CHAPTER II

WORKING GROUP 1-THE FORM OF ASSESSMENTS

II. 1 Working Group 1 Background Paper

Explaining the form of assessments: Why do we get the assessments we do?

Edward Parson, Shardul Agrawala, Anthony Patt, Robert Keohane, Ronald Mitchell, Liliana Botcheva,

William Clark, Elizabeth DeSombre, James McCarthy, and Eileen Shea

ABSTRACT

This paper provides background for the discussions of Working Group 1 at the Summer Workshop of the Global Environmental Assessment Project. It considers why we see the variations in assessment form that we do. In particular, it examines how the context in which assessments take place interact with assessors' decisions over time to yield the form that assessments take.

The paper sketches a classification of assessments in terms of their substantive, procedural and contextual characteristics. From an analysis of a wide range of climate assessments, it then advances a number of propositions concerning the factors affecting the degree of consensus an assessment is likely to entrain, the credibility of an assessment for various groups, the ways in which an assessment is likely to address policy issues. Further observations are advanced on the likely implications for assessments when the political saliency of an issue increases, or the range of sponsoring organizations, interests or countries increases.

Finally, the paper poses a series of potential discussion questions for the Workshop regarding the negotiation of assessment content between assessors and their sponsors, the implications for assessment of issue maturation (for example from a purely scientific to a political question), and the role of entrepreneurship and leadership for the assessment process.

ACRONYM LIST

CIAP - Climatic Impacts Assessment Program

DOE - United States Department of Energy

EPA - United States Environmental Protection Agency

FCCC - Framework Convention on Climate Change

GCM - general circulation model

GEA - Global Environmental Assessment

GFDL - Geophysical Fluid Dynamics Laboratory

GHG - greenhouse gas

IGO - Inter-governmental organization

IPCC - Intergovernmental Panel on Climate Change

NAPAP - National Acid Precipitation Assessment Program

NGO - Non-Governmental Organization

NRC - National Research Council

OTA - Office of Technology Assessment

SCOPE - Scientific Committee on Problems of the Environment

SEI - Stockholm Environment Institute

UNEP - United Nations Environment Program

WMO - World Meteorological Organization

WRI - World Resources Institute

1. INTRODUCTION

This paper serves as background and provocation to the discussions of one Working Group at the Global Environmental Assessment Project's June 1997 workshop. It is motivated by the observation that environmental assessments have very different characters. Even within a single issue such as climate change, assessments have varied greatly in the components of the problem on which they have focused, how they have organized and discharged their work, what results they have presented, and how they have marketed their results. We contend that these variations in the form of assessments are important, and are not mere accidents. They may suggest why different assessments of the same problem sometimes reach different conclusions. They

may also help to explain why different assessments differ in their influence on policy, or in some cases in their contribution to the advancement of knowledge.

The Working Group will pose the question, why do we see the variations in assessment form that we do – or with more detail, how does the context in which assessments take place interact with assessors' decisions over time to yield the form that assessments take? This background paper seeks to stimulate discussion on these and related questions. We draw primarily on assessment experience for global climate change, the focus of the project's first year of research. Assessment experience in other related issues, in particular acid deposition and ozone depletion, is also considered.

2. ASSESSMENTS AND THEIR FORMS

While any attempt to define "assessment" inevitably yields some ambiguity and some difficult border cases, a workable definition is necessary to proceed with attempts at explanation and analysis. The GEA project defines assessment as the entire social process by which expert knowledge related to a policy problem is organized, evaluated, integrated and presented in documents to inform policy or decision-making. This definition has several important implications. First, assessment includes both products and reports, and the processes that generate them (e.g., the management and organization of the endeavor, who participates, and how assessment reports are reviewed). Second, assessment bridges expert knowledge and policy; it draws on the current state of relevant knowledge, usually from the natural sciences but also potentially from engineering, economics, or other disciplines, *and* it seeks to shape policy debate or inform policy makers. We confine our scope to activities that have this bridging character. For example, we would not count as assessment a special issue of an interdisciplinary scientific journal such as *Climatic Change* o

Many past climate-assessment efforts fall within this definition: e.g., NRC (1977, 1979, 1983, 1992), DOE (1985), EPA (1983, 1986, 1989, 1990), OTA (1991, 1993) and others at the US national level; and the World Climate Conferences (WMO 1979, Jaeger and Ferguson 1991), the IPCC (1990, 1992, 1994, 1995) and others at the international level. Assessments conducted by non-governmental organizations (NGOs) such as SCOPE (1986), WRI (Mintzer, 1987) and SEI (1990) are also included.

Assessments such as these show great variation in the questions they address, the procedures by which they operate, and the substance of the outputs they deliver. Indeed, assessments vary so much that a central challenge is to identify a short list of characteristics that capture the most important variation.

This section sketches one such list of characteristics, which is illustrated in Figure 1. We present these dimensions of variation in two groups, those which are primarily substantive and those which are primarily procedural. This division is principally a convenience of exposition, in that some of the dimensions combine elements of substance and procedure. The goal of this section is to classify variation across assessments, in order to prompt discussion both about how assessments vary, and the causes of such variation.
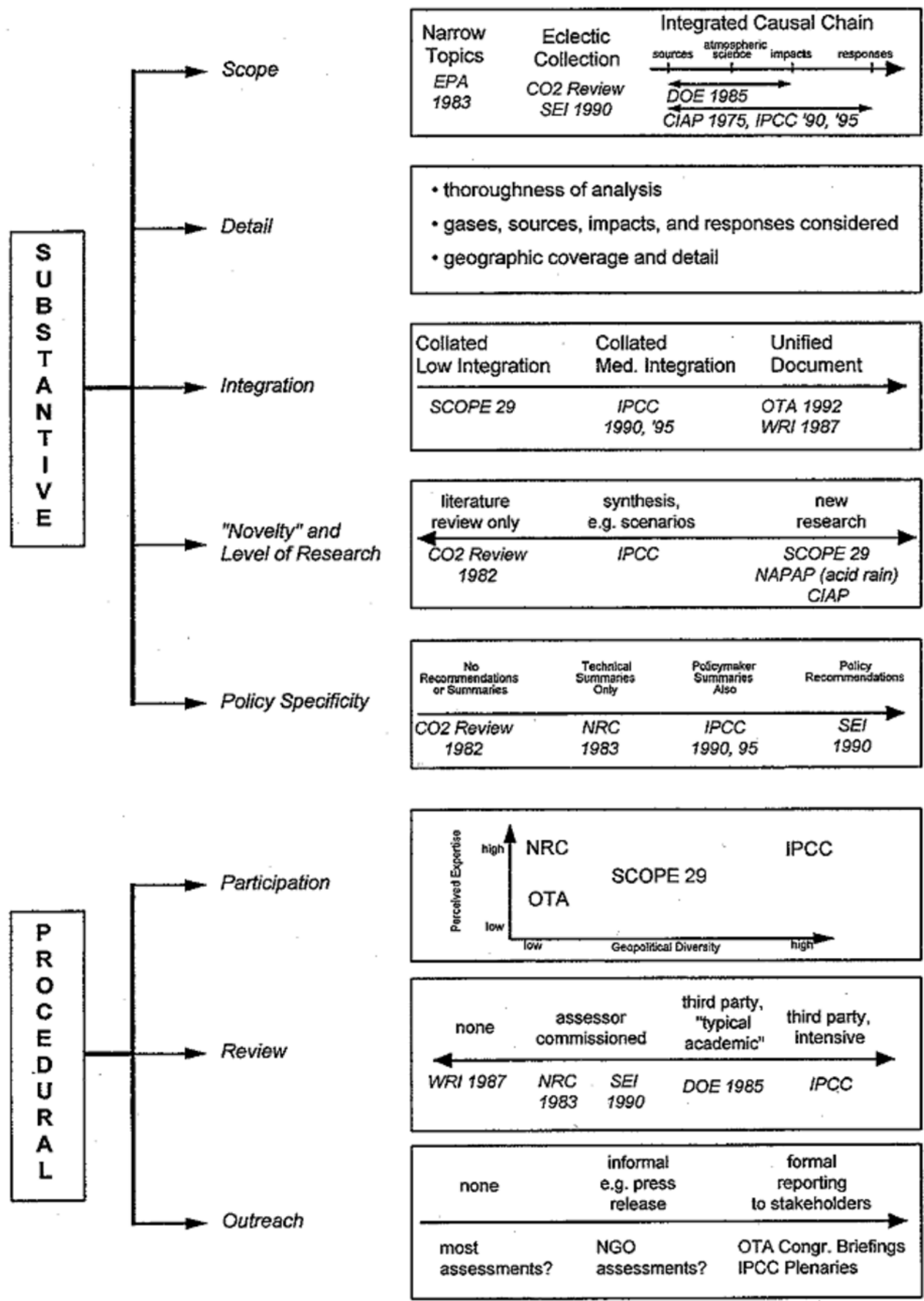
Figure 1—Assessment Characteristics. Source: Agrawala, Shardul and Anthony Patt, "Variation in the

2.1 Assessment Characteristics: Primarily Substantive

Scope: Assessments vary widely in which elements of an environmental problem they include and exclude. Climate change can be represented by a crude causal chain that extends from greenhouse-gas emissions, to atmospheric concentrations, to changes in climatic variables, to impacts on natural and socioeconomic systems. Possible response strategies might reduce emissions, reduce vulnerability to impacts, or intervene at intermediate points in the causal chain. CIAP (1975) and IPCC (1990, 1995) include the causal chain from end to end, while other assessments consider various sub-sections. For example, EPA (1983) focused narrowly on the feasibility of greenhouse gas mitigation, while DOE (1985) stressed the upstream elements of the causal chain (emissions-concentrations-climate change), with limited discussion of impacts and none of responses. Other assessments consider selected topics at various points along the causal chain, e.g. the Carbon-dioxide Review (Clark, 1982), and the reports of the Stockholm Environment Institute (1990).

Detail: Assessments also differ in the detail they accord to each component they do include. There can be more or less detail in the thoroughness of scientific argument presented; in the breadth of emissions, source activities, impacts, and response options considered (e.g., emissions of CO_2 vs. all greenhouse gases; impacts from sea-level rise, agriculture, or other sectors; specific response options for abatement, adaptation, or geoengineering); and in the geographic scope and detail at which sources, impacts, and responses are discussed (e.g., global, regional, national, or grid-cell).

Integration: Assessments vary in the extent to which the information they present is made consistent and integrated across the elements of the causal chain that they consider. Few are truly unified documents. Rather, most collate contributions on different pieces of the problem by different experts or teams. The degree of internal consistency imposed on these separate contributions can be low or high. Some, such as SCOPE (1986) have *low integration*, assembling largely uncoordinated chapters from separate authors. Others, such as IPCC (1990, 1995) have *medium integration,* employing various devices to promote unified and consistent treatment among chapters, and providing integrated technical summaries and synthesis reports. Assessments with *high integration* would include staff assessments in which one or a few authors collaborate closely to write the entire report (e.g., Mintzer 1987, OTA 1991), or assessments based on a single formal integrating model.

"Novelty" and Level of Research: Assessments vary in whether, and how far, they go beyond reviewing existing knowledge and literature. Some are primarily literature reviews, such as the CO_2 Review (Clark, 1982); some develop new interpretation or syntheses based on available literature, such as climate-change scenarios (e.g., IPCC 1990, 1992) or identify research priorities; still others actively undertake or present new research (e.g., SCOPE 1986 presented new research conducted at the International Meteorological Institute, Stockholm; both CIAP (1974) and the NAPAP (1991) acid-rain program combined assessment and research within one institution).

Policy Specificity: Some assessments make little attempt to express their results in policy-relevant terms (e.g., DOE, 1985). Others attempt to summarize or synthesize the information they present in a way that they believe to be policy-relevant, sometimes through summaries (e.g., IPCC 1990, 1995), while avoiding explicit policy recommendations. Only a few assessments, mostly done by NGOs, make explicit policy recommendations.

2.2 Assessment Characteristics: Primarily Procedural

Participation: Who participates as assessment authors and how they are chosen can vary in several ways: number, stature, and disciplinary and geopolitical diversity. Many assessments divide into two broad categories in their participation: *invited panels* of eminent independent scientists, e.g. the IPCC and NRC panels; and *staff assessments* that draw on in-house expertise and paid consultants, e.g. OTA, WRI, and SEI. Invited panels usually involve more, and more senior, participants than staff assessments. Author diversity often reflects a combination of the scope of problem the assessment is addressing, and among whom the assessment seeks to achieve credibility. Thus, NRC panels, which seek credibility in US policy-making, pursue some required diversity of disciplinary participation, but do not need international diversity; the IPCC, which also seeks global

political credibility, also requires diversity of international participation.

Review: Assessments vary in the extent of their external peer review. Staff assessments such as OTA, WRI, or SEI may have no review or may rely on informal internal review procedures. Other assessments conduct review processes of varying levels of transparency and rigor. For example, NRC (1983) had about 30 "volunteer" reviewers; DOE (1985) conducted a third-party review similar that employed by a journal administered by the AAAS; and the IPCC conducts an intensive third-party review process involving hundreds of reviewers.

Outreach: Assessments devote various levels of effort to outreach, and have various channels available to them. One important distinction is between assessments that are mandated by policy-making bodies and those that are not. Mandated assessments often require formal presentation of their findings. For example, OTA reports were submitted to their commissioning Congressional Committees; the first IPCC assessment was presented to member governments at the IPCC plenary, and at the 1990 World Climate Conference; IPCC assessments are now also informally presented to the Climate Convention bodies. Un-mandated assessments, such as Mintzer (1987) and SEI (1990), typically lack such a designated formal recipient, and are instead disseminated informally according to the energy, resources, and contacts of their proponents. Whether an assessment is mandated or not, the effort its principals devote to press releases, briefings, presentations, educational events, and pursuit of media attention vary greatly, and may be strongly associated with the assessment's subsequent influence.

3. THE CONTEXT OF ASSESSMENT

Assessments take place in the context of existing knowledge, research communities, politics, institutions, and history. Although an assessment's context cannot fully determine how it is done, context does influence assessment in various ways. For example, the existing set of institutions or political traditions may determine who is deemed qualified to undertake an assessment. National or institutional factors determine the resources available to do assessments, and the competing demands on the time of those who are qualified to do them. The current politics of an issue may influence what questions are included in an assessment's mandate or who participates. The current state of knowledge limits what questions can be answered—with what confidence, precision, and level of consensus—however strongly political actors may desire answers.

This section seeks to categorize and operationalize the elements of context so as to facilitate analysis of how context influences the substantive and procedural characteristics of assessment. Although this analysis focuses on systematic features of context, one should also recognize the importance of idiosyncratic and chance events on assessment form. The form and influence of an assessment may depend strongly on the vision or energy of individual leaders or participants, or on random, high-profile external events like droughts in agricultural regions or icebergs calving in Antarctica. Still, in our investigations we should try to avoid idiosyncratic explanations, as useful insights and advice require some level of generality.

3.1 Knowledge Context

For any particular assessment, certain bodies of knowledge will be deemed more relevant, others less so. This choice is not automatic, but interacts with choices regarding the character of an assessment. For example, whether an assessment broadly characterizes the climate issue as one of variability or of global trends will determine which bodies of knowledge are relevant, as will decisions regarding the assessment's scope and depth. In turn, the current state of particular bodies of knowledge can influence decisions regarding the scope and depth of assessments. For particular areas of knowledge, these interactions are likely to be influenced by several characteristics: whether it is data-rich or poor (e.g., we see a dearth of developing-country climate-impact studies); what issue-framing dominates (e.g., climate variability at the time of the First World Climate Conference in 1979 turned to climate change in later years; there was a shift from a "CO2 problem" before about 1983 to "CO2 and other greenhouse gases (GHGs)" in later years); whether uncertainty or disagreement

is high or low; whether the area is stable or changing rapidly; whether particular methods dominate current research (e.g., there are constraints on regional climate and impacts information due to downscaling limitations of General Circulation Models (GCMs); and whether the associated disciplines and research communities are large or small, rich or poor in resources, and high or low in scientific status.

3.2 Policy Context

Political agendas regarding the issue at hand – their current state, evolution to date, and recent changes – can affect the form of assessments. Important aspects of an assessment's policy context may include: the salience of the issue; whether decisions about the issue are on a current policy agenda; the contentiousness of proposed decisions; the existence of institutions responsible for managing the issue; and the history of prior decisions, including both general political declarations and decisions on specific actions.

Policy context can shape both substantive and procedural aspects of an assessment. For example, once the debate shifted to calls for realistic international responses to climate change in the mid 1980s, it became evident that developing-country participation was crucial for any successful international climate regime. This recognition contributed to the establishment of the IPCC, and to many aspects of its structure and process such as broad international participation, extensive peer review, and government approval of policy-maker summaries.

3.3 Stakeholder And Institutional Context

Any assessment has a local context consisting of relationships with particular outside actors, who may stand in various relationships to it. The outside actors may be *authorizers*, who grant formal permission or institutional authorization for an assessment to proceed; *sponsors*, who provide financial or other resources; *recipients*, to whom assessment results are formally presented; or *users*/*audiences*, a broad residual category of those who might use the assessment in various ways (e.g., other political actors who might use it to advocate or oppose certain policies or to advocate delay, scientists or research managers who might use it as a reference or to formulate and fund research priorities, or the public at large).

For any particular assessment, these roles may be filled by the same organization, different ones, or none. For example, the IPCC has high stakeholder differentiation: its authorizers are UNEP and WMO; its sponsors are UNEP, WMO, and about fifteen national governments; its recipients are about 150 member governments and the FCCC bodies; and its users/audiences are many political actors, firms, industrial and advocacy groups, academics, national research programs, etc. At the opposite extreme, an organization may authorize and sponsor its own assessment, with no formal recipient and an undetermined set of potential users/audiences, such as WRI (1987).

Many assessments are conducted under the auspices of government agencies, NGOs, and international organizations. These settings, like the differentiation of stakeholders, can influence the form of an assessment. First, there might be implicit inducements for the assessment to support the position of the host institution. This is especially true for assessments done by advocacy groups, but might also hold for certain government assessments. Second, the disciplinary mix and levels of expertise of "in-house" assessors might influence the scope and quality of assessments. For example, this might explain the absence of a discussion on socio-economic impacts and response options in DOE (1985).

4. UNDERSTANDING CAUSAL RELATIONSHIPS: TWO FRAMES FOR ANALYSIS

In the following sections, we sketch a series of arguments, hypotheses, and questions to help account for variations in assessment form. In this section, we identify two approaches to explanation: static and dynamic analyses. In the following two sections, we propose a few hypotheses using each approach, and

identify a series of questions for further discussion in the Working Group.

4.1 Static Analysis

To explain assessments in static terms, we seek to account for variation in the form of assessments in terms of characteristics of their context, at the time they are done. We ask, for example, if different forms of assessment are undertaken when the underlying knowledge base is stronger or weaker; when the issue is more or less politically salient or contentious; or at earlier or later stages in the history of policy-making on the issue.

In addition to these broad aspects of their context assessments are also, perhaps primarily, shaped by active managerial decisions by, and negotiations among, assessment managers and participants, and the outside stakeholders who authorize, sponsor, and receive the assessment. These actors have interests that they seek to advance through the assessment; and they have discretion to make and negotiate choices regarding its form, content, and process. The discretion accorded to assessment managers to design their own assessment may range from very little to near-total. Skillful assessment managers may actively seek to be given certain instructions and not others, or to be asked certain questions and not others. They may try to anticipate likely stakeholder response even when not under explicit instructions. To further complicate matters, particular individuals may change roles, moving between assessment bodies and various outside stakeholder bodies over time.

For example, some aspects of the IPCC were imposed by its outside stakeholders, e.g., its intergovernmental nature, comprehensive scope, and global participation. Many of its features, however, reflect strategic decisions by the Chair and Bureau, e.g., the elaborate peer review and the progressive shift in products toward interim reports and Technical Papers (Agrawala, 1997).

The interests of both assessors and outside stakeholders, and the interactions between them in negotiations over the establishment and form of an assessment, can be complex. Through an assessment, outside stakeholders may seek to speed or slow the pace of policy-making; to highlight particular components of the issue (e.g., sources, impacts); to support or oppose particular responses; to inform policy or improve deliberations; to increase, reduce, or channel dissent; or to keep (or avoid) control over the issue or decisions to be made on it. Assessors may seek some of these same goals. Assessors may also seek to attract policy attention to an issue; to keep specious scientific dispute out of policy debates or transform political deliberations in the supposed image of science; to advance relevant knowledge or participate in an intellectually rewarding activity; or to gain enhanced status or resources for individuals, disciplines, institutions, or the scientific community at large. Since outside stakeholders usually cannot fully control an assessment, the

Both assessors and outside stakeholders must manage tensions between their various desired goals. Assessors may have to balance their desire to attain policy relevance with their need to defend scientific credibility. Outside stakeholders may have to balance their desire to ground their preferred policy positions in positive "scientific" claims with the risk that their claim will be shown wrong. Stakeholders may also have to balance their desire for a preferred answer from an assessment with the need to give up control over an assessment if it is to be credible; or their desire to retain control over an issue in a political body with their desire to dispense with persistently divisive questions by calling them "scientific" and asking an assessment body to resolve them.

4.2 Dynamic Analysis

Examining assessments in dynamic terms, we seek to identify trends in assessment substance and procedure over time. We also seek to account for their evolution in terms of prior decisions and events, and to understand how early choices may influence subsequent assessment trends. The evolution of assessment over time may be path-dependent, and may reflect feedbacks that drive the process toward certain extreme forms, or to equilibria.

For example, the reputation of assessment participants and sponsors creates positive feedback in credibility of

assessment processes over time. The decision of the IPCC founders to invite an eminent and apolitical scientist to chair the process made many other leading scientists willing to participate, and set in motion a self-reinforcing cycle lending high credibility to the institution. In contrast, the reputation of advocacy organizations may render them unable to attract eminent mainstream experts to participate in assessments they sponsor *even if they intend to run an objective process*, allowing them to draw only those who are willing to be associated with the organization's ideology. Resultant assessment bias would compound the organization's inability to attract the experts necessary to conduct a credible assessment. In these cases, we see assessment credibility *tipping* toward one extreme or another. Which direction the tipping goes may depend on contextual factors—e.g. levels of funding, pr

5. HYPOTHESES ON THE FORM OF ASSESSMENT

This section presents a few hypotheses about how contextual factors and the choices of assessors and outside stakeholders interact to shape the characteristics of assessments. These hypotheses are intended to be sharply drawn, provocative, and plausible. In drafting these, we have sought to avoid conjectures that appear to be obviously true but of little practical importance, favoring instead statements that are more arguable, but might have significant implications. In its discussion, the Working Group is invited to critique, revise, or drop any of them, and to propose new ones to add.

5.1 Static Hypotheses

The state of relevant knowledge and of the research communities that hold it, shape and limit what any assessment can do or say. Some aspects of this effect appear to be obvious. For example, some minimum threshold level of consensual knowledge on particular questions is surely necessary for an assessment to say anything about them. When this threshold is not met, we would expect to see no assessments, or assessments that avoid the territory in question. As understanding advances, it may become possible for tightly managed staff assessments to speak to the questions before broader consensus assessments are able to do so.

The more interesting aspects of the influence of knowledge context on assessment will likely be those that arise when the state of relevant knowledge is adequate for assessment to proceed. Here, we propose a hypothesis regarding the influence of the social structure of the relevant research communities.

Hypothesis 1:

> When the social networks of the research communities holding knowledge relevant to an assessment are more tightly connected:

a. it is easier to obtain consensus within the network, because the signals sent by reputation are stronger, so group members know who to defer to on specific questions. [Note that this claim admits a sharply drawn contrary: that the tightest networks make for the most intense rivalries, and hence for the greatest persistence of nit-picking disagreements.
b. it is harder for new entrants to gain credibility within the associated assessment community, because reputation is self-reinforcing and newcomers lack it. Consequently, there may emerge stable hegemony of certain prominent actors (e.g., certain climate modeling groups such as GFDL; or certain assessment processes such as IPCC). Since newcomers find it difficult to gain entry into the community based on their expertise, they are more likely to enter based on claims of breadth of participation, and political legitimacy.

Aspects of the policy context will also shape and constrain assessments in various ways. As with the knowledge context, some of these effects may be obvious but unimportant, others more important but more contestable. For example, it appears obvious that the resources necessary to do assessments become increasingly available

as the issue appears on political bodies' agendas and grows in salience. Consequently, very early in an issue's history, we expect that if assessments are done at all, it will be by advocacy groups trying to get the issue onto an agenda; or in some cases by scientific bodies trying to attract policy attention. (And the important aspects of these two cases may be similar or different. Issues differ in their early histories: some cook for a long time in scientific circles, until many scientists come to think that it holds high policy relevance; others are promoted from very early by advocacy groups.)

As an issue develops and its salience grows, assessments may be sponsored by advocacy groups or by political bodies (governments and IGOs). The interests of these two kinds of groups in doing an assessment, and the kinds of assessments they do, are likely to differ sharply. Advocacy groups normally have strong and clearly defined policy preferences, and sponsoring an assessment is one of several means they may employ to advance their preferred policy agenda. On the one hand, these groups lack the stature and convening power of governmental bodies, and have a specific policy goal in mind. Consequently , they will likely neither want, nor be able, to attract eminent independent scientists to participate in a consensus panel assessment. On the other hand, to help them advance their agenda an assessment must not only support the agenda, but must also attain more widespread credibility than the group could gain for its cause through some other means. Pursuing broader credibility will typically require some movement from the group's preferred policy position toward a broader scientific or political consensus. From this argument, we propose a hypothesis about the kind of assessments that advocacy groups will sponsor.

Hypothesis 2:

> When the authorizer/sponsor of an assessment is an advocacy group:
>
>> a) The assessment will be undertaken through staff or consultants;
>>
>> b) Its policy recommendations will be explicit, or easily discernible;
>>
>> c) Its contents will be less extreme than the group's policy stance as represented in other forms of advocacy.

Part c) appears to be the most interesting part of this hypothesis. It suggests the possibility that highly polarized issues may tend to become less so to the extent that advocacy groups choose to pursue their goals through commissioning assessments and hence seeking wider credibility, rather than through grassroots, litigation, or lobbying campaigns.

As salience grows, assessments are increasingly likely to be sponsored by bodies that hold policy responsibilities, and that are large, heterogeneous, and divided. Such bodies are likely to include some, perhaps many, members who lack the capacity for independent substantive evaluation of the content of an assessment's report. These people must consequently decide whether to trust an assessment according to its institutional affiliation. Consequently:

Hypothesis 3:

> When the issue on which assessment is undertaken is politically salient and contentious, only assessments authorized by the most inclusive level of relevant political authority can gain credibility with all significant actors.

While such authorizing bodies embed highly heterogeneous, divided political interests, they do share some common goals in undertaking an assessment. While they are most unlikely to share a specific preferred outcome, they may share some interests in assessment form and output, at least in advance of the assessment being conducted. For instance, they may share a desire for broad enough participation to prevent dominance by a few, or the most extreme, actors, even if this detracts from the efficiency or "quality" of the assessment

science. Alternatively, they may wish to use the assessment to resolve persistently divisive questions. For the assessment to resolve disputes, it must be designed so that its outputs are widely perceived as credible. This in turn requires that the assessment sponsors renounce the opportunity to exercise significant control over its substance, through various aspects of procedural design that protect the assessment's independence. Certain forms of procedural control (e.g., requiring broad participation, transparent rules of procedure, or consensus process) are necessary to guard against attemptsuse they tend From the perspective of assessors, their sponsors' desire to give away contentious questions can pose serious problems. While diverse sponsors may easily agree *in advance* to give away a divisive question, they may still be expected to protest those specific answers that they do not like. Assessors must consequently balance their interest in serving their sponsors and providing policy-relevant answers, with their need to maintain scientific credibility. Maintaining this balance may require that they limit their results to those points on which very solid and wide consensus among relevant experts is attainable. It may also require that they decline to answer questions posed to them, when they judge that they are unable to secure such consensus on them.

The interactions among these interests of political bodies who authorize and sponsor assessments, and of assessors, yields:

Hypothesis 4:

> As the authorizers and sponsors of an assessment are more multiple, heterogeneous in interests, and divided in their preferred policies:
>
> > a) assessments will have broader participation, and more formal and explicit rules of procedure;
> >
> > 1. b) assessments will more likely use consensual processes;
> >
> > c) assessments will more likely have vague mandates, and few explicit substantive constraints on their work;
> >
> > d) assessors will restrict the scope of the assessment to obtain stronger consensus;
> >
> > e) assessors will more likely decline to answer some questions posed to them;
> >
> > f) assessments are more likely to drift away from policy relevance toward questions of scientific interest.

5.2 Dynamic Hypotheses

As argued above, various self-reinforcing factors can operate in the early period of an assessment process to drive it toward either high or low scientific credibility. We can describe a plausible path that yields high credibility. If the most powerful actors involved in initial assessment design believe that their preferred policy positions have legitimate scientific foundation, they will be more likely to push for the appointment of eminent and independent scientists to lead the assessment. These scientists, who have valuable reputations to protect, are likely to defend aggressively a science-dominated assessment process and resist attempts at political interference. Other scientists of stature will be more likely to participate if the leaders and initial participants are eminent and independent, because this signals that the process is likely to be intellectually rewarding and reputation-enhancing. Consequently, the more scientifically credible the process, the stronger the incentives for any actor who wishes to influence it to do so according to scientific norms -- sending their best scientists, and not attempting political influence over them. An equivalent but opposite self-reinforcing process in the early days can push an assessment toward low scientific credibility. A question for further research is whether assessments that follow a downward spiral of credibility simply die out, or continue in some modified form, perhaps being claimed and

Franz, Wendy E. "The Development

Kandlikar, Mi

National Research Council (NRC). 19

Scientific Committee on Problem

World Meteorological Organization. 1986. *Atmospheric Ozone*

Hypothesis 5:

> *Initial conditions and early events in the establishment of an assessment process drive it toward equilibrium at either very high, or very low, levels of scientific credibility. Assessment processes do not persist at intermediate levels of credibility.*

Finally, within the GEA project various hypotheses have been proposed regarding how continuing assessments, or repeated assessments on the same issue, tend to broaden or narrow over time -- in their scope; in the number of sources, impacts, and response options they consider and the depth with which they consider each; and in their participation. These hypotheses have taken several forms.

Hypothesis 6: "Broadening and narrowing"

> *Continuing assessment processes on the same issue over time, or repeated assessments done for the same domain (e.g., US, International) on the same issue, tend to:*
>
> > *a) **Narrow in scope**. Early over-reaching toward policy-relevance by politically inexperienced natural scientists coming into the issue is often followed by increasing contraction of causal-chain of analysis, as assessors attempt to develop bulletproof consensus.*
> >
> > *b) **Broaden in participation**. Increased political salience and contentiousness over time generates increasing demands for broad national and sectoral representation.*
> >
> > *c) Policy assessments, when undertaken at all, get stuck for long periods doing narrow comparisons between a small number of competing technological options.*

6. PROPOSED DISCUSSION AREAS FOR WORKING GROUP

This paper has identified a few dimensions of the form of assessments and the context in which they are done. It then has proposed a few static and dynamic hypotheses about relationships between assessments and their context.

This paper is intended to launch discussion of the Working Group, without constraining it. In particular, participants are invited to consider whether the dimensions of form and context identified in the paper are the important ones, whether others should be added, and whether the distinction between static and dynamic effects is justified. Are the proposed hypotheses reasonable, and are their predictions interesting and important? Participants may also want to consider additional hypotheses drawn from their experience with assessments and policy-making.

At this stage of the GEA Project's work, it would be premature to offer conclusions. Instead, this closing section sketches a few questions and areas of inquiry that may be appropriate areas for investigation either in Working Group discussions, or in the next year of Project research. These address aspects of the

relationship between assessments and their context that are of clear importance and on which the Project's initial research has identified preliminary questions or provocative observations, but on which we do not have enough, or specific enough information yet to present hypotheses.

6.1 Negotiations Over Assessment Form

Assessors and outside stakeholders may be motivated by different interests and have different tensions to manage. At times, the resultant negotiations have been problematic. Conflicts between assessors' and sponsors' objectives, or the appearance of such conflicts, can undermine the credibility of the assessment process. More empirical work to elaborate the character and consequences of these negotiations could be of high value. Questions for the Working Group to consider might include how such tensions have arisen, and how and with what consequences they have been resolved.

For example, in what specific ways have outside stakeholders tried to shape or constrain assessments? When have outsiders tried to broaden or narrow the questions that are addressed? When have they tried to require that particular things be included or excluded?

In what specific ways have the goals of assessors and outside stakeholders diverged? In what ways, and for what reasons, have outside stakeholders not obtained what they wanted or expected from assessments? How do outside stakeholders think about how much control to give up to an assessment? How, if at all, do they think about balancing their interest in shaping the result with their interest in a credible assessment that reduces policy dissent?

Where have the conflicts between stakeholder demands and scientific process been sharpest? Where have different assessment leaders drawn the line, with what consequences for the quality, policy relevance, credibility, and subsequent effects of the assessment?

How does the balance of discretion between assessors' and outside stakeholders shift over the lifetime of a single assessment and over repeated assessments? How, and how actively, do assessors re-negotiate their mandate? Are there consistent patterns in what they seek to revise and why?

One particularly interesting aspect of negotiations over assessment form concerns the presence or absence of policy recommendations in assessments. It appears that explicit policy recommendations in assessments are extremely rare. Under what conditions will assessments make policy recommendations, or other normative policy-relevant statements? What are the likely consequences of doing so? When assessments do assess policy options, how have assessors managed the tradeoffs involved in influencing policy decisions? The Project's research, and other analysis, all suggest that assessments are most likely to be cited, and used to substantiate decisions, if they confine themselves to options presently on the policy agenda, follow prevailing framings, and support the prior policy preferences of the recipient. To follow this path precisely is to renounce any attempt at influence, but to diverge too far from the present policy debate is to risk being unpersuasive, or even incomprehensible, to the policy actors. How have assessors attempted to manage this trade-off, to what effect?

6.2 Assessment And Research

Some assessments have explicitly articulated research priorities, or undertaken research themselves, while most have not. Plausible arguments can be constructed in favor of either stance. On the one hand, assessment processes are likely to be well situated to see most clearly what the primary policy-relevant research priorities are, and it may be more efficient for them to support such research directly than for them to wait for somebody else to pick it up. On the other hand, allowing assessments to shift their effort toward conducting, or even planning, research may risk having assessment drift toward the most scientifically interesting questions and away from policy relevance. Subsequent discussions and investigations might fruitfully ask under what conditions assessments should, or are more likely to, explicitly articulate research

priorities, or even undertake research themselves. How does this strategic assessment choice affect assessors' attempts to manage the tradeoff between policy relevance and scientific credibility?

6.3 Entrepreneurial Assessments

The vision of assessment as a bridge between science and policy suggests that assessments undertaken without some form of mandate or sponsorship from the policy community are at greater risk of being ineffective. But such assessments are occasionally undertaken, and some appear to have had significant effects on policy. Possible examples might include the ozone "Blue Books" (WMO, 1986) or the Villach proceedings (UNEP, 1985). What are the conditions that lead scientific bodies to undertake self-initiated assessments without official mandates? Under what conditions are such assessments likely to be effective, and by what means?

6.4 Assessments And Issue Framing

Issues are framed in particular ways that change over time, including the basic character of an issue (e.g., climate variability vs. global trends), the linking and un-linking of substantively related issues (e.g., climate change with stratospheric ozone depletion, acid deposition, and tropospheric air pollution), and the set of particular sources, impacts, and responses that are considered relevant.

The Project's research might usefully investigate the relationships, over time, between assessments and broad prevailing issue framing. Do assessments always follow prevailing framings? How is the scope of assessments shaped by prevailing policy framings and by the current state of knowledge? Under what conditions, if at all, do assessments make significant changes in how the issue is framed? Under what conditions do assessments make significant innovations in assessment methods? How do outside stakeholders constrain innovation in assessment, with what consequences?

What strategies have assessors used to shift the focus of assessments over time, check for continued relevance of current focus areas, discard those that are no longer useful, and incorporate those that are becoming important? How does maturation of bodies of relevant knowledge affect their inclusion and treatment in assessment? How do assessments influence shifts in popular and political framing of the issue?

6.5 Non-Assessments: Why Assessments are Done or Not Done

The broad question for this Working Group, "why do we see the assessments we do" really has two parts: accounting for variation in form among assessments that are done; and explaining why assessments are done on certain issues at certain times, and not others. On this second question, the Project's initial round of research has little to say, because we have little data on cases where assessments might well have been done but were not. But the Working Group might consider situations in which plausible or desirable assessments were not done, and ask what aspects of the context -- whether knowledge, politics, or institutions -- appeared to be responsible for their being blocked. Are there prominent instances of issues on which we would expect to see assessments but do not? For example, why was no US NRC assessment done on climate between 1983 and 1992, and no US Executive Branch assessment other than DOE (1985)?

6.6 What Kinds Of Assessment Are Appropriate At What Times?

Are there particular forms of assessment that are most appropriate, or most likely to contribute usefully to policy-making, at particular stages in the development of a policy issue? e.g., raising initial policy salience; helping to frame issue; identifying most important risks, impacts, sources, responses to include; supporting the choice and design of particular responses.

6.7 The Role Of Leadership In Assessment

What are specific opportunities for the exercise of leadership in assessment? Are there specific recognizable

types, with identifiable influences? How much can we understand it in a generalizable way?

REFERENCES

Agrawala, Shardul. "Explaining the Evolution of the IPCC Structure and Process." Environment and Natural Resources Program Discussion Paper E-97-05, Kennedy School of Government, Harvard University, August 1997 and International Institute for Applied Systems Analysis Interim Report IR-97-032/August.

Agrawala, Shardul and Anthony Patt, "Variation in the Form of Climate Assessments," Unpublished Manuscript, Harvard University, May 1997.

Cash, David. "Local Response to Global Change: Information Transfer and U.S. Agriculture." Draft Paper for the Global Environmental Assessment Project, Belfer Center for Science and International Affairs, Kennedy School of Government, Harvard University, June 1997.

Climate Impact Assessment Program (CIAP), US Department of Transportation 1975, Development and Accomplishments 1971—1975. Washington, DC: US Government Printing Office.

Clark, W. (ed.). Carbon Dioxide Review: 1982, New York: Oxford University Press.

Department of Energy (DOE), 1985. State of science reports. (4 vols). Detecting the Climatic Effects of Increasing Carbon Dioxide, DOE/ER-0235; Characterization of Information Requirements for Studies of CO2 Effects: Water Resources, Agriculture, Fisheries, Forests and Human Health, DOE/ER-0236; Glaciers, Ice Sheets, and Sea Level: Effect of a CO2-induced Climatic Change, DOE/ER/60235-1; Projecting the Climatic Effects of Increasing Carbon Dioxide, DOE/ER-0237. Washington, DC.

Environmental Protection Agency (EPA). 1983. Can We Delay a Greenhouse Warming?: The Effectiveness and Feasibility of Options to Slow a Build-up of Carbon Dioxide in the Atmosphere., Washington, DC: U.S Government Printing Office.

Environmental Protection Agency (EPA) and UNEP. 1986. International Conference on Health and Environmental Effects of Ozone Modification and Climate Change. Washington, DC

Environmental Protection Agency (EPA). 1989. The Potential Effects of Global Climate Change on the United States. Smith, J.B., and D. Tirpak, (eds.). Washington, DC: US EPA, Office of Policy, Planning, and Evaluation.

Environmental Protection Agency (EPA). 1990. Policy Options for Stabilizing Global Climate: Draft Report to Congress. Lashof, Daniel and Dennis Tirpak (eds.) EPA 21P-20003.1. Washington, DC.

Fisher-Vanden, Karen. "International Policy Instrument Prominence in the Climate Change Debate: A Case Study of the United States." Environment and Natural Resources Program Discussion Paper E-97-06, Kennedy School of Government, Harvard University, August 1997 and International Institute for Applied Systems Analysis Interim Report IR-97-033/August.

Franz, Wendy E. "The Development of an International Agenda for Climate Change: Connecting Science to Policy." Environment and Natural Resources Program Discussion Paper E-97-07, Kennedy

School of Government, Harvard University, August 1997 and International Institute for Applied Systems Analysis Interim Report IR-97-034/August.

Intergovernmental Panel on Climate Change (IPCC). 1990a. Climate Change: The IPCC Scientific Assessment. Cambridge; New York: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC). 1990b. Climate Change: The IPCC Impacts Assessment. Report from Working Group II to IPCC. Canberra: Australian Govt. Pub. Service.

Intergovernmental Panel on Climate Change (IPCC). 1990c. Climate change: The IPCC Response Strategies. Report from Working Group III. Washington, DC: Island Press.

Intergovernmental Panel on Climate Change (IPCC). 1992a. Climate Change 1992: The Supplementary Report to the IPCC Scientific Assessment. Cambridge; New York, NY, USA: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC). 1992b Climate Change 1992 : The Supplementary Report to the IPCC Impacts Assessment. Cambridge; New York, NY, USA: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC). 1994a. Guidelines for National Greenhouse Gas Inventories. Cambridge; New York, NY, USA: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC). 1994b. IPCC Technical Guidelines for Assessing Climate Change Impacts and Adaptations. Cambridge; New York, NY, USA: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC). 1994c. Climate Change 1994 - Radiative Forcing of Climate Change and An Evaluation of the IPCC IS92 Emission Scenarios. Cambridge; New York, NY, USA: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC). 1996a. Climate Change 1995 - The Science of Climate Change. Report of the IPCC Working Group I. Cambridge; New York, NY, USA: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC). 1996b. Climate Change 1995 - Scientific-Technical Analyses of Impacts, Adaptations and Mitigation of Climate Change. Report of the IPCC Working Group II. Cambridge; New York, NY, USA: Cambridge University Press.

Intergovernmental Panel on Climate Change (IPCC). 1996c. Climate Change 1995 - The Economic and Social Dimensions of Climate Change. Report of the IPCC Working Group III. Cambridge; New York, NY, USA: Cambridge University Press.

Jaeger, Jill and Howard Ferguson, eds. 1991. Climate Change: Science, Impacts and Policy. Proceedings of the Second World Climate Conference, Geneva, 1990. Cambridge; New York: Cambridge University Press, 1991.

Kandlikar, Milind and Ambuj Sagar. "Climate Change Science and Policy: Lessons from India." Environment and Natural Resources Program Discussion Paper E-97-08, Kennedy School of Government, Harvard University, August 1997 and International Institute for Applied Systems Analysis Interim Report IR-97-035/August.

Long, Marybeth and Alastair Iles. "Assessing Climate Change Impacts: Co-evolution of Knowledge, Communities, and Methodologies." Environment and Natural Resources Program Discussion Paper

E-97-09, Kennedy School of Government, Harvard University, August 1997 and International Institute for Applied Systems Analysis Interim Report IR-97-036/August.

Mintzer, Irving. 1987. A Matter of Degrees: The Potential for Controlling the Greenhouse Effect. Washington, DC: World Resources Institute.

National Acid Precipitation Assessment Program (NAPAP). 1991. NAPAP 1990 Integrated Assessment. Washington, DC: US Government Printing Office.

National Research Council (NRC). 1977. Energy and Climate. Geophysics Study Committee, Geophysics Research Board, and Assembly of Mathematical and Physical Sciences. Washington, DC: National Academy Press.

National Research Council (NRC). 1979. Carbon Dioxide and Climate: A Scientific Assessment. Report of an Ad Hoc Study Group on Carbon Dioxide and Climate, Woods Hole, MA, July 23-27, 1979. Assembly of Mathematical and Physical Sciences, Climate Research Board. Washington, DC: National Academy Press.

National Research Council (NRC). 1983. Changing Climate. Report of the Carbon Dioxide Assessment Committee, Board on Atmospheric Sciences and Climate. Washington, DC: National Academy Press.

National Research Council (NRC). 1992. Policy Implications of Greenhouse Warming. Washington, DC: National Academy Press.

Office of Technology Assessment (OTA). 1991. Changing by Degrees: Steps to Reduce Greenhouse Gases, OTA-0-482. Washington, DC: US Government Printing Office.

Office of Technology Assessment (OTA). 1993. Preparing for an Uncertain Climate, OTA-0-568. Washington, DC: US Government Printing Office.

Parris, Thomas, Charles Zracket and William C. Clark. "Usable Knowledge for Managing Responses to Global Environmental Change: Recommendations to Promote Collaborative Assessments and Information Systems." Environment and Natural Resources Program Discussion Paper E-97-11, Kennedy School of Government, Harvard University, August 1997 and International Institute for Applied Systems Analysis Interim Report IR-97-038/August.

Patt, Anthony G. "Assessing Extreme Outcomes: The Strategic Treatment of Low Probability Impacts of Climate Change." Environment and Natural Resources Program Discussion Paper E-97-10, Kennedy School of Government, Harvard University, August 1997 and International Institute for Applied Systems Analysis Interim Report IR-97-037/August.

Scientific Committee on Problems of the Environment (SCOPE), 1986, The Greenhouse Effect, Climatic Change, and Ecosystems. (SCOPE 29). Bert Bolin et al. (eds.) (Wiley, Chichester).

Stockholm Environmental Institute (SEI) 1990, [four volumes]. Responding to Climate Change: Tools for Policy Development; Options for Reducing Greenhouse Gas Emissions; Targets and Indicators of Climate Change; Usable Knowledge for Managing Global Climate Change Stockholm: Stockholm Environment Institute.

United Nations Environment Program (UNEP), WMO, ICSU. 1985. An Assessment of the Role of Carbon Dioxide and of other Greenhouse Gases in Climate Variation and Associated Impacts. Villach, Austria.

World Meteorological Organization. 1979. Proceedings of the World Climate Conference: A

Conference of Experts on Climate and Mankind, Geneva, 12 -23 February 1979. WMO (Series); no.537. Geneva: Secretariat of the World Meteorological Organization.

World Meteorological Organization. 1986. Atmospheric Ozone 1985. Global Ozone Research and Monitoring Project Report No. 16. Geneva: World Meteorological Organization.