The Science/Policy Interface and The Role of Participation in Assessment Processes

Draft 4.0

Jill Jäger, Alex Farrell, and Stacy VanDeveer


1. Introduction

This paper reports on the deliberations of a working group that met during the 1998 Global Environmental Assessment and Public Policy Workshop held in Bar Harbor, Maine in June 1998. The discussions focussed on the characteristics of particular institutions and their role in influencing the assessment process and the outcomes of assessment; the extent of political control and influence in the assessment process; and a number of questions about participation in assessment processes. The group also considered briefly the implications of their findings for the design of assessment processes.

The GEA project has come to understand assessment as a social process, including but not restricted to its formal products. As pointed out in the report of the 1997 Workshop (The Global Environmental Assessment Project, 1997), this social process perspective directs attention beyond the content of assessment reports to encompass questions regarding participation, presentation, evaluation and how the boundaries between the scientific and policy dimensions of assessment are negotiated and legitimized. These topics are at the center of the present report.

In addition to recognizing that assessment is a social process bridging the realms of science and policy, the working group at the 1998 workshop acknowledged that there are different types of assessment. For example, a distinction can be made between insight-oriented assessments carried out within a policy-oriented context and decision-oriented assessments that can be investigative in nature (in which the answer is not known at the start of the assessment process) or to provide justification (i.e., the answer is known at the start). A further classification of assessments distinguished between three types: classic decision-making -- trying to get the answer as to what to do; advocacy -- determining what has to be believed in order to have a position accepted or rejected; and, marketing -- laying out in an orderly way the argument for other people to see. (See Figure 1) It was also noted that industry performs many high quality assessments but these have not, so far, been studied within the GEA project. Many of the assessments studied in the GEA project and discussed at the Workshop are consensus-based. This opened up the discussion on the role of consensus and possible alternatives and this discussion is reflected in this report. Finally, the group noted that a distinction can be made between assessment processes in which there are...

---

1 Participants in the Working Group were: Liliana Botcheva, Barbara Connolly, Ellis Cowling, Alex Farrell, Wendy Franz, Leen Hordijk, Jill Jäger, Greg Macrae, Edward Miles, Ronald Mitchell, Granger Morgan, Tony Rosenbaum, Milton Russell, Michael Uhart, Stacy VanDeveer, Derek Winstanley. The authors are grateful to Wendy Franz for note-taking in one session and Ron Mitchell for facilitation throughout the Working Group discussions.
parallel streams of disciplinary, analytic activity, in which an attempt is made to integrate these streams, versus truly integrated assessment activities, which are much harder to carry out.
2. Characterizing the Science-Policy Interactions in an Assessment Process

The starting point for this discussion was the question of why scientists and policy-makers need to 'negotiate' the boundary between their domains in any given process (Jasanoff, 1990). Jasanoff uses examples of standard-setting for pollutants and of decisions to allow the use or not of drugs to show how the challenge of producing politically relevant but scientifically credible advice is resolved by "boundary negotiation". The experts themselves often acknowledge that what they are doing is not "science" in any ordinary sense but a hybrid activity that combines elements of scientific evidence and reasoning with large doses of social and political judgement. Importantly, however, the product of expert groups, such as the EPA’s Science Advisory Board (SAB), must still be labeled as “science” and not as “policy”, and boundary negotiations involve agreements between regulatory bodies (such as the EPA) and expert advisory groups (such as the SAB) as to what issues each will deal with and what issues will be shared between them.

It is important to note that Jasanoff’s analysis can be classified as “constructivist” and is countered by a “positivist” view that “the best thing for scientists to do when they are far apart is to speak the truth exactly as they understand it to the powers that be. “ (Wildavsky 1992 p. 512). The Workgroup was divided as to which of these two views is most accurate and useful, replicating the tension which exists in the literature. This division often falls along the disciplinary lines of physical scientists (usually positivists) and social scientists (often constructionists).

The GEA project is concerned with a larger and more diverse set of activities than standard-setting, however. In the area of air pollution, for instance, assessment activities include examinations of atmospheric chemistry, meteorology, transportation demand, power plant engineering and economics, and many others beyond the fields of toxicology and epidemiology which dominate standard-setting. Assessments of other environmental issues further increase the
range of activities examined. These provide a useful basis for exploring further the boundaries between "science" and "policy" in a range of assessment activities.

As shown below, examples from a number of assessment processes illustrate the kinds of judgements that are made throughout assessment processes. Jasanoff concluded that since the issue is not whether there are boundaries between "science" and "policy" but where they are, one of the key roles of expert advisory bodies is to provide a forum to stake these boundaries out. Allowing expert bodies to perform this function is crucial to obtaining the political acceptability of advice and it appears that when the process is successful (and scientific credibility is also obtained) it is not possible for political adversaries to deconstruct the results or attack them as "bad science". Thus, Jasanoff's advice includes the idea that the problem of politically-motivated bias in appointing members of expert committees is unavoidable, and should be countered with administrative devices to limit it, not impossible attempts to banish it (pp. 244-245).

The Intergovernmental Panel on Climate Change (IPCC) is frequently cited as an assessment process where there is a solid boundary between the science and policy realms -- the Panel itself produces assessments based on the state of the art of knowledge about the climate issue. The panel does not make policy recommendations, so the scientific work is to that extent insulated from the decision-making process. Only the process of producing policy-makers' summaries of IPCC report increases the interactions between the science and policy realms but it still takes place in an institution where the boundary between science and policy is quite strictly drawn. In contrast, the assessment process within the bodies of the Convention on Long-range Transboundary Air Pollution (LRTAP) provides opportunities for formal and informal interactions between scientists and policy-makers, cooperative sharing of information and cooperative redefinition and elaboration of questions. Similarly, the Ozone Transport Assessment Group (OTAG) process in the US was characterized by intensive, although somewhat brief, interactions between politicians and technical staff aimed (in the end) at improving the politicians’ understanding of the physical characteristics of the problem. So, unlike IPCC, the LRTAP and OTAG assessment processes are characterized by more porous boundaries between the science and policy realms.

Experience with a range of assessment processes indicates that the interactions between the science and policy realms in an assessment process can be characterized along a number of dimensions. An indicative list of such characteristics includes:

- Cooperative vs. Conflictual
- Reciprocal vs. Unidirectional (a two way discussion or flow of information across the boundary vs. a one-way discussion)
- Hierarchical vs. Egalitarian (one side assumed to be stronger than the other vs. assumed equality)

For example, the first two rounds of IPCC assessments were characterized by a largely unidirectional flow of information, in which the Chair of the Panel informed the policy-makers (negotiators of the climate regime) of the results of the assessment LRTAP, conversely, has much more cooperative dynamics between scientific and technical assessors and policy makers and information, ideas and preferences flow across the boundary in much more reciprocal fashion.
Moreover, as indicated in the introduction above, the assessment process itself can be characterized according to a number of criteria which can also influence the nature of the science/policy interface:

- Open-ended, insight-oriented analysis (investigative) vs. Position-supporting advocacy (justification)
- Private vs. Public
- Transparent vs. Opaque

Clearly a single assessment process could be characterized along several, if not all, of these dimensions. Furthermore the kind of interaction can be scaled in some cases according to the frequency of the interaction. The working group tried hard, in using these dimensions, to avoid explicit and implicit “good” or ”bad” distinctions. For example, it cannot be assumed that a unidirectional boundary is automatically better or worse than cooperative exchange of information between scientists and policy makers. Within the working group, some members were very strong in their belief that science and policy should be kept very separate (solid boundaries) and interactions should be minimal and unidirectional, while others were quite happy with porous boundaries and lots of complex interaction between science and policy within assessment processes.

2. Why are assessment processes initiated?

Part of the explanation for the characteristics of an assessment process and the role that individuals and institutions play in it can be traced back to the origins of the assessment processes. Who called for a particular assessment process and why? The 1997 GEA Workshop produced a broad catalogue of assessment goals from the perspectives of assessors and users, individually and jointly (GEA, 1997). The 1998 Workshop began by looking at specific assessments and why they were initiated. Examples included the National Climate Assessment in the Netherlands, which was initiated by the Environment Minister in order to convince his own political party that the issue required attention and the OTAG process that was initiated by decision-makers because of a combination of a perceived political crisis and a technical crisis. In the US there are examples of assessments initiated in response to lawsuit threats (e.g. some assessments set up by EPA, including some of the assessments discussed, for example, by Jasanoff (1990)). The NAPAP was initiated in order to increase knowledge but there was also a second reason, which was to delay action on the issue. Power companies have initiated assessments in order to be able to rebut Best Available Technology Standards and other stakeholders have initiated assessments to provide guidance for their own research and development activities.

More generically, the initiation of assessment can be the need for insights into complex problems. This knowledge might be required generally because of its policy relevance or specifically in response to a crisis or a specific request. The assessment can be called for in response to another assessment or to block, avoid or promote action. The origin of an assessment is an important determinant of the selection of people involved in the assessment process and the design of the process itself. However, it is clear that assessments rarely have one goal or a definitive moment when a person or group called for an assessment. Instead, in most of the cases looked at in GEA,
multiple actors wanted some kind of assessment process (often different kinds) for a variety or reasons. As such, the task for analysts, at least in part, has been to try to parse out how certain goals or frames and certain patterns of participation arise or come to dominate within assessment processes.

4. The roles of individuals or institutions in an assessment process

Kai Lee (1993) discussed the dilemmas implicit in pursuing science in a political setting and pointed out that "Science and politics serve different purposes. Politics aims at the responsible use of power; in a democracy, ‘responsible’ means accountable, eventually to voters. Science aims at finding truths - results that withstand the scrutiny of one's fellow scientists.” As a positivist, Lee uses the idea of a spectrum of occupational roles and social functions originally proposed by Price (1965) to identify the distinctive contributions made by different groups to a technological society. Lee argues that the roles of individuals as politician, administrator, professional analyst and scientist are separate and that a single person cannot play several different roles at once at least without the risk of losing legitimacy.

The working group discussions recognized that it is possible to play multiple roles in some situations, but it is difficult to do so and the main element in successfully playing multiple roles is trust. This trust can be vested in a person or in a process. It was suggested that decision-makers can play several roles without losing legitimacy but that scientists might not be able to do so. At least in the case where there is a rigid separation of science and politics in an assessment process, the role of the scientists is to discover facts, whereas the role of the policy analyst is to weigh facts and communicate options and the decision-makers (stakeholders) have the role implied in their name.

In the Ozone Transport Assessment Group (OTAG) (described in more detail in the GEA paper of Farrell and Keating (1998)) institutions did play multiple roles without much loss of the legitimacy of the outcome because of the relatively open and transparent process (although other factors had negative influences on legitimacy). It was argued that in the National Acid Precipitation Assessment Program (NAPAP) individuals and institutions currently play multiple roles but there is no apparent loss of legitimacy because of the low salience of the issue, i.e., if the issue were high on the political agenda then it would matter much more if individuals were playing multiple roles. Using the IPCC as an example, in which the scientists in Working Group 1 were insulated from the policy realm and maintained their legitimacy, it was argued that if the policy/politics realm is very turbulent, playing multiple roles can be very dangerous.

Clearly, the question of whether individuals or institutions can play multiple roles in an assessment process without jeopardizing legitimacy (e.g., carry out scientific assessment, make policy recommendations and perhaps even participate in the implementation of policies) is very much dependent on the context within which the assessment is carried out. As pointed out above, it is also very much dependent on whether there is trust in the individual or process. In the LRTAP process, individuals play multiple roles, especially by carrying out scientific assessments and participating in the decision-making process but there is confidence in the process and the individuals involved so that the outcome is considered to be legitimate.
5. Value judgements made in assessment processes

As pointed out in Section 2, the experts involved in an assessment process often recognize that they are involved in a hybrid activity in which scientific expertise is accompanied by a considerable amount of social and political judgement. Others apparently do not recognize this aspect of their activity and think that all of their contributions to the assessment process are 'objective scientific facts' uninfluenced by any value judgement.

The discussion of the working group illustrated clearly that value judgements are made at all stages of an assessment process. During the initiation phase, the framing of the question(s) to be considered in the assessment involves value judgements or decisions about what will be considered and what not and about who will be involved and how and why they participate. Value judgements are also embedded in decisions about which results will be used in the assessment, about data analysis and interpretation. The process of making assessment summaries and making recommendations is also based on value judgements.

An example of the value judgements in the framing of the issue was provided with reference to NAPAP, which decided to look first at local problems. NAPAP thus focussed on models with a short time-frame. In contrast Scandinavians focussed studies on the effects of acidification on forests and fish -- long-term problems that do not need minute-by-minute pollution data. The judgement about whether the study was about problems that had a short or a long time scale determined the kinds of models that were used, the kinds of data that were collected and the conclusions drawn with respect to dealing with the issue. LRTAP’s prioritization of acidification effects had implications for national participation as well. Eastern and Southern European countries simply did not perceive these problems to be their problems. CEE pollution concerns, such as urban air pollution and human health effects, were not at the center of LRTAP. Today, LRTAP is expanding the scope of its interests to include POPs, heavy metals, eutrophication, and tropospheric ozone, yet the assessment processes participation patterns largely reflect national interests in acidification (VanDeveer, 1998).

During the assessment process one of the main ways in which value judgements are introduced is through the use of models. Scientific inquiry is replete with mathematical modeling, so it is not surprising to find that many assessments involve modeling in some form, including LRTAP, NAPAP, IPCC, and OTAG. Indeed, in some ways the modeling effort is at the heart of a number of integrated assessments. In these processes modelers must produce a mathematical simplification of the physical (and sometimes economic) problem that they can solve with the tools they have. Producing a model is a selective effort which necessarily involves judgements about what to leave in the model and what to ignore. In order to communicate with non-technical people, the modelers must perform a second stage of simplification. As discussed below, assessments could benefit from documenting more explicitly the kinds of judgements made throughout the process.

Another example of judgements made during the assessment process was provided by LRTAP, where during the 1991-1992 time period there was a dip in interest in the negotiations. The
modelers thought that this could be due to the use of three different models to support the negotiations, which was possibly increasing the sense of uncertainty. The Task Force on Integrated Assessment Modeling decided to go ahead only using the RAINS model to reduce the perceived uncertainty but there were no explicit selection criteria for choosing this particular model. A further example of judgements made in the LRTAP process is provided in the GEA paper of Patt (1998), who describes how actors involved chose to center analysis on ecosystem damage and not on human health impacts. This choice significantly influenced the assessment process and the design of policy solutions. The decision to base policy on critical loads for ecosystems rather than on emission reductions to improve human health appears to have been based on political considerations, since according to studies cited by Patt, the economic damages from health effects in Europe far exceed those from damages to materials or crops.

A number of other recent publications have looked at the topic of value judgements in assessment processes. Schneider (1997) commented on the dangers that analytic methods with limited capabilities bring to the public debate given that not all potential users of models (Integrated Assessment Models -- IAMs) in assessment processes will be aware of hidden values or assumptions that are inherent in all such tools. Schneider provided a checklist of issues or practices to bear in mind when building or applying IAMs and this list included:

- Specify clearly at the outset and in the conclusions of presentations or publications the limited context of each particular IAM exercise.
- Cite alternative approaches and contrast them to your approach, stressing how each treats uncertainty and deals with the many value-laden components of the analysis.
- Provide as many menu options as practical, especially for those choices which deal with culturally-dependent components or “imaginable surprises”.
- Perform as many “validation” tests as possible, and when not practical, discuss, based on qualitative reasoning, the credibility of structural assumptions, input data, and model parameters, and their relevance to policy issues are being considered.
- Note components of the IAM which are particularly sensitive (or insensitive) to aspects of the problem that are controversial and thus likely to change with evolving research.

Morgan and Dowlatbadi (1996) summarized insights from five years of integrated assessment activity in which they played the leading role. On the basis of their experience, Morgan and Dowlatbadi presented basic principles, which, they believe, should guide all integrated assessments and these included:

- The characterization and analysis of uncertainty should be a central focus of all assessments.
- Parts of the problem about which we have little knowledge must not be ignored. Order-of-magnitude analysis, bounding analysis, and carefully elicited expert judgment should be used when formal models are not possible.
- Treatment of values should be explicit, and when possible parametric, so that many different actors can all make use of results from the same assessment.
Morgan and Dowlatabadi concluded that integrated assessment faces methodological and philosophical challenges. The incorporation of uncertainty and of values are extremely challenging but also crucially important if the assessments are to be useful for decision-making.

The cases examined in the GEA project this year revealed significant gaps between these goals and actual practice, some of them being virtually ignored, particularly the issue of uncertainty. Assessment processes discussed in the GEA summer workshop provided examples of value judgements made at all stages of the process, sometimes without the participants and often without the users of the assessments being aware that these judgements were being made or of their implications. Recent literature and discussions at the workshop suggest that explicit treatment of values and assumptions made during the assessment process is one important way that assessment processes can be improved.

6. The role of consensus in assessment processes

Assessment, as a bridge between science and policy, is bringing together knowledge from more than one disciplinary approach in a way that is useful for the policy-making process. How do assessment processes deal with dissenting views about the state of knowledge both within and between disciplines (or between groups with different interests as appears to have been the case in OTAG) and with dissenting views about the interpretation of that knowledge? Empirically it seems that most of the assessments that the GEA project has looked at use a consensus approach among the participants in the assessment process. However, there are a number of different interpretations of what is meant by 'consensus'. Some assessment processes have dealt with dissent through the production of a 'minority report' but this seems to be rare for global environmental issues.

The OTAG process documented by Farrell and Keating was consensus-based but the meaning of consensus varied from sub-group to sub-group. In addition, it was important to the dominant participants (the States) that OTAG not fail, partly to demonstrate that the States involved in the process could work together and partly to bound future actions by the federal government. The IPCC process was also consensus-based but as Franz (1998) points out: in the climate debate, a vocal group of skeptics emerged to disagree with the views presented by the IPCC and labeled as consensus statements on the science of climate change. There were some scientists that did not participate in the IPCC process who wished to point out that the broad consensus developed by the IPCC did not represent a consensus of all scientists knowledgeable about climate change.

The discussions in the working group suggested that consensus plays differing roles depending on whether the assessment is "insight-oriented" or "decision-oriented". Qualitative consensus can lead to subsequent challenges if the assessment is "decision-oriented". In this case, qualitative assessment could be used to mask disagreements and in doing so mask some real issues, which are subsequently uncovered and lead to a challenge to the credibility of the assessment. This would suggest that quantitative consensus would be preferable, except for assessments that are set up to support a decision that has already been made. Lack of consensus in insight-oriented assessments
is not necessarily a failure, since it could point to an area where there is a need for further research.

As we look at examples of assessment processes, we can see that consensus does not necessarily give an assessment greater legitimacy. A lot depends on the decision rules used to reach consensus, which vary because there are various ways of judging that consensus has been reached. For example, in some assessment process consensus is simply claimed, if nobody speaks up against a proposition, while in others a majority vote is accepted as consensus. Consensus can mean that "nobody argued loudly enough against a point" or it could mean "everyone felt that they could live with that point in one way or another". In the IPCC, reaching consensus on scientific issues leads to homogenization and 'a lowest common denominator' on controversial issues. This might not be the case if the issue were less salient.

The group agreed that in decision-oriented assessments, consensus should aim for a fair representation of what is known and not known. The decision rules for consensus would then be clear. It was noted, however, that reaching consensus should not be at the cost of the quality of the assessment. Especially if there is broad participation in an assessment, non-scientists should not be able, for example, to "vote" on scientific questions. Furthermore, it was pointed out that dissent can be used to the benefit of the assessment. For instance, the Netherlands has nurtured the emergence of dissenters, since they keep public interest in an issue alive. The media pays attention to loud dissenters but then often asks other experts for their views. Consensus is not so newsworthy, so the dissent, if it is "kept under control" can be used to maintain public interest in an issue. Consensus about the process of an assessment can also be as important as consensus about the outcome.

VanDeveer (1998) argues that consensus within LRTAP cooperation and assessment frequently relies on consensus between a small number of the most interested states and national research communities. That is, once the most involved parties reach consensus, the others tend to follow along. Furthermore, this “big player” consensus is recognized as such by many participants from the periphery.

While agreeing that the process of consensus-building can be important in creating a shared view of a problem (i.e. "buy in"), the group concluded that not enough attention appears to have been given in global environmental assessments to other possible approaches. For example, rather than a consensus-based lowest common denominator opinion on a scientific topic, expert elicitation (e.g. Morgan and Keith, 1995) could be used to show the range of opinions. Testing of multiple hypothesis, an approach with a long tradition in the science realm, would be another alternative approach.

7. Who participates in the scoping, conducting and communicating of assessments?

The question of "participation" has interested most working groups at the 1997 and 1998 workshops. It is also a topic that is covered in several of the GEA fellows' papers. For example, the paper by Botcheva elaborates the mechanism through which the nature of participation is likely to affect the credibility of the information being communicated through an assessment process. The study is based on information about the use and perceived credibility of different
economic assessments in Poland and Bulgaria. Botcheva shows that in Poland the inclusion of multiple political perspectives in a knowledge-building process enhances its credibility and communication power to multiple audiences. Building such a participatory assessment process involves complex interactions between relevant actors and technical experts and requires a considerable institutional capacity to facilitate such interactions without sacrificing academic quality. Domestic institutional and expert capacity played an important role in this participatory process. In summary, Botcheva’s paper shows that while wide political participation might be a commendable quality of an assessment process, it is often difficult to achieve and the ability to design a sufficiently participatory process without undermining its technical quality depends on the institutional framework within which it is embedded. These are clearly points that can be discussed further with regard to their implications for the design of future assessment processes.

Clearly the kind of participation in the various phases of an assessment depend on the goals of the assessment and on the design of the process. In some cases, special concessions have to be made to allow particular actor groups, such as environmental NGOs, to participate, while other more formal processes use public hearings, for instance. Participation can be used to maximize diversity of the actors involved in an assessment process and is of value if clarification from different groups is required. Broad participation is also seen by some as a way of allowing power-sharing. It might also be a means to assuring easier implementation of measures, by ensuring "buy-in". However, it was noted that if participation is used as a means to engage a wider group in a dialogue there might be capacity constraints. This appears to be the case for the countries of Central, Eastern and Southern Europe (“peripheral nations”), whose participation in LRTAP assessments are constrained (as documented by vanDeveer, 1998 and Farrell, 1998). Further, the breadth of participation in LRTAP seems to be exaggerated. LRTAP has institutionalized the environmental interests of northwestern European nations (e.g. Germany, Norway, and the Netherlands), especially the concern about acidification. It does not appear that the peripheral European nations had a significant impact on any meaningful decisions within LRTAP. To a large degree, this doesn’t matter to the peripheral states, however, because LRTAP is more about image than environment. That is, these nations are more interested in appearing and becoming westernized than in environmental protection, and their participation should be seen in this light.

More generally, expanding participation in an assessment process can be a way to increase input into the process, or at least appear to do so. In addition, participation could be stimulated by more political considerations, such as building a political constituency, responding to the need to shift the policy agenda or to create or remove political cover. An important motivation for participation can be self-protection. Participation can be enhanced in order to generate new insights, shape a research agenda, practices and methods, or to build new issue networks. On the other hand, there can also be reasons for limiting, rather than expanding, participation, especially if it is felt necessary to insulate the scientists from the decision-makers.

The Working Group discussed the participation in three assessment processes: NAPAP, the Skipjack Tagging and Assessment Program in the South Pacific and OTAG. Table 1 summarizes the results of the discussion in terms of the participants, the objective of participation and the actual effects.

<table>
<thead>
<tr>
<th>Who participated?</th>
<th>What was the objective of participation?</th>
<th>What was achieved?</th>
</tr>
</thead>
</table>

- in scoping
- in conducting
- in communicating participation?

**NAPAP - Enabling Legislation**
Joint Task Force (20 different Federal and state agencies, 3 Presidential appointees, DOE National Labs)

**NAPAP - operational**
Executive Director
Two teams Science and Policy

| Science group - all scientists, federal and universities | Science Group - generated considerable new knowledge. |
| Policy group - federal employees, largely determined what science got funded. | Policy Group - synthesis of "state of the science". Shaped the direction of science group by determining funding. |

EPA in 1988/9: delay the report (communicating).

Increased scientific understanding of the problem. Political appointees may create obstacles to legitimacy. Timeliness matters, but blocking the report does not mean blocking the message, other mechanisms exist-like testimony and news reports. EPA cut off the program after Title IV was passed.

**SKIPJACK TAGGING AND ASSESSMENT PROGRAM (S. Pacific)**
(an open analytic system, policy insight oriented process, formally strong boundary between science and policy realms, informally more interactions)

| Scientists - 10 to 15 from AUS, NZ, US, FR, no UK or Pacific Islanders | Insulate the scientists from the political realm. |
| Get expertise where it existed even though that meant the "predators" |
| Get good answer to the "how big" question so that the "who gets" argument could be usefully defined. |
| Develop a training program for fisheries managers. |
| No build up of capacity in fisheries sciences. |
| Australian interest in keeping the region together - provides funds |

Trained cadre of fisheries managers.
Shift in political question from "who gets" to "how to best get fast".
Institutionalized management training program.
Development of informal network of scientific advice from research team to the individual nations. This benefit goes beyond the benefit of the formal structure.

**OTAG**

| Scoping – | Objective: fulfill the ideals of federalism (state participation); industry (raise credibility – defensive posture). Some of |
| | Achieved: Better data collection, and consistent data collection across the eastern |
| States and EPA at first, industry came in late but since the OTAG process evolved up to nearly the end, industry was involved in scoping. NGOs participated in scoping in the last few months | Industry and some States were defensive, especially those in the Midwest and South. Others (mostly in the Northeast) participated in order to try and get emissions reductions from upwind sources. Reframing of issues for the governors – Technical people understood transport, but most political people did not, and OTAG became an effort to help change those ideas. | States. Connect to participation: peer representation by political appointees (if commissioners were chairing a meeting – technicians would then describe the same results and those at the original meeting serve as representatives to the wider policy group. Could go back to governors with the information that it was a credible process. Boundary of the "laugh test" shrunk. Undesirable effects?: long delay that bought you to a weak recommendation (85% to status quo). Delay through need to get 37 states to agree to this band. Technical concept of transport is understood by more people – OTAG defused a crisis situation in 1994-95 during which environmental advocates were concerned that the Congress would significantly weaken the Clean Air Act if given the opportunity. Increased the State-level understanding of what was happening in other States. |
| Conducting – States, EPA, industry | EPA Air: scoping (technical experts) – ECOS: political appointees at the top of environmental agencies. Those actively involved had backgrounds in environmental issues. Conducting: became much broader: state Air offices (technocrats) – their capabilities increase over time; contractors; variety of businesses and other actors. NGOs almost entirely absent (couldn’t afford, time or money – NRDC was key NGO which was prepared to sue EPA to prevent something like OTAG from happening) NGOs came back once the implications of the findings were discussed. Communicating: all participated, but wasn’t really communicated to the outside. 54 page exec. Summary - only exists in electronic version – not paper (and probably will never exist in that form) Some parts of report were voted on: recommendation |
The table shows that for each of the three assessment processes examined it was not always the same set of actors that participated in the scoping, conducting and communicating of the assessment. An interesting point that came out of the discussions was the way that participation influenced communication of an assessment outcome. Often the participation for this phase of the assessment process was strongly politically influenced and decisions about the content of reports or executive summaries or even about whether to have them at all, were made at this stage, so that the communication of the assessment was politically influenced.

As will be discussed again in the final section, participation is an important design criterion of an assessment. The workshop discussion showed the importance of recognizing that participation and the role it plays are not constant throughout an assessment process. The selection of participants at the various stages of the assessment process depends on the objectives at that stage and very much upon the context within which the assessment is being carried out.

8. How can assessment processes be sustained in the long term?

This question arises from the observation that some assessment processes have indeed been sustained over the long term, even when public and political interest in the issue had waned. For example, the LRTAP assessment process is currently addressing the issues related to the negotiation of the second Nox protocol, although transboundary air pollution receives little public and political attention in the 1990s. Similarly, the discussion at the Workshop of the Skipjack Tagging and Assessment Program showed how capacity had been built up slowly to sustain the assessment process. Obviously not all assessment processes should be sustained over the long term, in which case the assessment process should be closed down. Experience from NAPAP suggested that planning early for the closing down of an assessment process was an important step.

However, the working group agreed that there were a number of good reasons for sustaining at least the long-term capacity for performing assessments. It was noted that it is very difficult to recreate the interdisciplinary, integrated assessment teams to deal with the complex issues of global environmental change. Similarly, it can be difficult to replace the expertise of individuals in an assessment process. Furthermore, if the capacity for performing assessments is lost, then institutional memory is also lost, which eliminates valuable possibilities for learning from past experience in dealing with these complex issues. Importantly, maintaining capacity ensures continuing monitoring of both the environment and human behavior. At a meeting to review the World Climate Research Programme in 1997, several participants noted that the capacity to monitor the global climate system was deteriorating. It was suggested that in the year 2007, climate scientists would be able to say less about the climate of the past 10 years than they were able to in 1997. While it is true that the UN Framework Convention on Climate Change, signed in 1992, commits signatories to maintain monitoring systems and exchange data, there is no
commitment to funding of this activity. In contrast, within the LRTAP regime a protocol was signed in 1984 to provide long-term financial support for the European Monitoring and Evaluation Programme (EMEP).

Maintaining assessment capacity over the long term can, of course, also have disadvantages, primary among which are ossification and routinization. The challenge, therefore, is to design a process that can stay flexible and adaptive over the long term. The process should also be designed to avoid the temptation to 'study for study's sake' rather than provide input to a decision-making process. Furthermore, it was noted that there can be opportunity costs of maintaining a long-term capacity for performing assessments, especially in resource-scarce societies.

What factors contribute to maintaining long-term capacity? The first factor identified was "broad framing". Within the LRTAP regime, the broad frame of "air pollution" has meant that a series of protocols on sulfur dioxide, nitrogen oxides, volatile organic compounds, persistent organic pollutants and now on multiple pollutants could be negotiated. Initially, LRTAP focused almost exclusively on acidification issues, however the broad framing allowed assessors and policymakers to pursue connections between various narrower air pollution issues. Broad framing provides opportunities for creative entrepreneuring over time, although it does not necessarily come without costs. It can make the assessment process vulnerable, especially to political whim. The second factor is a recurring demand for assessment, especially on problems that states find important. In this respect, assessment processes are also sustained by legitimacy. If the outcome is seen as effective and of value, it is more likely to be sustained. As noted above, an important factor in maintaining capacity is provision of resources that are guaranteed for the long term. In this regard the source of long-term funding should also be considered -- for instance, embedding the funding for an assessment process in a standard funding agency (e.g. the National Science Foundation in the US) is better than in agencies whose mandate is broader (e.g., the Environmental Protection Agency or the Department of Energy). The investment in capacity building to do assessments (education and training) is also important. The assessment process can also institutionalize continuation. For instance, within the LRTAP regime, the protocols that have been signed have defined the next issue to be negotiated, thus setting the path for the continuation of the assessment process. The assessment process can also be maintained through new assessment challenges, as was the case with stratospheric ozone depletion in the second half of the 1980s. Of course, the maintenance of assessment capacity, like the construction of it, requires resources. While maintenance costs may be small in comparison to rebuilding lost capacity, resources for such projects remain extremely scarce in many parts of the world and in some issue areas.

9. Implications for the design of assessments

At the 1997 GEA Workshop one of the working groups derived a number of recommendations on improving the outcomes of assessments. Two recommendations on process -- inclusion of self-evaluation and careful treatment of uncertainty were considered to be especially worthy of attention in the further development of the assessment process for the climate change issue. Building self-evaluation into an assessment process would allow for an "adaptive assessment process". Other recommendations were: being aware of the political, economic, social and
scientific context in which the assessment is embedded; finding out the needs of potential users; noting past experience and institutional history; ensuring appropriate participation in the assessment process; coordinating national and international science-policy frameworks; and, ensuring effective follow-up.

The brief discussions about lessons for assessment design at the 1998 Workshop built on the recommendations from the previous year. They began by returning to the point that design depends very much on the kind of assessment that is envisioned -- is the assessment more insight-oriented or more decision-oriented? This question should be asked at the beginning of the scoping phase of the assessment process. In the early phase of the assessment it is also advisable to consider what breadth and depth of assessment would be enough to achieve the objectives of the assessment. Finding out how much detail is enough -- or too much -- requires an iterative process.

Given that different kinds of assessment process will have different design criteria, it was recognized that more complete sets of recommendations will ultimately have to be limited to "ideal types". On the basis of the discussions at the 1998 workshop, the group suggested that four particular design criteria deserve attention:

- framing
- participation
- institutional context
- science/policy interface

In addition, the participants agreed that evaluation of the process (either self-evaluation or external evaluation), handling uncertainty, treating values explicitly, resource allocation and documented post-mortem analysis of the process are important.

_Framing_ was seen to be important in attracting relevant actors into the assessment process, as well as playing a role in ensuring that capacity for assessment of complex, interdisciplinary issues is maintained on the long term. An open discourse about research questions and policy questions at the beginning of an assessment process can help ensure that a broad framing is achieved. In practice, participants in assessments sometimes see framing as a crucial concept, even the idea that a problem needs to be assessed is a frame and can be contested. Thus, participants often contest frames. Furthermore, frames also operate as limits or boundaries. They may delimit issues and stakeholders, thereby impacting participation.

_Participation_ has to be designed to achieve the objective of the assessment and take account of the fact that different kinds of participation are usually needed at different phases of an assessment (scoping, conducting and communicating). It was noted that increasing participation does not necessarily benefit the assessment process. It can reduce the assessment’s quality and/or make the assessment logistically unmanageable. Designing participation so that dissenting views can be internalized in the assessment process was also felt to be important. In practice, participation is sometimes exaggerated, especially when participation is an objective as much as a means.
The institutional context within which an assessment process is to be carried out should be taken into account in the design. For example, in an assessment process which has to involve actors from developing countries, it has often been necessary in the past to institutionalize a training program or build up research or management capacity within the assessment process. However, there have been critiques of these efforts suggesting it is not so simple to simply “add science” to societies which do not have the cultural and technological support for scientific endeavor (Miller, 1998) The importance of institutional context cautions against the common practice of transplanting assessment design ideas and models across contexts.

The science/policy interface can have a number of different characteristics in assessment processes. For example, in some assessments there is regular interaction between the scientists and decision-makers participating in the process with a two-way exchange of information, while in other assessments the scientists and decision-makers interact rarely, if at all, and the interaction can also be unidirectional. Especially given the complex issues of global environmental change and the `turbulent` policy realm in which decisions about these issues have to be made, careful consideration has to be given to the design of the interface.
REFERENCES

Botcheva 1998 GEA paper


Farrell and Keating 1998 GEA Paper

Franz 1998 GEA paper


Patt 1998 GEA paper

Price (1965)


Vandeveer 1998 GEA paper
Roger Kasperson develops an idea of the “hierarchy of trust” – people trust the system of power and decision-making but maybe not today’s occupants of that system – which might be useful here. (reference?)

Leen Hordijk presented this example. He pointed out that although this selection was based on a value judgement, the participants did not seem to be aware that it was at the time.