

2 Section 2: the case studies

The authors chose to illuminate the value and pitfalls of the assessment process over time. To support their insights, they provided three diverse and superbly researched and instructive case studies:

1. *The National Acid Precipitation Assessment Program:*

The National Acid Precipitation Assessment Program (NAPAP from 1980 through 1991) is described in Chapter 2; it was a program designed to help understand how acid precipitation generated by sulphur emissions damaged lakes and forests and human structures. The short story is that resistance to responding at all produced claims that the science was riddled with too much uncertainty to support action, and this perspective persisted across a decade or so. Perhaps by design, therefore, the Program was exploited by decision-makers to prevent action on acid rain year after year on the basis of these claims even while they looked like they were actually concerned about the damages. Critics have argued that NAPAP “failed to protect the scientific integrity of the assessment process”, so that the delay in response could be advanced as the only rationale response.

Sulphur emissions were finally limited across the USA by a market-based cap and trade system installed during the G.H.W. Bush Administration, so uncertainty was accommodated by policy design. In other words, no action was taken until after the Reagan Administration finished its eight-year policy reign of tolerating scholarly studies while focusing primarily on uncertainty as a reason not to act; USEPA (2017) records the ultimate product in Title IV of the amended Clean Air Act.

2. *The Ozone Hole over Antarctica:*

Assessments of ozone depletion (from 1974 through 1989 when the policy was enacted and then through 1990 for the last relevant assessment) led more effectively to governmental control of chlorofluorocarbon (CFCs) emissions through the *Montreal Protocol on Substance that Deplete the Ozone Layer*. Concern had been generated that reduced ozone concentrations in the ozone were allowing increased exposure to UV-B radiation and therefore melanoma. The *Protocol* was clearly a product of the Vienna Convention for the

Protection of the Ozone Layer. It came into force in January of 1989 after the DuPont chemical company announced at a meeting of the Vienna Convention that it would stop manufacturing CFCs—tomorrow. “Why?” Because they had developed and patented an environmentally benign substitute and because their engineers were convincing in their argument that the substitute could be produced efficiently to global scale (source: personal experience).

Chapter 3 chronicles this history, starting in 1974 with two pathbreaking papers on chlorine and the stratosphere. Ten US and international assessments, many of which were chaired by Robert Watson, ultimately provided sufficient evidence to support a ban of the offending chemicals; DuPont provided evidence of the feasibility of filling what would be the resulting void on the supply side if a ban were enacted. Interestingly, some major players who had produced the scientific understanding of the chemistry and the measurement of ozone depletion had not been invited to participate in the background assessments because of their activist political statements. Organizers of the assessments worried about the perception that the results would be perceived as biased if these leading scientists had participated. Author omissions, including F. Sherwood Rowland, likely slowed progress toward a global policy, but they did not ultimately stop a global response. Just for reference, Rowland and coauthor Mario Molina won the Nobel Prize in Chemistry in 1995 ostensibly for publishing a 3-page paper in *Nature* in 1974 (Molina and Rowland 1974). Their diploma from the Nobel Committee read “For their work in atmospheric chemistry, particularly the formation and decomposition of ozone” (www.nobelprize.org/prizes/chemistry/1995/summary/).

3. *The West Antarctic Ice Sheet (WAIS):*

Chapter 4 reports on the history (1981–2007) of our scientific understanding of the role of the West Antarctic Ice Sheet (WAIS) in explaining the pace of global sea level rise (SLR). There was concern that melting or disintegration of parts of the WAIS would produce global sea level rise. Progress was slow in making this connection because of the enormity of the Antarctic and the dominance of thermal expansion as an explanation of observed SLR. It was difficult to tell whether or not the ice sheet was growing (reducing the pace of SLR) or shrinking (increasing the pace of SLR like a land-based glacier). And so, it was difficult to tell whether or not this trend was positive or negative vis a vis a confirmed upward trend in SLR.

Concerns that the WAIS melting and/or disintegration could exaggerate SLR had been suggested by John Mercer as early as 1968 and elevated to notoriety by his 1978 *Nature* paper (Mercer 1978). Persistent concerns that events in Antarctica could be a source of global risk were fueled, in part, by two 1983 assessments: one from the US National Research Council (NRC 1983) and another from the US Environmental Protection Agency (United States Environmental Protection Agency (USEPA) 1983). Both were harbingers of the WAIS issue’s coming front and center in global assessments by the Intergovernmental Panel on Climate Change (IPCC). Beginning with the First Assessment (FAR) in 1990, conservative estimates of WAIS contributions to SLR persisted through the Fourth Assessment (IPCC) in 2007. It was from then that new science would suggest an effect that would be noticeable on top of the well-understood manifestation of the thermal expansion of the oceans driven by ‘unequivocal’ warming of the planet (Ö Ö Bernstein et al. 2007).

3 Section 3: some common themes extended from insights from the case studies

The authors have drawn a number of themes from these case studies, and their importance has been interwoven throughout. Each was skillfully set up in the introduction, illustrated in the case studies, and explored more fully in two discussion chapters—Chapter 5 on the policy/science interface and Chapter 6 on what assessments strive to accomplish. The four themes that caught my attention are discussed below.

1. *The interface between science, scientists and policy.*

This interface was deemed worthy of an entire chapter—Chapter 5 on ‘Patrolling the Science/Policy Border’. Subthemes are abundant, and many were previewed in earlier chapters.

In the text, for example, it was noted that some observers had argued that the success of an assessment should be judged on whether or not policies drawn from their content had been enacted. That seems to me to prejudge the outcome of the assessment; or at least, it is an appeal to 20–20 hindsight based on current knowledge in 2019. People frequently do what they do on the basis of what they know at some point in time and/or what they anticipate might happen on the basis of what they know. However, the foundations of their responses need not be informed by modern analyses of decision options. They can, instead, be based on tradition and/or historical rules of thumb. Given this ambiguity, what do we know about the science/policy “border” and how has it played out from one assessment to another? Assessments can frame scientific understanding and explore response choices to well-established issues of concern. Assessments that bolster or replace pre-existing knowledge and/or assessments that explore new issues whose manifestations can produce significant harm represent other possibilities. There are other possibilities, but this volume gives us some insight into all three of these roles drawn from the three case studies:

- Work on the ozone hole certainly expanded and supported preliminary hypotheses and eventually supported an effective policy response even though advocating scientists were excluded in the assessment.
- Work on the WAIS ultimately replaced pre-existing knowledge, but the process was time-consuming and had to morph into the IPCC process to track a detectable impact—rising seas—that could be attributed to some degree to the WAIS. Since an existing explanation of SLR had been available, the contribution of a melting and/or disintegrating WAIS was more difficult to detect, and thus nearly impossible to attribute. Recent assessments have argued that neither detection nor attribution is now difficult—change in the WAIS contributes positively to SLR.
- Political perspectives on acid rain used assessments of uncertainty to discourage, or at least delay, a policy response based on what they labelled as questionable attribution. Success? “Yes”, for an administration that did not want to enlarge government intervention in private markets; but “No” for those who were sure that they had detected dying lakes and ravaged forests and believed that attribution to sulphur emissions was obvious. Ultimately, though, all of the science catalogued by repeated assessments confirmed attribution and elucidated an efficient response—‘cap and trade’ around a politically

tractable target; see, again, USEPA (2017) for the final language in an amendment to the Clean Air Act.

It seems to me that coverage of the policies drawn from these assessments goes to the heart of science policy relevance if not to its power to be policy prescriptive. It is here that inclusion of social science authors played a role by assessing the potential of positive analysis while recognizing that normative analysis was proscribed. The difference between the two is nuanced, but authors and readers of assessments should be made aware.

IPCC authors, for example, have always been proscribed from being policy prescriptive in a normative sense, but that has not impeded their work in Working Groups II and III (since the third assessment (IPCC 2000)) on impacts, adaptation and mitigation. In the present context, my reading is that the evolution of the three case studies covered in this volume show at least two trends in this regard.

First, assessments have generated an increasing number of scientists who understand the intersection of their work and policy debates that define social context. Some have engaged in these debates, and they were more likely to frame their research hypotheses to more fully populate policy context with scientific facts. Others were inspired by earlier assessments that described the policy context of missing scientific information—natural, physical, economic and social. They found a plethora of research questions related to those policy questions, and they knew that success in providing rigorous and honest research to a current generation of policy-makers and the next generation of assessment authors would pay dividends.

It follows, secondly as the future unfolds, that the next assessments will look to populate its author teams with scholars who have answered the policy call from the previous assessments. Why? Because assessments need a policy literature to assess and they need experts to assess it.

Sometimes, it is the community of policy-makers who set this evolutionary step in motion. For example, signatory nations who authored the Paris Agreement under the United Nations Framework Convention on Climate Change (UNFCCC 2015) focused on a 2.0 degree (Centigrade) limit to increases in global mean temperature to craft country-specific emissions targets. But, the same countries of the world also eventually asked the IPCC to compare impacts and risks between a 1.5-degree limit and a 2.0-degree limit. Indeed, the first paragraph of what became IPCC (2018a) read.

“This Report responds to the invitation to IPCC ‘... to provide a Special Report in 2018 on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways contained in the Decision of the 21st Conference of Parties of the UNFCCC to adopt the Paris Agreement.

The IPCC accepted the invitation in April 2016, deciding to prepare this Special Report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty (Synthesis Report of IPCC (2018b))”.

How did this happen? Negotiators at the Conference of the Parties of the UNFCCC in Copenhagen in 2009 (COP2009) were disappointed that the nations of the world could not agree on a Long-Term Global Goal (LTGG); specifically, that could not agree on 2.0 degrees as a target. In Cancun (COP2010), they agreed to an elaborate process, but only after it was

decided that LTGG's would be reviewed, periodically, in concert with the IPCC assessment schedule. The process had two parts: (1) a negotiation part (to be conducted at the annual COP's) and (2) a "Structured Expert Review (SED)" part. The first part converged on a desire to contemplate the prospective value of selecting 1.5° as the LTGG in lieu of 2.0°. The SED part engaged researchers in interactive and intensive sessions over many hours and multiple days. Participants were tasked not only with describing an effective SED structure for iterative LTGGs but also with evaluating the potential (net) value of moving along a 1.5-degree scenario. IPCC (2016) elaborates the foundations of this process—a process that can be credited with inspiring the nations of UNFCCC to make the 1.5-degree request – i.e., evaluate the relative scientific, economic, and practical merits of a 1.5-degree target in comparison with other possible futures.

Researchers who saw this coming in 2015 knew that Intergovernmental Panel on Climate Change (IPCC) (2014a, b, c, d) had covered the 2.0-degree LTGG reasonably well. They also knew that authors of any IPCC Special Report on a 1.5-degree LTGG would need to find some peer-reviewed literature to assess. And so, they responded—producing a growing collection of comparative peer-reviewed literature. Other scholars, having concluded that even a 2.0-degree limit was aspirational, brought 2.5 and 3.0-degree limits into their work. The point here is not the up or down target relative to 2.0°. It is, instead, that an interest inserted into assessment design by policy-makers was, and will always be, taken seriously; and it will be expanded depending upon where the science and the scientists say attention should be paid, especially if there are processes in place by which science and scientists can do just that.

2. *The meaning of consensus and confidence language.*

Throughout the text, but especially in Chapter 6, the authors speak frequently about achieving consensus as the path to univocal conclusions even in the face of enormous and not necessarily diminishing uncertainty. Various early sections of the text contrast consensus with majority voting (reporting out majority and minority opinions, just like the US Supreme Court), but the volume falls a little short in describing what consensus means in an international or even federal context and how it can be achieved in real time. IPCC, the US National Climate Assessment (NCA), international negotiations under the UNFCCC, other assessments, and even negotiations under the Vienna Convention all conformed to precepts, wherein consensus means that nobody in the room disagrees with any word or any line or any number in every sentence, graph or table. It is no surprise that many early assessments at the turn of the last century were criticized as being unbearably conservative.

Nowadays, though, it seems to me that the fundamental question of consensus might better be framed as 'How and why can authors of modern assessments (who are working under this definition of consensus) report conclusions with 'low' or 'middle' or even 'high' confidence'? The answer? Because, they (and their clients—the countries of the world or members of Congress, or whatever group signs on) all understand that consensus can extend to the confidence statements, themselves. So, a consensus conclusion that X causes Y with medium or low confidence does not mean that nobody in the decision room objects to the conclusion that X causes Y. It means, instead, that nobody in the room objects to the authors' more detailed analysis of process and evidence that can support only a medium- or low-confidence statement. It also means that nobody in the room objects to including the medium or low confidence conclusion in the assessment because the potential consequences are large. It is critical that private citizens also understand this meaning. It is also critical for all to understand

that the second ‘no objection’ consensus relies on a risk-based (risk management) perspective. More on that is just below in Sections 3 and 4.

And how can this happen? Because modern consensus statements are supported by extensive background work undertaken behind the scenes by framing institutions like IPCC, NCA and the NAS. Figure 1, for example, offers a visual display of how confidence (for any particular hypothesis or conclusion) can be explored as a function of agreement (process understanding of the subject) and evidence (the quality and quantity of available data). It replicates Fig. 1 in Mastrandrea et al. (2010) that was prepared by economists and scientists to help author teams of the three working groups of the Fifth IPCC Assessment (AR5) achieve rigorous, consistent and therefore comparable confidence conclusions.

AR5 authors were asked, as they considered any proposed hypothesis or any proposed result for their various chapters, to subjectively locate its characteristics along these two critical axes of Fig. 1 Very high confidence could then be supported by significant and strong agreement about process supported by a multitude of quality data that supported that agreement (the upper-right part of the matrix). Judgments of very low confidence could similarly be supported if experts disagreed across competing understandings of processes (and therefore attribution) while data were scarce (the lower-left part of the matrix); this is the early history of the WAIS story when data were scarce and a second hypothesis had been suggested.

Other judgements lay in between, and the shading suggests a tradeoff between agreement and evidence for medium or high confidence arcing through the middle of the matrix. For example, to a group of economist assessors, there would be medium confidence in any description of how the macro economy works. Why? Because economists always disagree, but more precisely because there is serious disagreement about the general process (monetary views versus neo-Keynesian views; see Krugman and Wells 2015). This argument persists even though there are more quality data available about major economic indicators and drivers than anybody can fully exploit. A different collection of authors facing a different reality could also assign medium confidence to a well-understood phenomenon even if data were sparse and scattered.

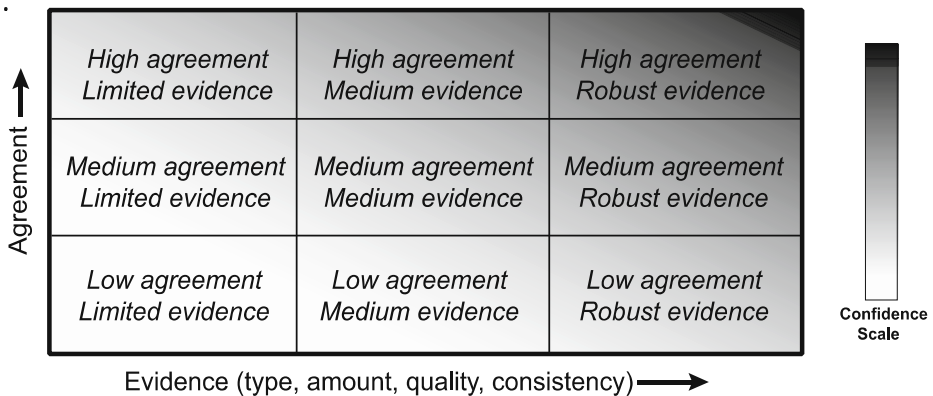


Fig. 1 A working matrix for determining confidence. Author teams were asked to locate each of their major conclusions and/or hypotheses somewhere within this matrix. The shading would them assert a level of confidence, but their work would not be complete until they defended their location and assertion to their peers. Source: Fig. 1 in Mastrandrea et al. (2010)

Training authors to think and organize their thoughts in terms of these two dimensions is only the first step in producing coherent and internally consistent top-down assessments like IPCC—more policy prescriptive efforts like *America's Climate Choices* (2010) and even a series of assessments like the WAIS or NAPAP experiences. The next step is to insist that authors defend their confidence conclusions (their chosen location on the matrix) before.

- An entire chapter author team (experts on the same topic)
- A diverse collection of authors from different chapters (experts in other topics relevant to a broader issue), and then
- Representatives of disciplinary working groups (e.g., in the IPCC WGI (science), WGII (impacts and adaptation) and WGIII (mitigation)).

This arduous and time-consuming work has the potential of producing a workable degree of consistency for decision-makers' considerations across multiple contexts. But it does not explain why consensus assessments now publish conclusions to only which low- or medium-confidence conclusions can be assigned. Nor does it explain why many decision-makers now insist on reading about such conclusions.

3. *Risk analysis and risk management.*

The idea of casting assessments largely in terms of analysing risk permeates this study—starting in the introduction, appearing intermittently in the case studies, but highlighted more intensively in both the policy-border discussions of Chapter 5 and the ‘what assessments do’ presentations in Chapter 6. Risk, in its most elementary form, is the product of likelihood and consequence. Figure 1 in chapter 4 of NRC (2011) displays a companion matrix to Fig. 1. It is again useful in organizing thoughts, this time about the sources of risk.

The applicability of risk analysis to climate change is, as it turns out, the result of a bit of ‘new science’ that was produced in 2007 by three authors of the Summary for Policymakers of the Synthesis Report for the AR4. The following sentence, crafted over many days in Estes Park, Colorado in 2007 by Stephen Schneider, Gary Yohe, and William Hare, achieved word for word consensus approval from more than 160 countries in the subsequent IPCC plenary meeting in Valencia, Spain (page 22):

Responding to climate change involves an iterative risk management process that includes both adaptation and mitigation, and takes into account climate change damages, cobenefits, sustainability, equity and attitudes to risk. (IPCC 2007s, pg 22; emphasis added by this author).

It took 5 days to convince the nations of the world that these words were the products of synthetic thought across information reported by 3 working groups and not some forbidden new knowledge hidden on the last page of a proposed document in the middle of the night. To rephrase its content,

- Response to climate change by reducing emissions (mitigation) or ameliorating damages (adaptation) is most informatively framed in terms of managing risk.
- The decision process has to be designed to be iterative because knowledge evolves (a harbinger of SEDs?).
- Uncertainties do not necessarily diminish over time, and tails can become more important.

- The catalogue of risk metrics must include both damages and co-benefits. That part is easy.
- But decision metrics should take account of social objectives like sustainability, equity, and risk aversion. Surely, more factors of effective evaluation will be added to this list as the future unfolds.

These words may or may not be the most important from all of the IPCC assessments but they certainly opened the way for reporting conclusions located nearly anywhere in Fig. 1. The clients of the IPCC assessments—the countries of the world who were signatories of the UNFCCC—had, by consensus, instructed their scientific assessors not to shy away from low likelihood conclusions if they carried large consequences (i.e., high potential risk). Since I am writing for Stephen Schneider’s journal, and since he was a co-author of these words, I feel comfortable in rephrasing this conclusion yet again. It was our experience in negotiating for more than 4 days across more than 160 countries at the Valencia plenary that they all (eventually and ultimately by consensus) wanted to be informed about the dark tails of the distributions of what the future might hold—even the USA, China, Saudi Arabia, and.... These possible futures in the tails of general distributions of damages or benefits calibrated in one of many metrics (currency, human lives, ecosystem diversity...) occur either because the climate system could be moving toward extreme outcomes, or because the damages associated with even the most benign climate future could be extreme, or *both*.

The original and endorsed words prepared the world to come to grips with modern forensic attribution of some of the worst manifestations of climate change: enormous forest fires, severe and persistent droughts, increasingly frequent flash and riverine flooding episodes, increasingly extreme hurricanes, other hurricanes that would turn into precipitation disasters by getting separated from steering air currents and therefore stall for 48 h over one location, and more. Having ‘detected’ an increase in intensity and/or frequency, what proportion of those events could be attributed to climate change? Some answers are ‘a lot’—e.g., wildfires in California that used to be controlled but now erupt in hours because of beetles’ damaging standing forests. Of course, though, other human activities matter (like not maintaining proper fire barriers around properties and roads. It is become clear, though, that such confounding factors cannot explain all of the statistical deviations that are being observed and calibrated.

Current events are, more succinctly, being drawn from a new distribution of possibilities. Answers are emerging from scholars armed with risk-based assessments of confidence in detection (conclusions that we have observed a change that holds the potential for severe damage) and confidence in attribution of observed changes to their underlying sources (from changes in local conditions like site-specific climate change all the way up to laying significant blame on global climate change caused by human activities).

4. *New Knowledge*

Commentary about new knowledge born of assessments, especially those whose rules prescribed authors’ bringing forth any new science, was scattered throughout the volume. Nothing in the text did violence to my recollections of my experiences, but I think that I can add some insight designed to fill in what has happened since 2014.

Let me assert from the start that any assessment worth its salt assesses all of the most recent information available and melds new insights from this more recent literature into the historical context. Some inclusions confirm existing knowledge, but some do not (the WAIS story versus NAPAP). Deciding how and what to include is, therefore, a tricky business with its own set of

rules, and the output of any assessment can easily depend on those rules. What follows tries to synthesize insights from the various mentions of knowledge in the volume and to provide a more direct perspective from my own experience.

At the scoping meeting for the AR5, for example, a chair of a working group proclaimed that his authors would not be allowed to assess a result born of a single peer-reviewed paper. In the context of thousands of other references for other conclusions that would be cited by authors of a single working group (e.g., WGII) in their support of many confidence conclusions, this seemed reasonable. This was his (it was a ‘he’) attempt to craft a rule, but chaos broke out from authors who remembered the risk conclusion of AR4. These authors argued that such exclusion from above without reference to context would be a clear violation of the intent of a risk-based approach to assessment literature. It was his authors, and authors of other working groups, who should be (and actually were) empowered by acceptance of a risk-based approach to assess even very limited literature—informed by instructions about how to write about low confidence hypotheses that might have enormous consequences. In fact, a few issues with support from only a few papers can be used to support an emergent concern.

Could one paper be an outlier? Sure. Is there danger in attributing risk to a single event? Sure. Would that be an unsupportable outlier? Maybe. But these are questions for the author teams to decide without prejudice from the chairs. Given the pattern of scientific literature, many lines of inquiry frequently come from one paper. Given the artificial but necessary literature cutoff dates of assessments, that one paper may be all there is (even though many more may be in process and soon to be in print). Given that four or six+ years generally pass from one assessment to another, should a potentially consequential outlier not be highlighted for scholars if not for decision-makers because, otherwise, important news could be delayed for a half of a decade? The countries attending the Valencia IPCC plenary had said “Yes”—by consensus.

If this sort of risk-based support structure had been in place earlier, perhaps the various WAIS assessments might have been legitimately less conservative. The synthetic lesson here is that the leaders of existing and future assessments should be deputized to engage and *instruct* authors of future assessments about how to cope with new knowledge, historical perspective and the interface coloured by historical memory—all from a risk-based perspective.

How might these *instructions* codified by the major institutions that run large assessments? By writing rules and produced by preliminary and sponsored workshops. What would those rules cover? My personal history on such things provides, I think, some insight of best practices:

- Perhaps, the *first* notable rule would define the publication deadlines beyond which new information from the current literature cannot be considered (unless suggested by an outside expert reviewer of an early draft).
- The *second* rule typically would define the dominating role of peer-reviewed literature in comparison to the acceptability of grey literature. Some elements of the grey literature are topic-specific reports. Many (but not all) will have passed significant and multifaceted reviews of their own. In either case, my impression of the rules protocol was ‘beware’. Not necessarily because the content was questionable, but because the adjective ‘peer’ could not be applied.
- Other thoroughly reviewed but troublesome documents include published studies that are really assessments in their own right; they are the source of a *third* rule.

Here, the rules suggest not trusting the judgement of these authors—not because they were not wise but because they drew from an earlier literature whose coverage stopped well before the publication of the ongoing assessment. This, in my experience, meant that modern authors should work backwards from assessment conclusions and consider carefully the supporting literature that was then available. They can then produce rough working documents—preliminary time-constrained confidence evaluations that they would use as consistent inputs for the next generation of assessments.

As an example, here, consider a regularly scheduled IPCC assessment or Special Report. Previous IPCC assessments would be appropriate input, but only if the current authors took time and context into account. Earlier assessments set the context (e.g., low confidence that changes in the WAIS contributed positively or negatively to global SLR from Chapter 4), but new assessments (IPCC AR5, 2014s) can change the conclusion (changes in the WAIS have contributed positively to SLR and will continue to do so).

- The *fourth* rule generally would require that all cited papers and volumes and reports be deposited in a common electronic site so that readers can find the underlying literature if they want.
- The *fifth* related rule concerns creating ‘traceable accounts’ for deposit at the same site—separate descriptions of how current assessment authors had reached their confidence conclusions (historical and current) and defended them before their colleagues.
- *Finally*, a review process should certainly be enumerated. Each chapter of any IPCC assessment report, for example, goes through at least three and sometimes four drafts. Each draft is reviewed, in turn, by internal government experts, by an external expert community, by the public at large, by governments (not necessarily government experts) which are signatories of the UNFCCC, and finally in person at a giant plenary before representatives of those governments.

Summaries for Policymakers that synthesize the content of the entire collection’s chapters with direct reference back to chapter language are frequently major sources of new knowledge. Every word is usually accepted by consensus in the giant plenaries. There, consensus means that no government in the room disagrees with any word in sentences that are displayed one after another for approval (including words that report confidence conclusions). As cumbersome as this synthetic process is, it produces summaries of important content that sometimes comes close to and sometimes steps over the red lines of the ‘thou shall not do new science in writing this assessment’ rule.

The sentence from AR4 that was highlighted above is a perfect example. Its content was gleaned from material scattered across three working groups, but its message was new (social) science and its importance was not recognized until it became a consensus conclusion from all of the signatories. After that, it actually became an organizing principle for subsequent assessments—two US National Climate Assessments (NCA 2014, 2018), the five volumes of the NAS *America’s Climate Choices*” (NRC 2010a, b), communications from the New York (City) Panel on Climate Change to public and private decision-makers across all five boroughs, and all subsequent IPCC assessments and special reports.

Assessments also create new knowledge when they identify gaps in process understanding and/or relevant data and thereby focus attention on conclusions with low or medium confidence; i.e., if they include high consequences, then they identify research topics for future scholars who are looking for places where their work would find fertile ground and also be most significant. Indeed, in fulfilling this role, assessments save countless hours of justification language in proposals for external funding. Something as simple as “The IPCC thinks that there is a gap here. This is how they reached that conclusions. This is why they think it an important gap to fill. And here is where my work will help advance knowledge.”

4 Section 4: some concluding thoughts

To repeat my lead: Oppenheimer et al. (2019) have produced an assessment of assessments. Its last few chapters add new knowledge, on the basis of three diverse case studies, about how assessments that are created from the top down filter their conclusions into the decision-making context. Two of the case studies are part of the climate change literature, so readers of *Climatic Change* should see considerable value in having a look. The other, on sulphur emissions, is certainly important for readers, as well, even though it contrasts markedly from greenhouse gas emissions; greenhouse gas emissions are not ‘traditional pollutants’ for which damages are local and reversible. They are emissions that span the planet regardless of source and produce effects that are felt global and frequently irreversible.

As noted, this volume should be essential reading for participants of any large environmental assessment. Past participants in top-down assessments like IPCC and NCA will see themselves in the pages. The anecdotal evidence that they tell their friends, families and students will also fit well into these pages. More importantly, natural and physical scientists will see how their work can be transmitted across humanity to help inform opinion about what is going on and perhaps what to do—on the basis of rigorous science.

Perhaps, the largest value will be found among the young scholars who do their homework after being invited to participate in their first assessment. After they read this volume, they will understand what to expect and why their signing on is a valuable investment of their time.

Finally, external commentators of past and future assessments will, if they take the time to read this whole volume, come to understand.

1. What assessments do (and what they do not do)
2. What ‘consensus’ means
3. What taking a risk management approach to climate change (or really any environmental hazard) means

When they draw from this reading of this volume to write reports about future climate change and other environmental risks, I hope that they touch each of these three bases firmly. If they do, maybe they can steal home by creating something really important.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Blake R, Klaus J, Yohe G, Zimmerman R, Manley D, Rosenzweig C, Solecki W (2019) “Chapter 8: Indicators and Monitoring”, in New York City Panel on Climate Change 2019 Report, Annals of the New York Academy of Sciences, New York Academy of Sciences. <https://nyaspubs.onlinelibrary.wiley.com/toc/17496632/2019/1439/1>
- Intergovernmental Panel on Climate Change (IPCC) (2014a) Synthesis Report, (TAR – 2000-s), (AR4-2007-s), and (AR5 – 2014-s), www.ipcc.ch
- Intergovernmental Panel on Climate Change (IPCC) (2014b), Report from Working Group I, (TAR - 2000-I), (AR4 - 2007-I), (AR5 - 2014-I), www.ipcc.ch
- Intergovernmental Panel on Climate Change (IPCC) (2014c) Report from Working Group II, (TAR- 2000-II), (AR4–2007-II), (AR5–2014-II), www.ipcc.ch
- Intergovernmental Panel on Climate Change (IPCC) (2014d), Report from Working Group III, (TAR - 2000-III), (AR4– - 2007-III), (AR5–2014-III), www.ipcc.ch
- Intergovernmental Panel on Climate Change (IPCC) (2016) A new science-policy interface: the structured expert dialogue of the 2013–2015 review of the UNFCCC, available at ... https://www.ipcc.ch/site/assets/uploads/2016/02/Structured_Expert_Dialogue_2013-2015
- Intergovernmental Panel on Climate Change (IPCC) (2018a) Global warming of 1.5°C., Masson-Delmotte V, P Zhai, H-O Pörtner, D Roberts, J Skea, PR Shukla, A Pirani, W Moufouma-Okia, C Péan, R Pidcock, S Connors, JBR Matthews, Y Chen, X Zhou, MI Gomis, E Lonnoy, T Maycock, M Tignor, and T Waterfield (eds), Cambridge University Press
- Intergovernmental Panel on Climate Change (IPCC) (2018b) Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, www.ipcc.ch
- IPCC TAR WG2 (2001) McCarthy JJ, Canziani OF, Leary NA, Dokken D J, and White K S (eds), Climate change 2001: impacts, adaptation and vulnerability, contribution of working group II to the third assessment report of the Intergovernmental Panel on Climate Change, Cambridge University Press, ISBN 0-521-80768-9 (pb: 0-521-01500-6)
- Krugman P, Wells R (2015) Economics, fourth edn. MacMillan Publishers, New York
- Mastrandrea M, Field C, Stocker T, Edenhofer O, Ebi K, Frame D, Held H, Kriegler E, Mach K, Plattner G-K, Yohe G, Zwiers F (2010) Guidance notes for lead authors of the IPCC fifth assessment report on consistent treatment of uncertainties. IPCC. www.ipcc.ce
- Mercer J (1978) West Antarctic ice sheet and CO2 greenhouse effect: a threat of disaster. *Nature* 271:321–325
- Molina MJ, Rowland FS (1974) Stratospheric sink for chlorofluoromethanes: chlorine atom catalysed destruction of ozone. *Nature* 249:810–812
- National Climate Assessment (NCA) (2014) www.nca2014.globalchange.gov
- National Climate Assessment (NCA) (2018) www.nca2018.globalchange.gov
- National Research Council (NRC) Committee on Atmospheric Transport and Chemical Transformation in Acid Precipitation (1983) Acid deposition: atmospheric processes in eastern North America: a review of current scientific understanding. National Academy Press, Washington, DC
- National Research Council (NRC) (2010a) Report of the adaptation panel, America’s climate choices. National Academy Press, Washington, DC
- National Research Council (NRC) (2010b) America’s climate choices. National Academy Press, Washington, DC
- National Research Council (2011) Climate stabilization targets: emissions, concentrations, and impacts over decades to millennia. Washington, DC: The National Academies Press. <https://doi.org/10.17226/12877>
- Ö Ö Bernstein L, Bosch P, Canziani O, Chen Z, Christ R, Davidson O, Hare W, Huq S, Karoly D, Kattsov V, Kundzewicz Z, Liu J, Lohmann U, Manning M, Matsuno T, Menne B, Metz B, Mirza M, Nicholls N, Nurse L, Pachauri R, Palutikof J, Parry M, Qin D, Ravindranath N, Reisinger A, Ren J, Riahi K, Rosenzweig C, Rusticucci M, Schneider S, Sokona Y, Solomon S, Stott P, Stouffer R, Sugiyama T, Swart R, Tirpak D, Vogel C, Yohe G (2007) Climate Change 2007: Synthesis report (for the Fourth Assessment Report of the Intergovernmental Panel on Climate Change), Cambridge University Press, Cambridge
- Oppenheimer M, Oreskes N, Jamieson D, Brysse K, O’Reilly J, Shindell M, Saxeck M (2019) Discerning experts. University of Chicago Press, Chicago
- United Nations Framework Convention on Climate Change (UNFCCC) (2015) The Paris Agreement <http://unfccc.int/resource/docs/2015/cop21/eng/109r01.pdf>
- United States Environmental Protection Agency (USEPA), (1983) Projecting future sea level rise, Washington, DC

United States Environmental Protection Agency (USEPA) (2017), Acid Rain Programs and Regulations – Title IV of the Clean Air Act – Acid Deposition Control available at https://www.ipcc.ch/site/assets/uploads/2016/02/Structured_Expert_Dialogue_2013-2015

World Resources Institute (WRI), National Climate Plans (INDC's) by the Numbers (2015) <https://www.wri.org/blog/2015/10/national-climate-plans-indecs-numbers>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.